



Kent Academic Repository

Doko, Denisa (2012) *Biophysical characterization and NMR analysis of the PDI fragment b'xa'c*. Doctor of Philosophy (PhD) thesis, University of Kent.

Downloaded from

<https://kar.kent.ac.uk/94313/> The University of Kent's Academic Repository KAR

The version of record is available from

<https://doi.org/10.22024/UniKent/01.02.94313>

This document version

UNSPECIFIED

DOI for this version

Licence for this version

CC BY-NC-ND (Attribution-NonCommercial-NoDerivatives)

Additional information

This thesis has been digitised by EThOS, the British Library digitisation service, for purposes of preservation and dissemination. It was uploaded to KAR on 25 April 2022 in order to hold its content and record within University of Kent systems. It is available Open Access using a Creative Commons Attribution, Non-commercial, No Derivatives (<https://creativecommons.org/licenses/by-nc-nd/4.0/>) licence so that the thesis and its author, can benefit from opportunities for increased readership and citation. This was done in line with University of Kent policies (<https://www.kent.ac.uk/is/strategy/docs/Kent%20Open%20Access%20policy.pdf>). If you ...

Versions of research works

Versions of Record

If this version is the version of record, it is the same as the published version available on the publisher's web site. Cite as the published version.

Author Accepted Manuscripts

If this document is identified as the Author Accepted Manuscript it is the version after peer review but before type setting, copy editing or publisher branding. Cite as Surname, Initial. (Year) 'Title of article'. To be published in **Title of Journal**, Volume and issue numbers [peer-reviewed accepted version]. Available at: DOI or URL (Accessed: date).

Enquiries

If you have questions about this document contact ResearchSupport@kent.ac.uk. Please include the URL of the record in KAR. If you believe that your, or a third party's rights have been compromised through this document please see our [Take Down policy](https://www.kent.ac.uk/guides/kar-the-kent-academic-repository#policies) (available from <https://www.kent.ac.uk/guides/kar-the-kent-academic-repository#policies>).

Biophysical characterization and NMR analysis of the PDI fragment b'xa'c

Denisa Doko

PhD Biochemistry 2012

A thesis submitted to the University of Kent for the degree of PhD in Biochemistry at the
School of Biosciences, Faculty of Science, Technology and Medicine.

University of
Kent

Declaration

No part of this thesis has been submitted in support of an application for any degree or qualification of the University of Kent or any other University or institute of learning.

Denisa Doko
September 2012

A. Acknowledgements

Firstly, I would like to thank my supervisors **Dr. Richard Williamson** and **Dr. Mark Howard** for all their help, guidance, support, time and patience throughout my PhD and thesis writing. I am very grateful to **Dr. Michelle Rowe** for assistance with running and analysing NMR data, as well as many software-related issues. I would also like to thank **Dr. Katrine Wallis** of the University of Warwick, who supplied me with DNA for various PDI constructs, **Prof. Lloyd Ruddock** for the generation of PDI constructs, as well as **Mr. Kevin Howland** of the protein research facility for help with mass spectrometry with fluorescence analysis.

Special thanks go to past and present members of the NMR group, especially **Dr. Jane Wagstaff** and **Samantha Taylor**, for continuous support, encouragement, endless science-related conversations and much needed tea breaks.

All my love and gratitude go to my family for their continued support and guidance throughout my life and studies. I would not have come this far in life without them! Finally, my deepest thanks go to Eris for all the love and encouragement that have kept me going over the years.

B. Contents

A. Acknowledgements	i
B. Contents	ii
C. Abbreviations	ix
D. Abstract	xi

Chapter1 – Introduction

1.1 Overview	1
1.2 Protein folding in the ER	2
1.3 PDI (Protein Disulphide Isomerase)	3
1.3.1 PDI Structure	3
1.3.2 Activity	13
1.3.3 Mechanism	14
1.3.4 Ligand binding	17
1.4 Human PDI family members	19
1.4.1 PDIp	19
1.4.2 ERp57	20
1.4.3 ERp72	21
1.4.4 P5	22
1.4.5 ERp44	22
1.4.6 PDIr	23
1.4.7 Other PDI family members	23
1.5 Introduction to Biomolecular NMR	25
1.6 Project Aims	27

Chapter 2 – Mutagenesis, Protein Production and Purification of b'xa'c

2.1 Introduction	28
2.1.1 <i>E. coli</i> strains and vectors use for protein expression	28
2.1.2 Mutagenesis by overlap extension	30
2.1.3 Protein expression and purification	34
2.2 Materials and Methods	35
2.2.1 Mutagenesis by overlap extension	35
2.2.1.1 Bacterial strains	35
2.2.1.2 Luria Bertani Media	35
2.2.1.3 SOB Media	35
2.2.1.4 Oligonucleotides	35
2.2.1.5 Preparation of <i>E. coli</i> competent cells for DNA transformation	36

2.2.1.6 DNA transformation of <i>E. coli</i> cells	36
2.2.1.7 Preparation of <i>E. coli</i> glycerol stocks	37
2.2.1.8 Preparation of template DNA	37
2.2.1.9 Digestion of DNA with restriction endonucleases	37
2.2.1.10 Agarose gel electrophoresis of DNA	38
2.2.1.11 DNA purification from agarose gels	38
2.2.1.12 PCR amplification	39
2.2.1.13 Splice by overlap extension PCR	39
2.2.1.14 Ligation of inserts into empty vector	40
2.2.1.15 Screening	41
2.2.1.16 DNA sequencing	41
2.2.2 Protein expression and purification techniques	42
2.2.2.1 Minimal Media	42
2.2.2.2 Recombinant protein expression in <i>E. coli</i>	42
2.2.2.3 Lysis	43
2.2.2.4 Metal Affinity Chromatography	43
2.2.2.5 Dialysis	44
2.2.2.6 Anion-exchange Chromatography	44
2.2.2.7 Gel Filtration	44
2.2.2.8 Mass Determination	45
2.2.2.9 SDS-PAGE analysis	45
2.2.2.10 Native PAGE analysis	46
2.2.2.11 Determination of protein concentration	47
2.3 Results	48
2.3.1 Preparation of vector DNA	48
2.3.2 Generation of PCR products	49
2.3.3 Generation of mutant DNA	51
2.3.4 Protein expression	51
2.3.5 Metal Affinity Chromatography	52
2.3.6 Anion-exchange Chromatography	53
2.3.7 Gel Filtration	54
2.3.8 Molecular weight confirmation by Mass Spectrometry	55
2.3.9 Characterization of WT b'xa'c and mutants by gel filtration	58
2.3.10 Final protein yields	61
2.4 Discussion	62
 Chapter 3 – Biophysical characterization of WT b'xa'c and mutants by intrinsic fluorescence and limited proteolysis	
3.1 Introduction	66
3.1.1 Intrinsic fluorescence	66
3.1.2 Determination of conformations of PDI by intrinsic fluorescence	68

3.1.3 Determination of the conformational stability of b'xa'c	71
3.1.4 Determination of the conformational flexibility of b'xa'c	71
3.2 Materials and Methods	73
3.2.1 Sample preparation	73
3.2.2 Fluorescence studies to determine the conformation of x	73
3.2.2.1 Determination of conformational stability by GdmCl Denaturation	74
3.2.2.2 Determination of the free energy of denaturation ΔG^0	74
3.2.3 Limited Proteolysis	76
3.3 Results	77
3.3.1 Determination of capping in b'xa'c by fluorescence	77
3.3.2 Determination of conformational stability of PDI constructs	81
3.3.3 Determination of conformational flexibility by limited proteolysis	89
3.4 Discussion	92
3.4.1 Determination of capping in b'xa'c by fluorescence	92
3.4.2 Determination of conformational stability of PDI constructs	93
3.4.3 Determination of conformational flexibility by limited proteolysis	95
 Chapter 4 – Characterization of b'x and xa'c by NMR Spectroscopy	
4.1 Introduction	96
4.1.1 $^{15}\text{N}/^1\text{H}$ HSQC experiments	97
4.1.2 Sequential backbone assignments by Triple Resonance	97
4.1.3 Chemical Shift Mapping	99
4.1.4 Secondary Structure Prediction from chemical shifts	100
4.1.5 ^{15}N Backbone NMR Dynamics	101
4.1.6 Model-free formalism	104
4.1.7 Spectral Density and Spectral Density Mapping	106
4.2 Materials and Methods	108
4.2.1 NMR sample preparation	108
4.2.2 NMR data acquisition and processing	108
4.2.2.1 $^{15}\text{N}/^1\text{H}$ HSQC experiments of PDI constructs	108
4.2.2.2 Amino acid sequential backbone resonance assignment	109
4.2.2.3 ^{15}N NMR Relaxation	109
4.2.3 NMR data analysis	109
4.2.3.1 Triple Resonance assignment	109
4.2.3.2 Chemical Shift Mapping	109
4.2.3.3 Secondary Structure Prediction using DANGLE	110
4.2.3.4 NMR Relaxation Analysis	111
4.2.4 NMR temperature and pH titrations	112
4.2.5 Ligand binding using Δ -somatostatin	113
4.3 Results	114

4.3.1 Expression and purification of b'x , WT and I272A, and xa'c	114
4.3.2 The effect of temperature and pH on PDI constructs	115
4.3.3 Sequential backbone assignment of reduced xa'c	121
4.3.4 Assignment of WT and I272A b'x	122
4.3.5 Secondary Structure Prediction using DANGLE	124
4.3.6 Ligand binding	128
4.3.7 ¹⁵ N NMR Relaxation Dynamics of xa'c and b'x	129
4.3.7.1 ¹⁵ N T ₁ , T ₂ and hetNOE data for xa'c	129
4.3.7.2 ¹⁵ N T ₁ , T ₂ and hetNOE data for b'x	131
4.3.7.3 ¹⁵ N T ₁ versus T ₂ relaxation data for xa'c and b'x	132
4.3.7.4 ¹⁵ N Model-free analysis of xa'c and b'x	134
4.4 Discussion	136
4.4.1 The conformational stability of the a' domain	136
4.4.2 Sequential backbone assignment of xa'c	137
4.4.3 Secondary Structure Prediction using DANGLE	138
4.4.4 Ligand binding	138
4.4.5 ¹⁵ N NMR Relaxation Dynamics of xa'c and b'x	139
4.4.5.1 Relaxation data	139
4.4.5.2 Model-free analysis	142
Chapter 5 – Characterization of b'xa'c by NMR Spectroscopy	
5.1 Introduction	146
5.1.1 NMR and conformational exchange	147
5.1.2 Evidence of capping in PDI proteins	149
5.1.3 Sequential backbone assignment of reduced b'xa'c	151
5.1.4 The redox-dependent conformational changes of bb'xa' and abb'xa'	152
5.2 Materials and Methods	154
5.2.1 NMR sample preparation	154
5.2.2 NMR data acquisition and processing	154
5.2.2.1 Amino acid sequential backbone assignment	154
5.2.2.2 ¹⁵ N NMR Relaxation	154
5.2.3 NMR data analysis	155
5.2.3.1 Triple Resonance experiments HNCA and HN(CO)CA	155
5.2.3.2 NMR temperature titrations	155
5.2.3.3 NMR pH titrations	155
5.2.3.4 Ligand binding using Δ-somatostatin	156
5.2.3.5 Secondary Structure Prediction using DANGLE	156
5.2.3.6 NMR Relaxation Analysis	156
5.2.3.7 Minimal Shift Mapping	156
5.3 Results	157

5.3.1 Initial NMR investigations on WT b'xa'c and mutants	157
5.3.2 The effect of temperature and pH on WT b'xa'c and mutants	161
5.3.3 Assignment of I272A b'xa'c and mapping assignments on WT b'xa'c	165
5.3.4 The effect of the mutations on b'xa'c	168
5.3.5 Determination of capping in b'xa'c by NMR spectroscopy	172
5.3.6 Secondary Structure Prediction using DANGLE	173
5.3.7 Ligand binding	176
5.3.8 Relaxation Dynamics	183
5.3.8.1 ¹⁵ N T ₁ , T ₂ and hetNOE data for WT and I272A b'xa'c	183
5.3.8.2 ¹⁵ N T ₁ versus T ₂ relaxation data for WT and I272A b'xa'c	185
5.3.8.3 Model-free analysis of WT and I272A b'xa'c	187
5.4 Discussion	191
5.4.1 The alternative conformations of b'xa'c studied by NMR	191
5.4.2 Assignment of I272A b'xa'c and mapping assignments on WT b'xa'c	194
5.4.3 Secondary Structure Prediction using DANGLE	195
5.4.4 Ligand binding	196
5.4.5 ¹⁵ N NMR Relaxation Dynamics of WT and I272A b'xa'c	197
4.4.5.1 Relaxation data	197
4.4.5.2 Model-free analysis	199
Chapter 6 – General Summary	201
References	207
Appendix Contents	215
Appendices	216

C. Abbreviations

^1H	proton (magnetic nuclei: 100 % natural abundance)
^2H	deuteron (0.015% natural abundance)
^{13}C	carbon isotope-13 (magnetic nuclei: 1.11% natural abundance)
^{15}N	nitrogen isotope-15 (magnetic nuclei: 0.36% natural abundance)
2D	two dimensional
3D	three dimensional
γ_{H}	proton gyromagnetic ratio
γ_{N}	nitrogen gyromagnetic ratio
τ_{e}	effective correlation time
τ_{m}	global correlation time
ω_{H}	proton larmor frequency
ω_{N}	nitrogen larmor frequency
φ	phi angle
ψ	psi angle
μ_0	permeability of free space
\hbar	Plank's constant
B_0	applied field
CCPN	Collaborative Computing Project for NMR
CPMG	Carr-Purcell-Meiboom-Gill
2DA	D346A/D348A
D_2O	deuterium oxide
Da	Dalton
DANGLE	Dihedral ANgles from Global Likelihood Estimates
dH_2O	distilled water
DNA	deoxyribonucleic acid
DTT	Dithiothreitol
EDTA	ethylenediaminetetraacetic acid
GdmCl	guanidine hydrochloride
hetNOE	heteronuclear Overhausen effect
hPDI	human Protein Disulphide Isomerase

HSQC	Heteronuclear Single Quantum Coherence
IPTG	isopropyl β -D-1-thiogalactopyranoside
$J(\omega)$	spectral density
LB	Luria Bertani
NMR	Nuclear Magnetic Resonance
NOE	nuclear Overhauser effect
PCR	polymerase chain reaction
PDB	protein databank
PDI	Protein Disulphide Isomerase
ppm	parts per million
R_{ex}	rate of chemical exchange
rpm	revolutions per minute
S^2	protein backbone flexibility order parameter
SDS	sodium dodecyl sulfate
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
T_1	longitudinal/ spin-lattice relaxation
T_2	transverse/spin-spin relaxation
UV	ultraviolet
yPDI	yeast Protein Disulphide Isomerase

D. Abstract

Protein disulphide isomerase (PDI) was the first catalyst of protein folding to be identified and is a key cellular chaperone required for the rapid formation of correctly folded disulphide bonded proteins in the secretory pathway. PDI is the archetype of a large family of ER-resident PDI-like proteins and contains four thioredoxin-like domains, two of which, like thioredoxin itself, have redox-active (CGHC) catalytic sites (**a** and **a'**) and two of which do not (**b** and **b'**). The domain order is **abb'xa'c** where **x** is a 19 residue linker between the **b'** and **a'** domains and **c** is a C-terminal acidic tail containing the KDEL ER-retention signal. The **x** linker has been shown to occlude the primary ligand binding site in the human **b'x** construct; an event called “capping”. In previous work, site-directed mutants were identified that stabilise both the capped and uncapped conformations of **b'x**. Furthermore, the binding of **x** was shown to compete with peptide ligands and could prove to be an important physiological mechanism controlling access to the hydrophobic binding site. The movement of the **x** linker in gating the binding site could also lead to wider structural arrangement of the protein and control other aspects of PDI structure and function.

This study focuses on the characterisation of the smallest PDI fragment with isomerase activity, **b'xa'c**, to further understand the role of **x** capping and its interaction with adjacent **b'** and **a'** domains. Mutants of **b'xa'c**, which in **b'x** favoured capping or uncapping of the ligand binding site, were generated by site-directed mutagenesis. All **b'xa'c** proteins presented as a mixture of monomer and dimer, which on gel filtration presented with larger hydrodynamic volumes than expected for globular proteins of similar sizes. Denaturation studies using guanidine hydrochloride showed that the capping I272A mutant of **b'xa'c** presented with a higher conformational stability than WT **b'xa'c**, whereas the uncapping mutant L343A was less conformationally stable than WT and I272A **b'xa'c**. **b'xa'c** proteins presented with a biphasic denaturation curve in which the first phase could be attributed to the unfolding of the **a'** domain and the second phase was due to **b'x** unfolding.

NMR studies of **b'xa'c** proved challenging due to broad and poorly resolved spectra. Studies on **b'x** and **x'ac** were carried out as a stepping stone to the investigations of **b'xa'c**. The oxidation state of the **a'** domain, temperature and pH had a significant effect on the line widths seen for **b'xa'c**, and allowed the backbone resonance assignment of **xa'c** and I272A **b'xa'c** to be carried out. Relaxation dynamics studies showed that the **a'** domain has a more flexible backbone than **b'x** and that neighbouring domains affect the conformational behaviour of one another.

NMR and intrinsic fluorescence showed no conclusive evidence of capping of the ligand binding site by the **x** linker region in **b'xa'c**, but mutants I272A, L343A and D346A/D348A had a significant effect on the conformational stability of **b'xa'c**.

CHAPTER 1

Introduction

1.1 Overview

Disulphide bonds are covalent linkages formed between the side chains of cysteine residues. They are an integral part of many of life's essential proteins as they not only help stabilise protein structure, but they are also vital to their function. It is crucial that during the formation of disulphide bonds, the correct cysteine residues are paired in order to form the desired three-dimensional (3D) structure. Protein disulphide isomerase (PDI) is a central player in this process as it is involved in the formation (oxidation), breakage (reduction) and rearrangement (isomerisation) of disulphide bonds and therefore helps newly expressed proteins assume their correct 3D structure. In addition to its role in disulphide bond formation, PDI is also the β subunit of prolyl-4-hydroxylase (Koivu et al., 1987) and microsomal triglyceride transfer protein (Wetterau et al., 1991).

PDI was the first folding catalyst to be identified (Goldberger et al., 1963) and is a member of a large family of thiol-disulphide oxidoreductants found within the endoplasmic reticulum (ER) of eukaryotic cells (Hatahet and Ruddock, 2009) which contain the C-terminal KDEL-ER retention sequence (Pelham, 1990). PDI and its related proteins are members of the thioredoxin superfamily, as they all contain at least one domain with the thioredoxin-like fold and often the thioredoxin-like active site with the

amino acid sequence CXXC (Edman et al., 1985). This thesis will focus on the characterisation of the minimal human PDI unit with isomerase activity, **b'xa'c** (Pirneskoski et al., 2001).

1.2 Protein folding in the ER

Protein synthesis in eukaryotic cells starts in the nucleus with transcription of specific DNA segments by RNA polymerase to produce messenger RNA (mRNA). mRNA is then translated on the surface of ribosomes in the cytoplasm with the aid of transfer RNA (tRNA) into an amino acid sequence which requires folding into a 3D structure in order to become a functionally active protein. Protein folding occurs soon after the synthesis of the polypeptide chain and proteins destined for the secretory pathway contain a signal sequence, usually at the N-terminus, which directs their translocation into the endoplasmic reticulum (ER) ready for folding. A ribosome binds to the endoplasmic reticulum when it begins the synthesis of proteins destined for the secretory pathway. It is estimated that approximately one third of human proteins fold in the ER (Chen et al., 2005) but it is still not fully understood how the amino acid sequence determines a protein's 3D fold (Dobson, 2003). The lumen of the ER is rich in enzymes that catalyse the oxidation, reduction and isomerisation of proteins containing disulphide bonds, among which PDI is one of the most abundant. Disulphide bond isomerisation is thought to be the key rate limiting step in the biogenesis of most proteins and it is the final steps of isomerisation of nascent proteins that are catalysed by members of the PDI family (Klappa et al., 1995). The most important member in this family of oxidoreductases is PDI itself, a protein found in all eukaryotic organisms.

1.3 PDI (Protein Disulphide Isomerase)

Since its discovery in 1963, PDI (EC 5.3.4.1) has been the subject of extensive studies, but its mechanism of action and structure/function relationship are still not fully understood. It was not until the mid-1980s that the sequence of rat PDI was published (Edman et al., 1985) revealing the initial structural information about PDI. This sequence revealed two distinct regions which were highly homologous to thioredoxin, a small ubiquitous protein which catalyses disulphide bond reduction, but most importantly it proved that PDI catalysed protein folding by oxidation and reduction of disulphide bonds due to the presence of vicinal thiol groups in the characteristic thioredoxin active site motif CGHC.

Human PDI is a 57 kDa multi-domain protein resident in the ER of a variety of tissues and organs. It is a key enzyme and chaperone (Wilson et al., 1998, Wang et al., 2010) implicated in a range of cellular functions such as catalysis of disulphide bond oxidation, reduction and isomerisation, degradation of mis-folded proteins in the ER and the unfolding of the cholera toxin before its retrotranslocation into the cytosol (Tsai and Rapoport, 2002). PDI can also be found in non-ER locations, such as the cell surface where it participates in the activation of fusion of HIV virus (Fenouillet et al., 2007) and mediates the adhesion, secretion and aggregation of platelets (Essex, 2009).

Over the years, a number of excellent reviews have been written discussing the role of PDI in disulphide bond formation, breakage and rearrangement and its implications in protein folding; as well as the structure/function relationship and the impact on human health (Ellgaard and Ruddock, 2005, Freedman et al., 1994, Ferrari and Söling, 1999, Hatahet and Ruddock, 2009). Recently, PDI has been recognized to play a critical role in a number of neurodegenerative diseases related to protein mis-folding and aggregation such as Parkinson disease, Alzheimer disease and Huntington disease (Uehara et al., 2006).

1.3.1 PDI Structure

Although PDI was discovered over 40 years ago, it is only in the last 15 years that the structures for each of the PDI domains have emerged and to date a structure for full length human PDI (hPDI) is yet to be solved. However, recent years have seen a number

of successful structural studies with the most exciting outcome of decades of research being the publication of two crystal structures of yeast PDI (yPDI) (Tian et al., 2006, Tian et al., 2008). Also, very recently, two crystal structures of oxidised and reduced hPDI (4EL1.pdb and 4EKZ.pdb respectively) lacking the C-terminal extension, **c**, are currently in press but the coordinates are on hold in the Protein Data Bank (PDB) until April 2013. The structure of **bb'xa'** of hPDI (3UEM.pdb) was also solved in 2011 and is to date the only available structure of hPDI consisting of both the ligand binding domain, **b'**, and the catalytic **a'** domain. Table 1.1 summarises the structures of PDI single and multi-domain constructs that have been solved by a mixture of NMR and X-ray crystallography.

PDI domain	PDB code	Organism	Method	Reference
a	1MEK	<i>H. sapiens</i>	NMR	Kemmink et al., 1996
b	1BJX	<i>H. sapiens</i>	NMR	Kemmink et al., 1999
b'x	3BJ5	<i>H. sapiens</i>	X-ray	Nguyen et al., 2008
a'	1X5C	<i>H. sapiens</i>	NMR	RSGL., 2005 (not published)
bb'	2K18	<i>H. sapiens</i>	NMR	Denisov et al., 2009
bb'xa	3UEM	<i>H. sapiens</i>	X-ray	Wang et al., 2011
abb'xa'	4EKZ (red) 4EL1 (oxi)	<i>H. sapiens</i>	X-ray	Wang et al., 2012 (on hold in PDB)
abb'xa'c	3B5E (4°C) 3BOA (22°C)	<i>S. cerevisiae</i>	X-ray	Tian et al., 2006 Tian et al., 2008
b'	2KP2	<i>H. insolens</i>	NMR	Serve et al., 2009
a'	2KP1	<i>H. insolens</i>	NMR	Serve et al., 2009

Table 1.1 PDI constructs for which structures have been solved. The coordinates for **abb'xa'**, oxidised and reduced, are available from the PDB but the structures are on hold until April 2013. The structure of the **a'** domain (1X5C) has not been published despite its deposition in the PDB in 2005 by the Riken Structural Genomics/Proteomics Initiative (RSGL).

Full length hPDI consists of four structural domains, each consisting of a thioredoxin fold, a C-terminal extension, **c**, and a linker region, **x**, connecting the **b'** and **a'** domains. PDI also contains an N-terminal signal sequence which is cleaved during transport of the newly synthesised protein to the ER. Numbering of PDI residues in this

thesis refers to that for the mature PDI after cleavage of the N-terminal signal sequence. The domain arrangement is **abb'xa'c** as shown in figure 1.1.

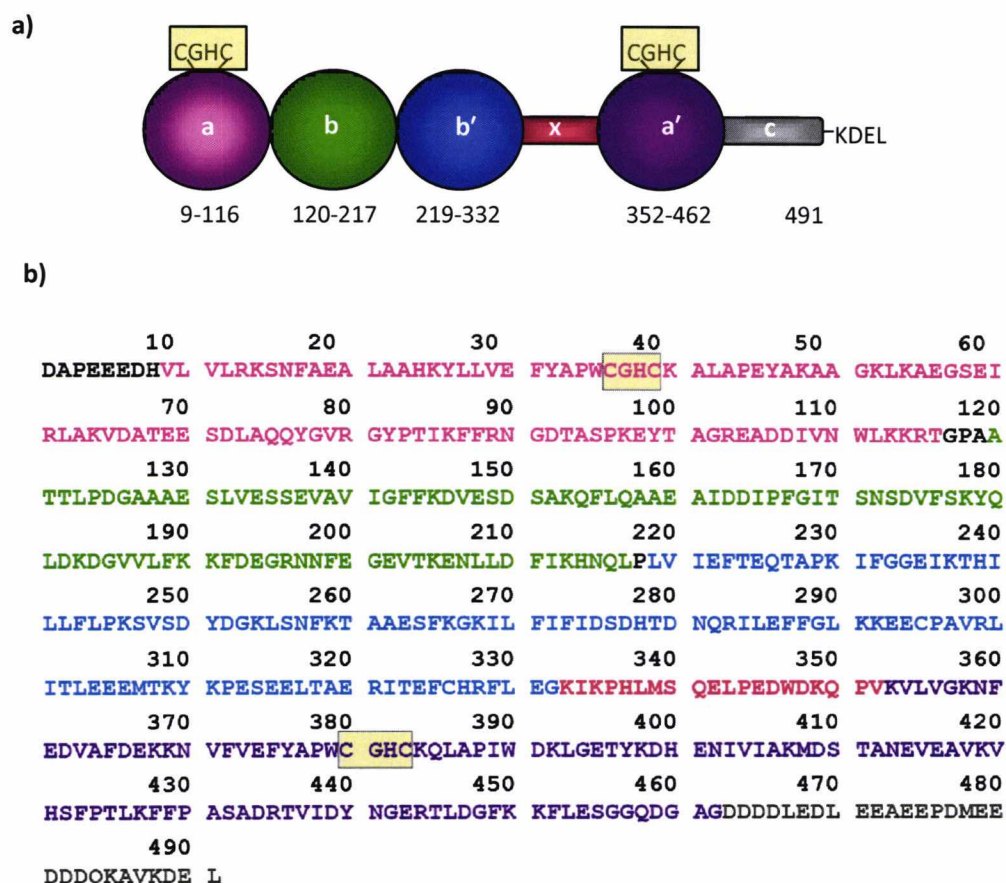


Figure 1.1 The domain architecture and amino acid sequence of mature hPDI. a) the arrangement of the domains and b) the amino acid sequence of mature hPDI. Residue numbering is for mature PDI after cleavage of the signal sequence. The domain boundaries are as reported by Wang *et al* (2010). The active site sequence is shown in the yellow box and the domains are coloured as follows: **a** in pink, **b** in green, **b'** in blue and **a'** in purple. The **x** linker and **c** extension are shown in red and grey, respectively.

The **a** and **a'** domains are also known as the catalytic domains as they contain the thioredoxin-like active site motif CGHC, whereas **b** and **b'** are the non-catalytic domains as although they both consist of a thioredoxin fold, they lack the active site sequence. The **b** domain, previously thought to play a structural stabilising role, has recently been reported to form part of the hydrophobic binding surface whereas, the **b'** domain has been shown to contain the primary ligand binding site (Byrne *et al.*, 2009, Pirneskoski *et al.*,

2001, Wang et al., 2012a). **b'** contains two cysteine residues (C295 and C326) but they are unlikely to form disulphide bonds as they have been found to be inaccessible to solvent (Darby et al., 1998, Pirneskoski et al., 2004). The **c** tail located at the C-terminus of PDI is rich in acidic residues and contains the ER-retention signal sequence, KDEL. Additionally, **x** is a flexible linker region located between the **b'** and **a'** domains (Pirneskoski et al., 2004, Tian et al., 2006). Domain boundaries have been determined by studying structures of individual domains, protease sensitivity, as well as looking at the solubility of recombinantly expressed domains (Yao et al., 1997, Ferrari and Söling, 1999, Sun et al., 2000).

For many years crystallisation of full length hPDI has proved challenging, so structural studies have been carried out on individual domains, or more recently double and triple domain constructs. The published structures of yeast PDI have aided in better understanding the domain rearrangement of hPDI, but there is only 37% sequence identity between the two species and therefore there are likely to be important structural differences. Figure 1.2 shows a sequence alignment between human and yeast PDI obtained by submitting the amino acid sequences of the proteins in the ClustalW2 tool in the EMBL-EBI website (<http://www.ebi.ac.uk/Tools/msa/clustalw2/>).

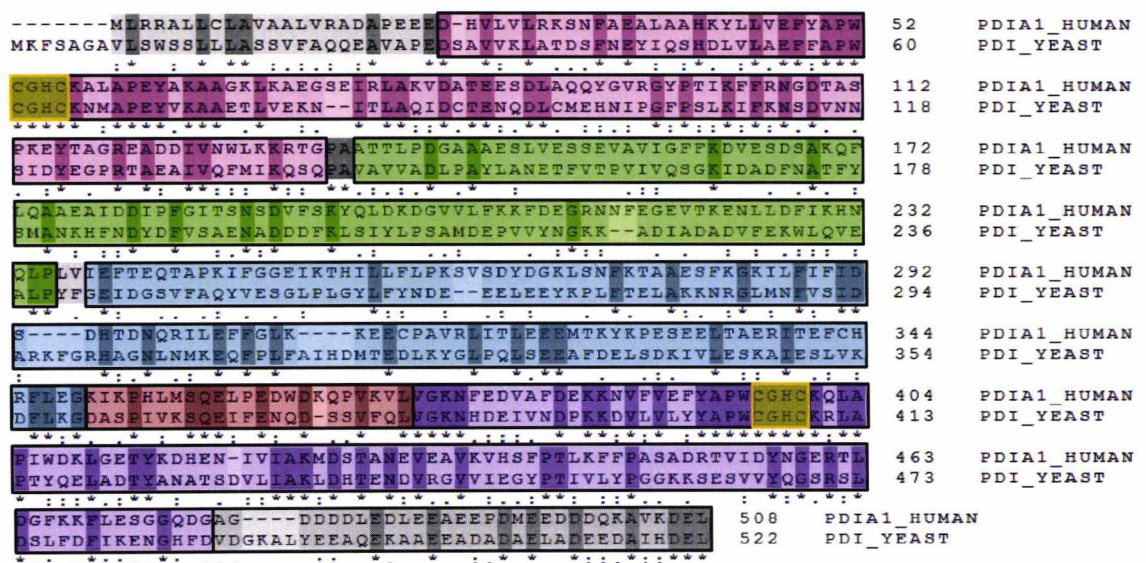


Figure 1.2 Sequence alignment of human and yeast PDI. Alignment was carried out using ClustalW2 in the EMBL-EBI website. Domain boundaries have been coloured as described in figure 1.1 with the active sites in the yellow boxes. Residues that are identical in both sequence are highlighted.

The first human PDI structure to be solved was that of the **a** domain (1MEK.pdb) (Kemink et al., 1996), which revealed the highly conserved thioredoxin fold comprising the $\beta 1$ - $\alpha 1$ - $\beta 2$ - $\alpha 2$ - $\beta 3$ - $\alpha 3$ - $\beta 4$ - $\beta 5$ - $\alpha 4$ structure, forming a five stranded β -sheet flanked by four α -helices (Ferrari and Söling, 1999). The active site thioredoxin motif is located between $\beta 2$ and $\alpha 2$. Sequence homology between the **a** and **a'** domain led to the hypothesis that **a'** also consisted of a thioredoxin fold. This was confirmed by the subsequent deposition of the NMR structure for the **a'** domain (1X5C.pdb) in the PDB by Riken Structural Genomics/Proteomics Initiative (RSGI) in 2005. Despite the low sequence identity between the catalytic, **a** and **a'**, and non-catalytic, **b** and **b'**, domains and the lack of the thioredoxin-like active site motif in the latter, **b** and **b'** were also shown to consist of thioredoxin folds as revealed by the NMR structure of the **b** domain (1BJX.pdb) and the crystal structure of **b'x** (3BJ5.pdb) (Nguyen et al., 2008, Kemink et al., 1999). The structures for all of the individual human PDI domains are shown in figure 1.3 against the human thioredoxin structure (1W4V.pdb) to illustrate the similarity of the fold in all cases.

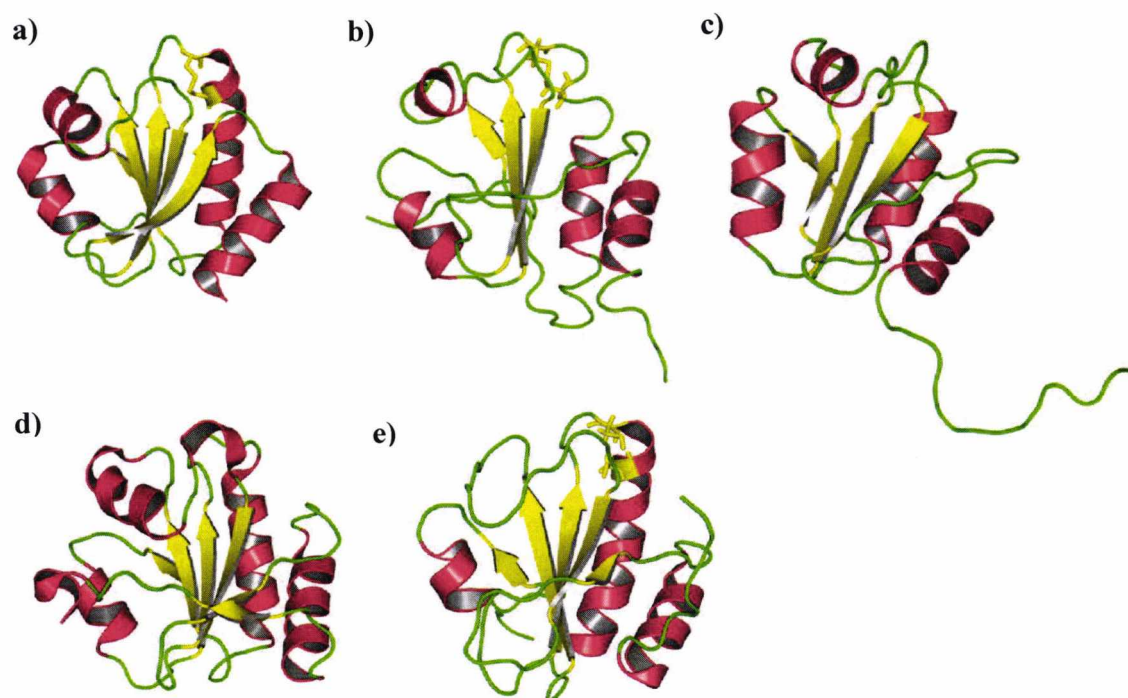


Figure 1.3 Structures of individual PDI domains deposited in the Protein Data Bank. a) human thioredoxin 1W4V.pdb is shown for fold comparison b) human **a** domain 1MEK.pdb c) human **b** domain 1BJX.pdb d) human **b'x** domain 3BJ5.pdb e) human **a'** domain 1X5C.pdb. The cysteines of the active site in thioredoxin, PDI **a** and **a'** domains are shown as yellow sticks. All of the above structures have been solved by NMR spectroscopy except **b'x** which is a crystal structure of the I272A mutant.

The first crystal structure of full length PDI from yeast (3B5E.pdb) revealed that the four thioredoxin domains are arranged in the shape of a twisted “U” with the non-catalytic **b** and **b'** domains forming the base of the “U” and the **a** and **a'** domains forming the arms of the “U” as shown in figure 1.4 (Tian et al., 2006).

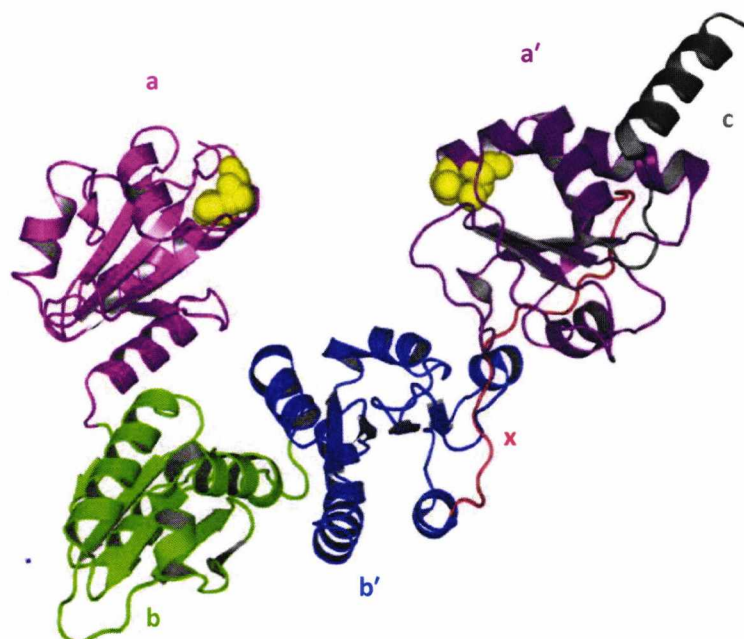


Figure 1.4 The crystal structures of full length yeast PDI from crystals grown at 4°C. Domains are shown as **a** in pink, **b** in green, **b'** in blue, **x** region in red, **a'** in purple and **c** tail in grey. Cysteines of the active sites are shown as Corey-Pauling-Koltun (CPK) space fill in yellow.

Figure 1.4 shows the active sites on **a** and **a'** face each other from a distance of ~ 28 Å (Tian et al., 2006). A hydrophobic patch on **b'** together with other hydrophobic residues around the **a** and **a'** active sites form a continuous hydrophobic surface inside the cleft of the “U” facilitating interactions with mis-folded proteins. All four domains of yeast PDI were found to be structurally similar to thioredoxin and the C-terminal extension, **c**, was partially ordered into a helix. The second α helix in the **a** and **a'** domains, the N-terminal end of which contains the active site, is the longest helix in the fold extending from the end of one parallel strand ($\beta 2$) to the next ($\beta 3$). This is facilitated by the presence of a conserved proline residue in both of the catalytic domains, which introduces a kink in the first half of the helix. The large interface between **b** and **b'** helps to form a rigid base to the “U” with a buried surface area of ~ 700 Å². In contrast, the interfaces between **a** and **b**, and **b'** and **a'** are considerably smaller with buried surface area of ~ 200 Å² allowing the

arms of the “U”-shaped molecule to be much more flexible than the base. This flexibility would also allow moderation of the hydrophobic cleft to accommodate substrates of varying sizes.

A second yeast structure solved from crystals grown at 22°C (3BOA.pdb), in contrast to the structure grown from crystals at 4°C, revealed PDI in a more open conformation forming a “boat” like structure as shown in figure 1.5 (Tian et al., 2008).

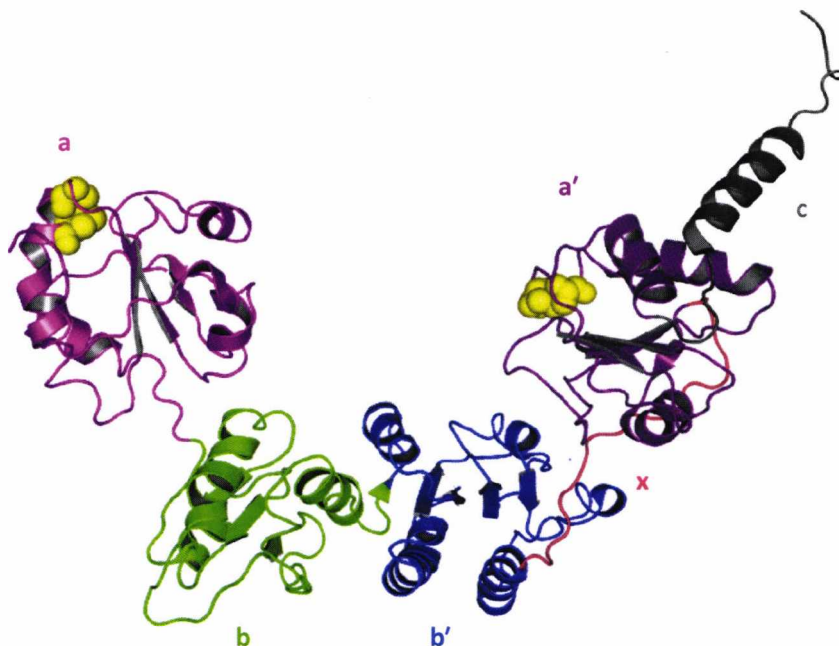


Figure 1.5 The crystal structures of full length yeast PDI from crystals grown at 22°C. Domains and active sites are shown as described in figure 1.3.

The spatial arrangement of the **b**, **b'** and **a'** domains is similar between the two structures, whereas the relative position of the **a** domain is considerably different in the 22°C structure, giving the whole molecule a more extended conformation. The most important feature of this second structure is that the active sites on **a** and **a'** no longer face each other as the **a** domain is dislocated by ~22 Å. This suggests that PDI is a highly flexible molecule that is able to adjust to substrates of different sizes and adopt a range of conformations (Tian et al., 2008).

Although PDI is highly conserved from yeast to mammals, with around 70% sequence homology, the full length structure of human PDI can be predicted to a certain extent, but there are likely to be important differences in the domain orientation between yeast and human PDI structures. In contrast to yPDI, where the flexibility of the molecule

is due to movement of the **a** domain, the C-terminal half of hPDI, **b'xa'c**, is more flexible than the N-terminal **ab** region (Wang et al., 2010). The NMR structure of the **bb'** construct (2K18.pdb) of PDI revealed extensive contacts between the **b** and **b'** domains forming a rigid structural unit as shown in figure 1.6a (Denisov et al., 2009).

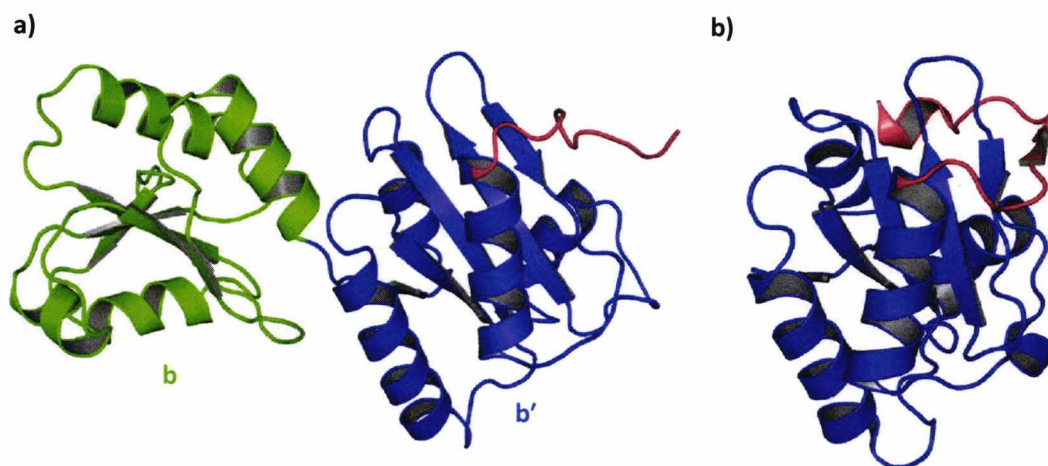


Figure 1.6 Structures of **bb' (2K18.pdb) and **b'x** (3BJ5) of hPDI.** **a)** NMR structure of **bb'** with the **b** and **b'** domains displayed in green and blue, respectively. Only part of the **x** linker region, shown in red, is present in this construct. **b)** Crystal structure of the I272A **b'x** mutant with the **x** linker folding back on the **b'** domain is shown for comparison.

The main structural difference between the human and yeast PDI **bb'** fragment is the extra helix, α_3 , in the **b** domain of hPDI and an extra helix in the **b'** domain of yeast PDI (Denisov et al., 2009). When the **b'** domain in the **bb'** structure was compared to the crystal structure of the I272A mutant of the **b'x** fragment of hPDI (figure 1.6b), differences were seen for a number of amino acids (Denisov et al., 2009). This could have been due to the presence of the **b** domain in **bb'**, the I272A mutation or the presence of the **x** linker region in **b'x**. The **x** linker in the **b'x** structure binds to the ligand binding site in the **b'** domain (Nguyen et al., 2008), whereas in the **bb'** structure, only half of the **x** linker is present and this could affect the conformation of the **b'** domain.

The **bb'xa'** structure of hPDI revealed a conformationally active molecule which was greatly affected by the oxidation state of the **a'** domain (Wang et al., 2012b). The reduced crystal structure showed the **a'** domain tightly packed with **bb'** to form a compact structural unit joined by the **x** linker region (figure 1.7).

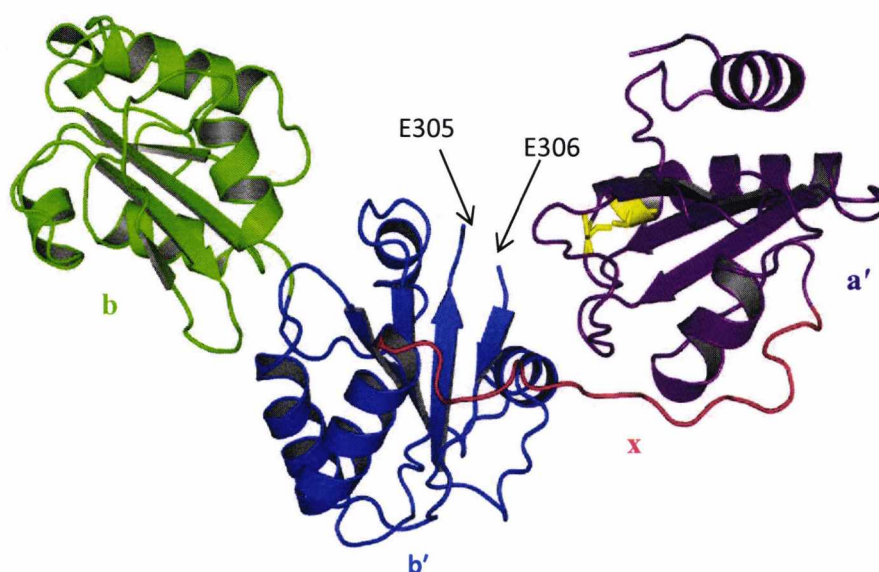


Figure 1.7 Crystal structure of the reduced **bb'xa'** of hPDI. The **b**, **b'** and **a'** domains are coloured in green, blue and purple, respectively. The **x** linker region is shown in red and the cysteines of the **a'** active site are displayed as yellow sticks. Residues E305 and E306 are missing from the structure.

The reduced **bb'xa'** crystal structure is the only available structure which allows the **b'xa'** region to be studied in detail. Superposition of the **bb'xa'** (3UEM.pdb) structure with the solution structure of the **bb'** fragment (2K18).pdb, showed that although the **b** and **b'** domains in the **bb'xa'** structure are similar to those in the **bb'** structure, inter-domain motions were visible between the two structures. The **a'** domain in the **bb'xa'** structure is similar to the isolated **a'** crystal structure (1X5C.pdb). The crystal structure of I272A **b'x** (3BJ5.pdb) was also very similar to the **b'** domain of the **bb'xa'** structure, but differences were seen in the **x** linker region, which in **b'x** is found capping the ligand binding site. Figure 1.8 shows a superposition of the individual **bb'**, **a'** and I272A **b'x** structures with the crystal structure of **bb'xa'**.

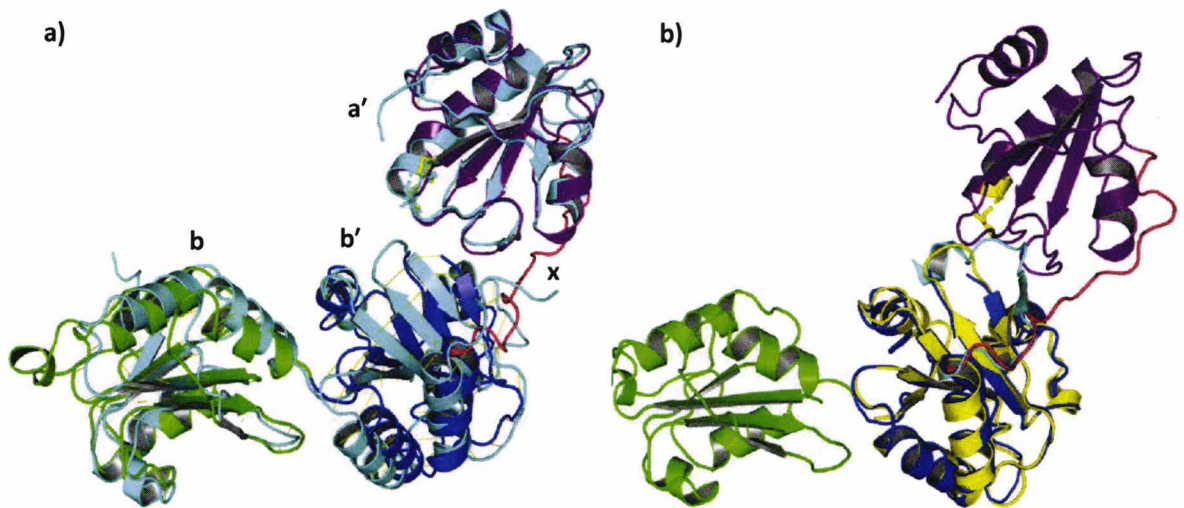


Figure 1.8 Superposition of bb' , a' and I272A $b'x$ with $bb'xa'$. Domains of the $bb'xa'$ structure b , b' , and a' are shown in green, blue and purple respectively. The x linker is shown in red. **a)** bb' (2K18.pdb) and a' (1X5C.pdb) structures are shown in cyan and are superimposed on their respective domains of the $bb'xa'$ structure. **b)** I272A $b'x$ with b' in yellow and x in cyan is aligned with the b' domain of $bb'xa'$. Structure alignment was carried out in PyMOL.

The $bb'xa'$ crystal structure also showed significant differences from both of the yeast crystal structures, especially in the position of the a' domain relative to the bb' base (Wang et al., 2012b). This is due to a $\sim 35 - 45^\circ$ rotation of the a' domain toward the hydrophobic surface of the b' domain. Although there is no structure for $bb'xa'$ in the oxidised state, Wang et al. (2011) showed through protease sensitivity and fluorescence studies that oxidised $bb'xa'$ undergoes redox-dependent conformational changes triggered by a redox switch of the active site in the a' domain (Wang et al., 2012b).

Recent structures of oxidised and reduced $abb'xa'$ of PDI, lacking the C-terminal extension, revealed the four thioredoxin domains arranged in a horseshoe shape with bb' forming the base and the catalytic a and a' domains at the two ends. The reduced structure shows the a , b and b' domains located in the same plane, whereas a' is twisted out of the abb' plane by $\sim 45^\circ$. The $b'xa'$ region of $abb'xa'$ is very similar to that of the $bb'xa'$ structure in the reduced state, and differences are seen only in the b domain, possibly due to contacts between domains a and b in the $abb'xa'$ structure. Comparison of the hPDI $abb'xa'$ with the yPDI structure at 4°C showed that the structures were very similar to each other. However, the catalytic a and a' domains of hPDI displayed rotations of $\sim 30^\circ$ and

$\sim 20^\circ$, respectively, with the **a'** active site moving closer towards the **b'** hydrophobic ligand binding site.

The oxidised **abb'xa'** structure also forms a horseshoe shape, but in contrast to the reduced structure, the thioredoxin domains are arranged in the same plane. The distance between the active sites increases from 27.6 Å in the reduced to 40.3 Å in the oxidised state. But the most important difference between the reduced and oxidised **abb'xa'** hPDI structures was in the **b'xa'** region, where the oxidised **a'** domain rotates $\sim 45^\circ$ around the **x** linker and away from the **b'** ligand binding site, forming a more extended structure than the compact reduced molecule. The redox-dependent conformational change of hPDI ultimately controls the access to the hydrophobic surface inside the horseshoe molecule, which consists of a hydrophobic pocket in the **b'** domain and extended hydrophobic patches around the active sites and the **b** domain. As the oxidised structure displays a more open conformation, revealing a cleft that is ~ 2 fold larger than that of reduced **abb'xa'**, oxidised hPDI can accommodate larger substrates than reduced PDI, as a result controlling the proteins chaperone activity (Wang et al., 2012a).

1.3.2 Activity

PDI displays thioredoxin-like activity although there is a difference in redox potential between thioredoxin and the **a** and **a'** domains due to variation in the CXXC active site sequences (Lundström and Holmgren, 1990). The thioredoxin active site, CGPC, is mainly responsible for catalysing the reduction of disulphide bonds, whereas that of PDI, CGHC, is involved mainly in disulphide bond oxidation and isomerisation. In addition to thiol-disulphide exchange catalysis, PDI has been suggested to have molecular chaperone activity, as it aids folding of disulphide bond-containing proteins (Winter et al., 2002, Yao et al., 1997) as well as assisting in the refolding of proteins lacking disulphide bonds (Cai et al., 1994, Song and Wang, 1995) in order to prevent aggregation of unfolded or mis-folded proteins (Freedman et al., 1994). This chaperone activity was previously reported to require full length PDI but fragments like **bb'xa'** have recently been shown to exhibit chaperone activity (Sun et al., 2000, Wang et al., 2012b). A C-terminal truncation, in which the **c** region and the C-terminus of the **a'** domain were deleted, resulted in the loss of chaperone activity (Dai and Wang, 1997). It is also possible that this truncation could

have resulted in loss of ligand binding by the **b'** domain due to destabilisation of **a'** and subsequent loss of chaperone activity.

PDI has also been shown to have antichaperone activity in instances when substrate concentration is too high for its chaperone capacity. In this case, PDI forms inactive, disulphide bonded aggregates that are retained in the ER for degradation (Puig and Gilbert, 1994). PDI has also been shown to be a subunit of larger multi protein complexes, such as its role as the two β -subunits of the ER luminal enzyme prolyl-4-hydroxylase (Freedman, 1989), involved in the catalysis of procollagen pro- α -chain prolyl hydroxylation.

The redox activity of PDI is dependent on the redox potential of the surrounding media. For example, the ER consists of a highly oxidizing environment caused by a higher proportion of oxidised glutathione (GSSG) than its reduced form (GSH) allowing the formation and rearrangement of disulphide bonds (Hwang et al., 1992). In more reducing conditions, such as the cytosol, PDI is able to reduce disulphide bonds of oxidised proteins, but thioredoxin is more efficient in this process than PDI (Noiva, 1999).

In an important study by Darby and co-workers, the contribution of each domain to the activity of PDI was explored using a number of substrates such as two different disulphide bonded BPTI non-native intermediates, a 28 amino acid long peptide based on the BPTI sequence, and the insulin reduction assay (Darby et al., 1998). They showed that all of the domains contributed to the thiol-disulphide exchange activity of PDI and that the minimal constructs required for catalysis were **abb'** or **b'xa'c**. This study also showed that the isolated **a** and **a'** domains had poor activity towards protein substrates but they were still capable of catalysing redox and isomerisation reactions. However, for full activity, PDI requires further domains in addition to **a** and **a'**, especially the **b'** domain which plays an essential role in substrate binding (Darby et al., 1998, Pirneskoski et al., 2004).

1.3.3 Mechanism

PDI is responsible for catalysing oxidation, reduction and isomerisation of disulphides in the ER. A schematic representation of the mechanism of catalysis of oxidation, reduction and isomerisation is shown in figure 1.9.

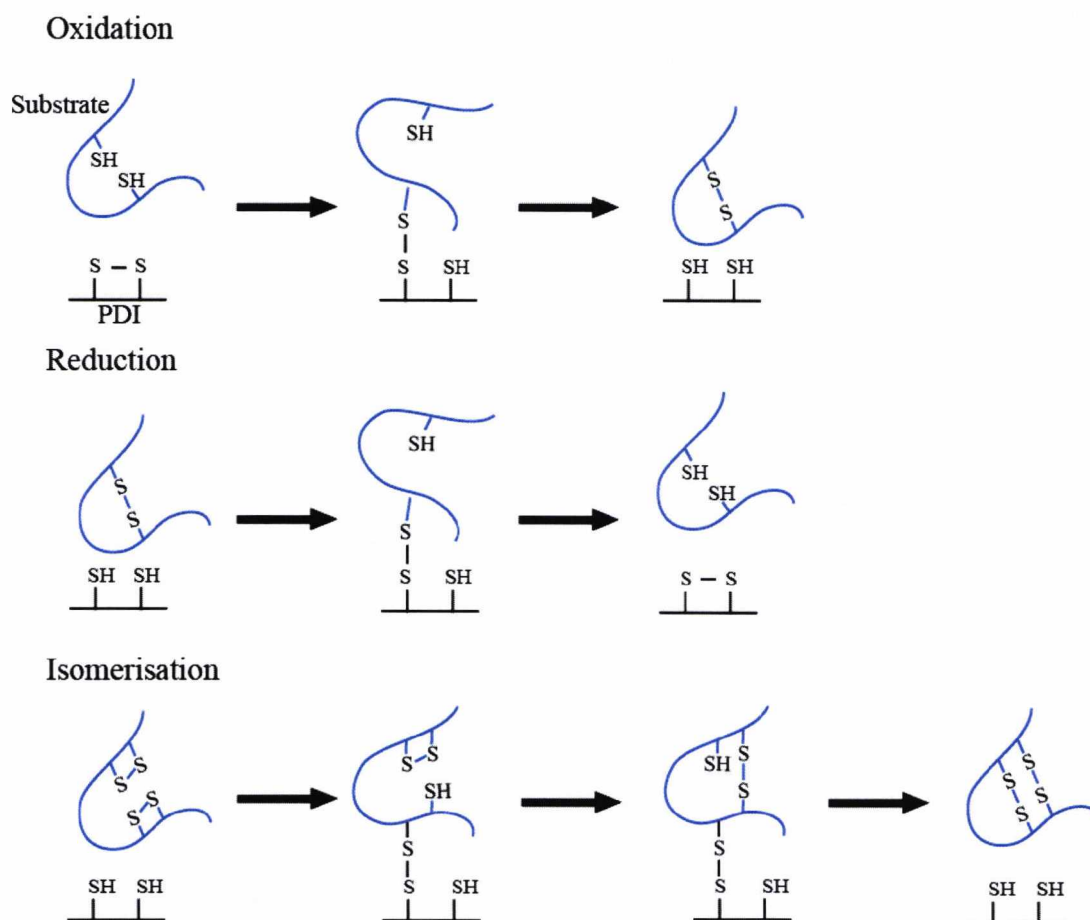


Figure 1.9 Mechanism of catalysis by PDI. Oxidation, reduction and isomerisation of disulphides by PDI. PDI is shown in black and the substrate in blue.

The redox activity of PDI is attributed to the presence of the two thioredoxin-like motifs, CGHC, in the **a** and **a'** domains, which account for the catalytic activity in thiol-disulphide exchange reactions (Edman et al., 1985). PDI consists of two active domains, so two reactions can happen simultaneously.

During oxidation, the substrate dithiol is oxidised to the disulphide state. This reaction requires the conversion of a disulphide CGHC active site to a dithiol. Consequently PDI is reduced, so in order to complete the catalytic cycle it must be reoxidised. Reoxidation *in vitro* can be carried out by using GSSG, glutathione disulphide which is made up of two disulphide bonded glutathione molecules, as the terminal electron acceptor. As a result GSH is generated and there is a change in reduction potential of the solution. If no electron acceptor is present, PDI is able to reduce a disulphide bond in a non-native protein molecule in order to subsequently form a disulphide in another substrate. PDI reoxidation *in vivo* is carried out by members of the ER oxidoreductin 1

family, Ero1 α and Ero1 β (Tavender and Bulleid, 2010). It is not clear why two family members are needed, but Ero1 α is expressed during hypoxia and ER stress, whereas Ero1 β expression is induced by stress due to accumulation of unfolded proteins in the ER (Marciniak et al., 2004, Gess et al., 2003, Pagani et al., 2000). Disulphide transfer from Ero1 to reduced PDI occurs through contacts between the PDI CGHC active site motif and the “outer” active site of Ero1. Following disulphide transfer to PDI, the Ero1 active site is reoxidised by disulphide exchange with its “inner” active site (Tavender and Bulleid, 2010).

Substrate reduction is the reaction in which the substrate disulphide is reduced to the dithiol state and consequently the PDI active site gains a disulphide and is oxidised. *In vitro*, PDI is reduced by DTT or GSH.

Isomerisation is the reaction in which the substrate disulphides are arranged to form a different combination of disulphide bonds (Hatahet and Ruddock, 2009). In this reaction, there is no net change in the redox state of the PDI active site, therefore PDI is recycled without the need for an electron acceptor or donor like in the redox reactions (Hatahet and Ruddock, 2009).

The pK_a of the active site cysteine residues plays an important role in regulating the mechanism of action of PDI. The N-terminal active site cysteines of **a** and **a'**, C36 and C380, have been reported to have pK_a values in the range of 4.4 to 6.7, lower than the pK_a value of a protein cysteine thiol (Hatahet and Ruddock, 2009). This abnormally low pK_a is partly due to its location in the protein structure. It is located at the N-terminus of a α -helix which has a permanent positive dipole associated with it leading to a decrease in pK_a (Kortemme and Creighton, 1995). The ionisable state of the this cysteine side chain is also partly stabilised by the histidine imidazole group located within the active site motif. A combination of these factors maintains the cysteine as a thiolate (S⁻) as opposed to a thiol (SH) which is a weaker nucleophile. The second cysteine, located at the C-terminus of the active site has been described to have limited solvent exposure as it is buried within the 3D structure of the protein (Darby and Creighton, 1995).

The C-terminal active site cysteines of **a** and **a'**, C39 and C383, require a higher than average pK_a value, calculated value of 128 (Lappi et al., 2004), to favour the formation of mixed disulphides with substrate molecules. Consequently, this allows a thiolate from the substrate to displace the mixed disulphide and cause the isomerisation of the substrate. However, this process inhibits the reoxidation of the PDI molecule which

requires nucleophilic attack by the C-terminal cysteine thiolate. In the **a** domain, this problem is circumvented by the presence of a conserved arginine residue, R120, located between $\beta 5$ and $\alpha 4$ of the thioredoxin fold. The side chain of R120 is thought to come into proximity with the active site creating a microenvironment and lowering the pK_a of the C-terminal cysteine, which acts as a nucleophile to enable reoxidation and generate the PDI disulphide bond (Lappi et al., 2004). This process generates the reduced form of the substrate which can undergo further reduction or oxidation during the isomerisation reaction. The **a'** domain of PDI uses the same mechanism.

1.3.4 Ligand binding

The **b'** domain of PDI provides the principal ligand binding site and is essential and sufficient for peptide binding, although additional domains are required for larger substrates (Klappa et al., 1998, Pirneskoski et al., 2004, Byrne et al., 2009). Studies using a range of peptides, including Δ -somatostatin and mastoparan, have revealed that interactions with the ligand binding site on the **b'** domain are based primarily on reversible hydrophobic interactions and to date there is no specific amino acid sequence to which PDI will preferentially bind (Okubara et al., 2003). However, whereas **b'** provides the ligand binding site, both the catalytic domains, **a** and **a'**, are required for binding of mis-folded proteins. Substrate binding by **b'** is sensitive to the conformational stability of other domains, as destabilizing mutations in the **a'** domain indirectly affect ligand binding on **b'**. This could be due to the disordered mutated region occupying the binding site and therefore inhibiting ligand binding (Klappa et al., 2000).

The ligand binding site has been mapped by NMR spectroscopy and it consists primarily of residues from the core β -sheet and α -helices 1 and 3 (Byrne et al., 2009). The isolated **b'x** construct of PDI has been found in conformational exchange and its structure can be stabilized by binding of peptide ligands or the **x** linker region (Byrne et al., 2009). NMR and intrinsic fluorescence data have shown that **b'x** can exist in at least two different conformations in solution. One of these conformations, termed the “capped” conformation, shows the **x** linker region bound to the hydrophobic ligand binding site on the **b'** domain. The other “uncapped” conformation consists of the **x** linker region free in solution leaving the ligand binding site on **b'** free for other substrates to bind (Nguyen et al., 2008). Figure 1.10 shows the **b'x** structure in the “capped” conformation.

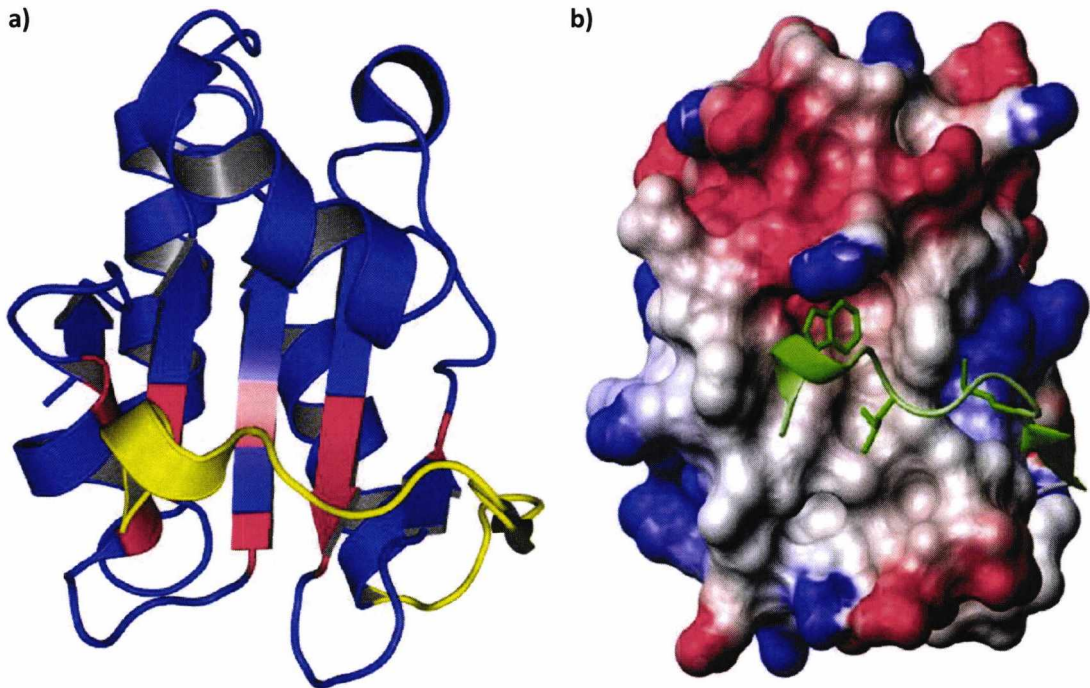


Figure 1.10 Crystal structure of the ligand binding domain **b'x** (I272A mutant 3BJ5.pdb). **a)** **b'x** is shown in blue with residues from the ligand binding site in red mapped by NMR (residues A228, F232, T238, I240, L270, F271, L300, I301, T302, Y310). The **x** linker region is shown in yellow capping the ligand binding site. **b)** Surface charge density map for the **b'** domain with **x** in ribbon format (blue = positive potential, red = negative potential, white = neutral).

The ligand binding site on **b'** is also required for the assembly of PDI with the β -subunit of prolyl-4-hydroxylase, but additional sites on **a** and **a'** are required for the assembly of the tetramer (Koivunen et al., 2005). Therefore substrate binding involves cooperation from at least three of the PDI domains. Additionally, the crystal structure of yeast PDI shows that the **b** domain is also implicated in substrate binding (Tian et al., 2006, Wang et al., 2012a).

1.4 Human PDI family members

Although PDI is considered the main catalyst of thiol-disulphide exchange reactions, it is the founding member of an ever-growing family, the PDI family. To date, there are 20 defined human PDI family members, each containing at least one thioredoxin-like domain with redox activity for ER folding and secretion of proteins. Not all of the PDI family members have been extensively studied, but because they share similarity to at least one PDI domain, they generally share similar catalytic abilities but can have distinct substrates. Figure 1.11 shows all of the current members of the human PDI family of proteins.

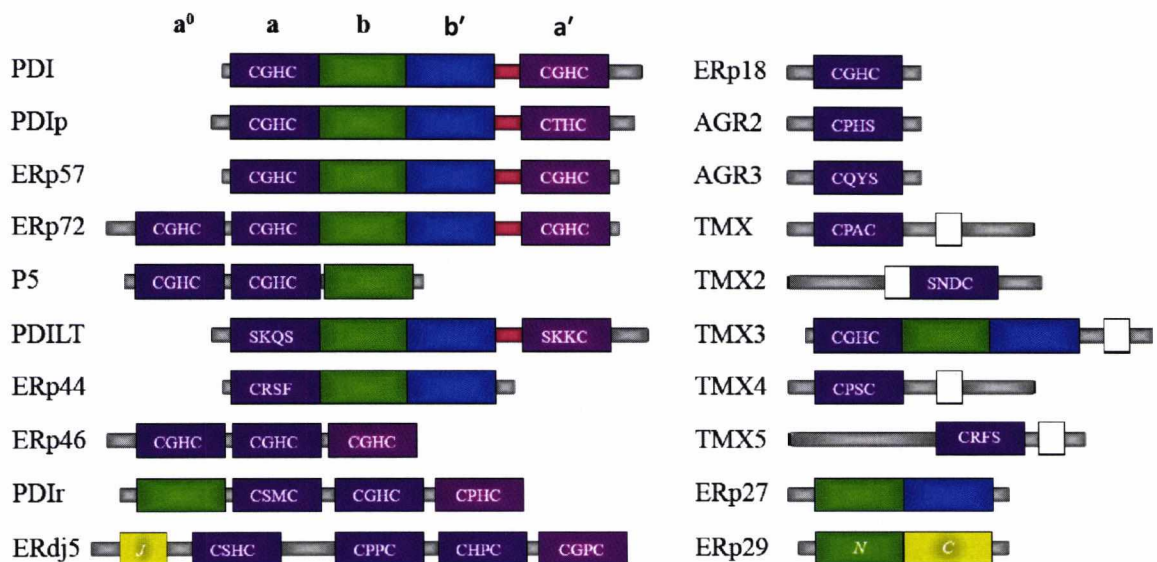


Figure 1.11 The human PDI family (adapted from Kozlov *et al* 2010). Catalytic thioredoxin like domains, **a** and **a'**, are shown in purple and pink, and non-catalytic domains, **b** and **b'**, are coloured in green and blue, respectively. The **x** linker region between domains **b'** and **a'** is shown in red. The DnaJ domain of ERdj5 and the C-terminal helical domain of ERp29 are shown in yellow. White boxes indicate trans-membrane domains. The active site sequence is written in each catalytic domain.

1.4.1 PDIp

PDIp, unlike most PDI family members, is specifically located in the aciner cells of the pancreas (Darby and Creighton, 1995), although recent database analysis implies that it may also be found in the brain and in low levels in organs surrounding the pancreas, such

as the bladder and stomach (Hatahet and Ruddock, 2009). It has the same domain structure as PDI with the **abb'xa'** fragment displaying 49.5% identity. The **a'** domain of PDIp has a unique WCTHC active site sequence. The **c** tail at the C-terminus is much less acidic than that of PDI, instead an additional acidic N-terminal extension occurs before the **a** domain (Ferrari and Söling, 1999). It is able to bind a range of peptides derived from pancreatic digestive enzymes such as pancreatic lipase and pancreatic trypsin inhibitor as well as the hormone somatostatin (Kortemme and Creighton, 1995).

1.4.2 ERp57

ERp57 shares a similar domain architecture to PDI, with the main difference being in the C-terminal extension, where **c** in PDI is rich in acidic residues, the **c** tail of ERp57 contains multiple lysine residues. It is located in the ER via the retention sequence QDEL and as in PDI, the active sites contain the CGHC thioredoxin motif (Frickel et al., 2004). The crystal structure of the **bb'** domain construct of ERp57 is shown in figure 1.12 (Kozlov et al., 2006).

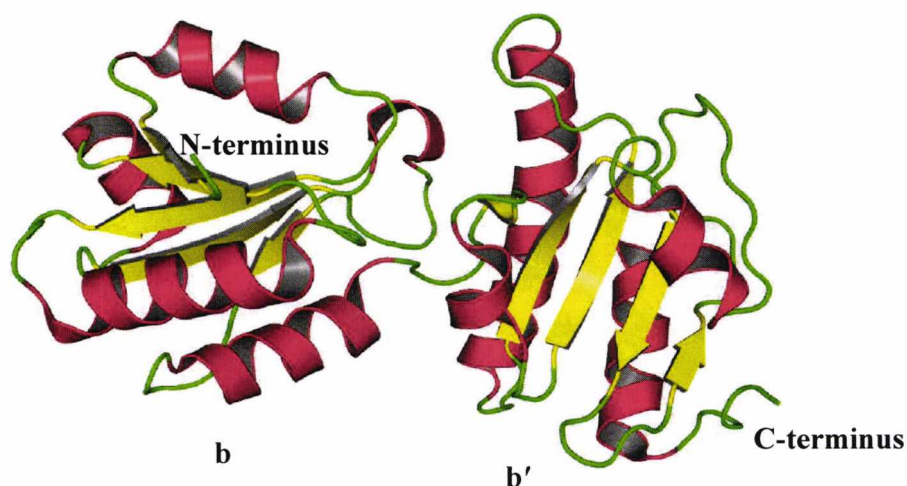


Figure 1.12 The crystal structure of the **bb'** fragment of ERp57 (2H8L.pdb). Domains **b** and **b'** as well as the termini of the construct are labelled in the figure.

ERp57 is thought to play an important role in the oxidative folding of N-glycosylated proteins such as hemagglutinin, rhodopsin and transferrin, as it interacts with the ER lectins, calnexin and calreticulin, which bind to these glycoproteins (Oliver et al., 1999, Hatahet and Ruddock, 2009). It is also involved in folding of MHC (Major Histocompatibility Complex) class I and forms part of the MHC I peptide loading complex

together with calreticulin, TAP and tapasin (Peaper and Cresswell, 2008). ERp57 is important in early development and differentiation of embryonic cells as its deletion has been shown to be lethal in embryonic mice (Garbi et al., 2006).

1.4.3 ERp72

ERp72, also known as CaBP2, was one of the first members of the PDI family to be reported, after PDI itself, but it has still not been studied as extensively as some other members of the PDI family. It shares considerable sequence homology with ERp57 and it has been reported that under physiological conditions, ERp72 can replace ERp57 in knock out cell lines (Hiniker and Bardwell, 2004). It is localised in the ER via a slightly unusual retention sequence, KEEL, and it is thought to possess significant redox and isomerase activity due to the presence of the three active site domains, each with the CGHC sequence (Mazzarella et al., 1990). There is no full length structure available for ERp72 but **bb'** (3EC3.pdb) and **a⁰a** (3IDV.pdb) have been solved by X-ray crystallography.

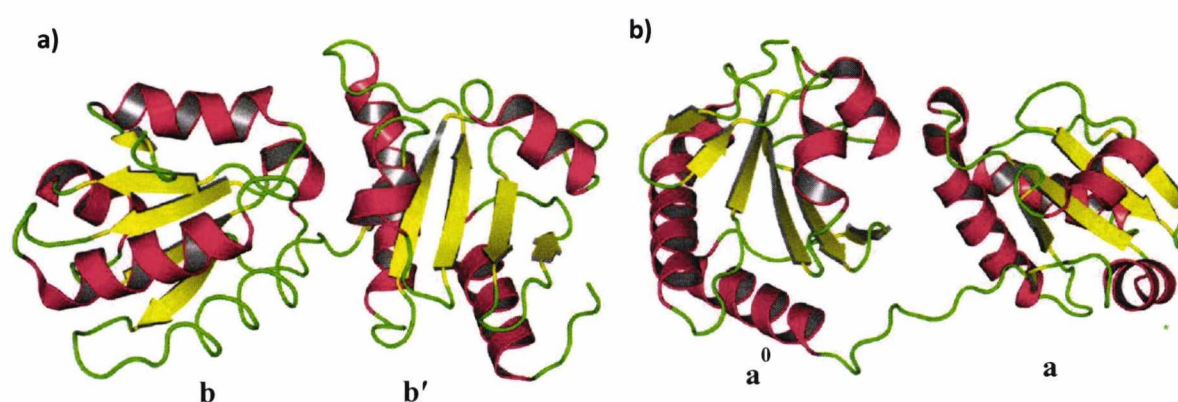


Figure 1.13 The crystal structures of the **bb'** and **a⁰a** fragment of ERp72. a) **bb'** (3EC3.pdb) b) **a⁰a** (3IDV.pdb).

ERp72 forms a complex with other PDI family members like PDI, P5, ERdj3, BiP as well as HSP40, GRP94 and GRP70 (Creighton et al., 1980) but it has been found not to bind peptides (Frand and Kaiser, 1998).

1.4.4 P5

P5, also known as ERp5, was the third member of the PDI family to be identified, but like ERp72 is less well studied than the other family members (Lundström and Holmgren, 1990). It is the smallest redox-active thioredoxin family member with a domain distribution a^0 -**a-b**. Although it contains the KDEL ER-retrieval signal, most reports for P5 function are outside the ER and include a role in platelet function (Song and Wang, 1995) and tumour metastasis (Winter et al., 2002).

1.4.5 ERp44

Human ERp44 was first identified in 2002, with homologues across metazoans (Anelli et al., 2002). It is thought to bind Ero1 α , a protein responsible for maintaining the oxidized state of PDI. ERp44 is highly expressed in secretory tissues and during B-cell differentiation (Wang et al., 2007). The crystal structure of ERp44 showed the **a-b-b'** domains forming a “V” shaped molecule with similar features to PDI, such as the thioredoxin folded domains (Wang et al., 2008). However, ERp44 contains an unusual active site sequence, WCRFS, which lacks the C-terminal cysteine residue.

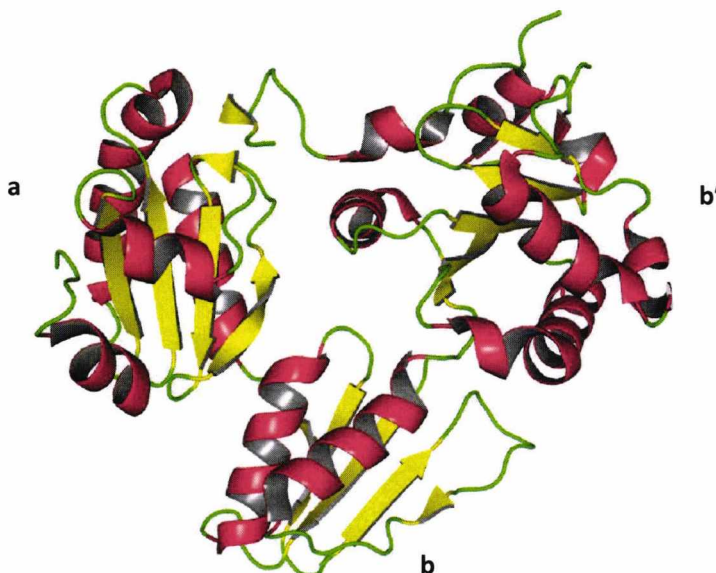


Figure 1.14 The crystal structure of ERp44 (2R2J.pdb) showing the **a**, **b** and **b'** domains arranged in a “V” shape.

1.4.6 PDIr

PDIr has an unusual **b-a⁰-a-a'** domain architecture as its three C-terminal domains are **a** like domains. Two of the thioredoxin motifs are similar to PDI, CXHC, suggesting that it may function as a thiol-disulphide oxidase/isomerase. The physiological function of PDIr is still unknown.

1.4.7 Other PDI family members

ERp29 is an ER resident protein found in rats which is highly expressed on ER stress (Mkrtchian et al., 1998). It consists of two domains and its crystal structure shows the N-terminal domain to be **b**-like, whereas its C-terminal domain is an all α -helical fold (Barak et al., 2009). ERp29 function has been linked to the folding of thyroglobulin within the ER of mammalian cells (Sargsyan et al., 2002).

ERdj5 is a PDI related protein whose function is still unknown. It is localised in the ER via the KDEL retrieval sequence and is induced during ER stress (Cunnea et al., 2003). It is classed as a type III DnaJ as it contains a J domain.

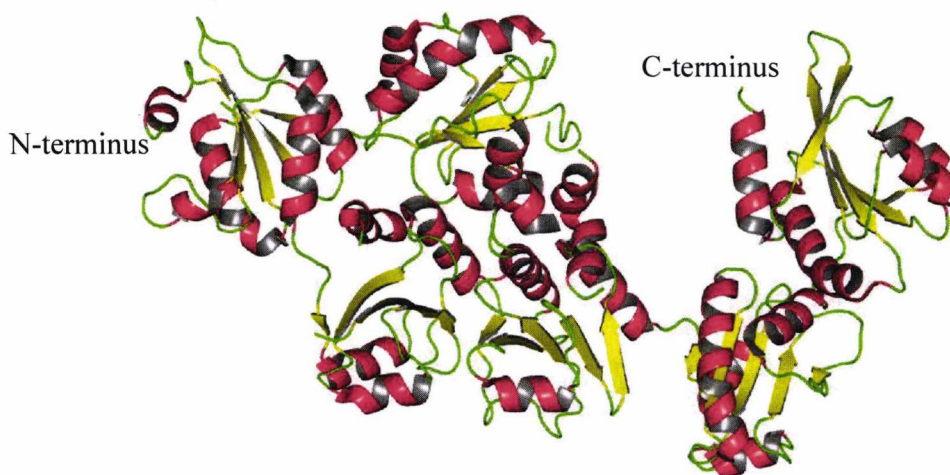


Figure 1.15 The crystal structure of full length ERdj5. The J domain is located at the N-terminus of the protein.

ERp18 is the smallest protein of the human PDI family members, consisting of a single catalytic domain as shown in figure 1.16 (Rowe et al., 2009). It has an unusual active site, with a CGAC motif, and is able to catalyse thiol-disulphide exchange reactions (Alanen et al., 2003).

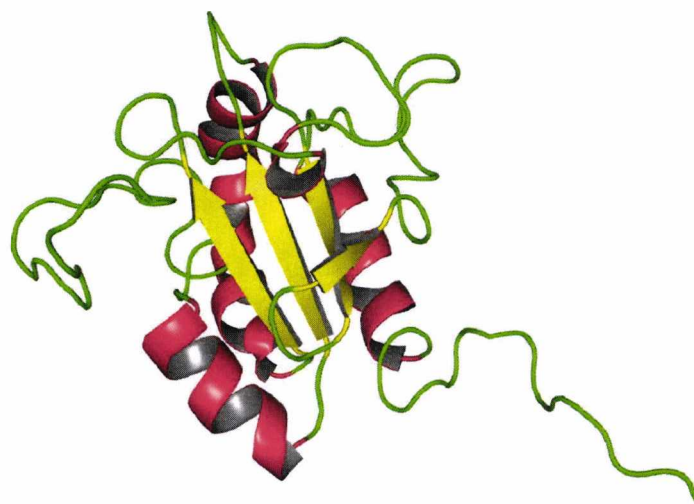


Figure 1.16 The solution structure of ERp18 (2KV8.pdb).

PDILT is highly tissue-specific as it is expressed mainly in the testes after puberty and it is the first family member to show developmental control (Van Lith et al., 2007).

ERp46, also known as endoPDI, has a unique domain arrangement with 3 catalytic domains linked by long spacers. It has been reported to protect cells against hypoxia and can functionally replace PDIp in mice (Knoblach et al., 2003, Sullivan et al., 2003).

ARG2 and ARG3 are single domain proteins localised in the ER and the latest members of the PDI family.

TMX, TMX2, TMX3, TMX4 and TMX5 are transmembrane PDI family members and have putative type I transmembrane protein ER localisation signals but they are not very well characterised (Hatahet and Ruddock, 2009).

ERp27 is a small PDI family member lacking a catalytic domain and its physiological function is still unclear. It contains a b'-like domain through which it binds peptides and non-native proteins, it also interacts with ERp57 via the same mechanism that ERp57 binds to calreticulin and calnexin (Alanen et al., 2006).

1.5 Introduction to Biomolecular NMR

Structural information is invaluable when it comes to understanding a protein's function and mechanism of action. Over the years a number of techniques such as X-ray crystallography and cryo-electron microscopy (cryo-EM) have been used to solve the 3D structure of not only individual proteins but also protein complexes, although cryo-EM is not as detailed as X-ray. The main technique used in this study is nuclear magnetic resonance (NMR) as it cannot only be used to solve 3D structures, but it can also provide information on dynamics and ligand interactions of proteins in solution.

Biomolecular NMR spectroscopy is made possible due to the fundamental property of nuclei known as “spin” which interacts with a static magnetic field. The spin of a nucleus is the same for any given isotope and has half integer value. Isotopes that are studied are either naturally abundant (eg. ^1H and ^{31}P) or if their natural abundance is too low for NMR, they are artificially enriched into the protein of interest (^{15}N and ^{13}C) as they have nuclei with spin $\frac{1}{2}$ and give rise to high resolution NMR data with narrow well resolved peaks. Isotopes like ^{12}C and ^{16}O have a spin of zero and are NMR inactive. NMR experiments collected in this thesis exploited the nuclei of three isotopes, ^1H , ^{15}N and ^{13}C , to gather information about the dynamic properties of PDI proteins as well as to explore the structural changes upon ligand binding. ^1H is the primary NMR isotope with a natural abundance of 99.99%. Naturally occurring ^{14}N and ^{12}C are not suitable for high resolution NMR as ^{14}N gives prohibitive broad peaks and ^{12}C is inactive. ^{15}N and ^{13}C isotopes were used instead, but their natural abundances are much lower at 0.23% and 1.10%, respectively. To circumvent this issue, the isotopes were enriched into the protein during recombinant expression in *E. coli*.

When placed within a magnetic field, the spin of NMR active nuclei generate a magnetic momentum, μ . The magnetic moment aligns with or against the field and will resonate at a frequency that is related to the strength of the applied magnetic field, B_0 . In this thesis, a 14.1 Tesla magnet was used where ^1H nuclei precess with a frequency of around 600 MHz, ^{15}N nuclei precess with a frequency of 60 MHz and ^{13}C at 150 MHz. Further theory into the NMR experiments used in this thesis are described in Chapters 4 and 5.

NMR is also very powerful in detecting subtle changes in the chemical and structural environment of individual nuclei using a property known as the chemical shift in

units of parts per million (ppm). Each nucleus has a potentially different local environment and therefore a unique chemical shift. The chemical shifts for each ^1H , ^{15}N or ^{13}C can be used for the assignment of resonances and subsequent NMR studies of the protein of interest.

1.6 Project Aims

This project aims to characterise the **b'xa'c** construct of hPDI by a number of biophysical techniques and NMR spectroscopy. This C-terminal half of PDI has been described as the minimal unit with isomerase activity due to the presence of the ligand binding site on **b'** and the active site on the catalytic **a'** domain (Pirneskoski et al., 2001, Wang et al., 2010). As the **x** linker region has been shown to act as a gate towards ligand binding, the aim is to investigate the capping mechanism when the **a'** domain is also present. One of the main objectives of this study is to determine what happens to the **x** linker region when it is tethered by both of its neighbouring domains, **b'** and **a'**, as capping could prove to be an important physiological mechanism controlling access to the hydrophobic binding site and could lead to wider structural arrangement of the protein as well as control other aspects of PDI structure/function.

The aims of this thesis were:

1. Site-directed mutagenesis of **b'xa'c** to generate the I272A, L343A, D346A/D348A (2DA) mutants in the WT and W390F backgrounds and optimisation of a protocol to allow expression and purification of **b'xa'c** proteins using the pET23b vector.
2. Ensuring isotopic enrichment of WT **b'xa'c**, its mutants and other PDI constructs such as WT and I272A **b'x**, **a'**, **a'c** and **xa'c**.
3. Determination of capping in **b'xa'c** using intrinsic fluorescence and characterisation of **b'xa'c** proteins using gel filtration and protease sensitivity.
4. Determination of the conformational stability of WT **b'xa'c** and mutants, **xa'c** and **a'c**.
5. Characterisation of **xa'c** and **b'x**, WT and I272A, by NMR including triple resonance assignments of **xa'c**, ligand binding and investigations into the dynamic properties of these proteins.
6. NMR characterisation of WT, I272A, L343A and 2DA **b'xa'c** through temperature and pH titrations, as well as further NMR investigations of WT and I272A **b'xa'c**, including triple resonance assignments of the I272A mutant, dynamic properties and ligand binding investigations to determine if capping of the ligand binding site by the **x** linker region occurs in **b'xa'c**.

CHAPTER 2

Mutagenesis, Protein Production and Purification of b'xa'c

2.1 Introduction

In order to carry out characterisation studies by nuclear magnetic resonance (NMR), milligram quantities of protein are required. For many years protein scientists have been exploiting bacterial expression systems in order to recombinantly over-express their protein of interest. *E. coli* is the most widely used host for the expression of recombinant proteins due to its ability to grow rapidly on inexpensive media at relatively high densities (Francois, 1999). A wide range of vectors and protocols are also available for expression of recombinant proteins in *E. coli*. Recombinant expression also allows uniform isotope labelling with ^{15}N and ^{13}C in order to obtain NMR signals due to the low natural abundance of ^{15}N (0.23%) and ^{13}C (1.1%).

2.1.1 *E.coli* strains and vectors used for protein expression

For this project, two different *E. coli* strains were used, DH5 α and BL21 (DE3) pLysS. *E. coli* DH5 α cells have a high transformation efficiency and produce high plasmid yields. This strain does not contain the T7 RNA polymerase gene, therefore eliminating plasmid instability due to the production of proteins potentially toxic to the host cell. For these reasons, *E. coli* DH5 α cells were used for production and manipulation of DNA.

In contrast, *E. coli* BL21 (DE3) pLysS cells are widely used in protein production. They carry a prophage, λ DE3, encoding the T7 RNA polymerase enzyme which is under the control of the isopropyl- β -D-thiogalactopyranoside (IPTG) inducible *lacUV5* promoter. BL21 strains lack the proteases *lon* and *ompT* which degrade proteins and are therefore ideal for the expression of target genes (Grodberg and Dunn, 1988). They also contain a chloramphenicol-resistant plasmid, pLysS, that provides a small amount of T7 lysozyme, which is a natural inhibitor of T7 RNA polymerase and inhibits leaky transcription. T7 lysozyme also assists in cell lysis so that soluble protein can be released for purification. T7 RNA polymerase binds specifically to the T7 promoter initiating transcription and consequently protein production.

The vector used for the expression of **b'xa'c** was the ampicillin-resistant pHIA241, a derivative of pET-23b(+) plasmid (Novagen), modified so that the region between *Xba*I and *Nde*I was replaced by synthetic oligonucleotides. Therefore, cloning into the *Nde*I site results in the addition of a non-cleavable His-tag (MHHHHHHM) where the final two residues are encoded by the *Nde*I site. Otherwise the vector is unchanged. pHIA241 was received as a gift from Prof. Ruddock (Oulu University, Finland).

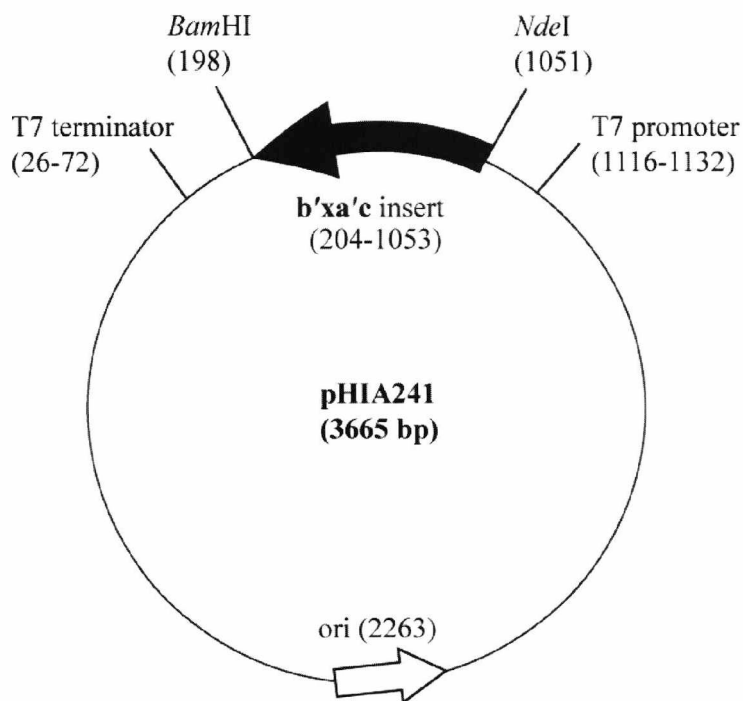


Figure 2.1 Map of pHIA241 (a derivative of pET-23b) showing the location of the restriction enzyme sites *Bam*HI and *Nde*I, where the **b'xa'c** coding sequence is inserted and the T7 terminator and promoter sequences used in the PCR reactions.

2.1.2 Mutagenesis by overlap extension

The pHIA241 vector containing the WT sequence for **b'xa'c** was used as a template for the generation of mutants via site-directed mutagenesis by overlap extension. The mutants generated for the **b'xa'c** insert were I272A, L343A and D346A/D348A (2DA). These are the same mutations that were previously engineered in the **b'x** construct of PDI in order to favour the capping of the ligand binding site by the **x** linker region, in the case of I272A and 2DA, or favour uncapping like in the case of L343A (Nguyen et al., 2008). The purpose of these mutations was to observe if they had the same effect in **b'xa'c** as they have been shown to have in the **b'x** construct. I272A, 2DA and L343A mutants were compared to the WT **b'xa'c** and **b'x** in NMR and fluorescence studies.

The environment of the **x** linker region in **b'x** has also been studied by intrinsic fluorescence. This has been possible due to the presence of the unique Trp residue in the **x** linker region, W347. In the capped conformation, where **x** is buried in the hydrophobic ligand binding site, the emission spectra are characterised by a blue shift in fluorescence. Whereas, in the uncapped conformation, where **x** is more free in solution, spectra display a red shift in fluorescence (Nguyen et al., 2008).

However, the **b'xa'c** construct contains three tryptophan residues, one in **x**, W347, and two in the **a'** domain, W379 and W390, as a result the fluorescence emitted will be a combination of the local environments of the Trp in the **x** region as well as those of the **a'** domain. W35 in the WCGHC active site motif of the **a** domain of PDI has been shown to be heavily quenched in both the dithiol and the disulphide states (Lappi et al., 2004). For that reason it is believed that W379, analogous to W35 and located in the WCGHC **a'** active site, is also heavily quenched and does not contribute to the fluorescence emission spectra of the protein. Therefore, W390 was mutated to a phenylalanine to ensure that W347 was the sole emitter of fluorescence in the **b'xa'c** construct. Thus any fluorescence emitted was representative of the environment of the **x** linker region and capping could be tracked.

For that reason, the W390F mutation was introduced into the WT **b'xa'c** and its three mutants I272A, L343A and 2DA. The mutations that were made are shown highlighted in yellow in the WT **b'xa'c** amino acid sequence in figure 2.2.

210	220	230	240	250	260
MHHHHHML	VIEFTEQTAP	KIFGGEIKTH	ILLFLPKSVS	DYDGKLSNFK	TAAESFKGKI
270	280	290	300	310	320
LFIFIDSHT	DNQRILEFFG	LKKEECPAVR	LITLEEEMTK	YKPESEELTA	ERITEFCHRF
330	340	350	360	370	380
LEGKIKPHLM	SQELPEDWK	QPVKVLVGKN	FEDVAFDEKK	NVFVEFYAPW	<u>CGHCKQLAPI</u>
390	400	410	420	430	440
WDKLGETYKD	HENIVIAKMD	STANEVEAVK	VHSFPTLKFF	PASADRTVID	YNGERTLDGF
450	460	470	480	490	
KKFLESGGQD	GAGDDDDLED	LEEAEPPDME	EDDDQKAVKD	EL	

Figure 2.2 Amino acid sequence of WT b'xa'c. The mutations introduced are highlighted in yellow and the **a'** active site is underlined. The different components of the constructs are shown in colours: His tag (blue), **b'** (green), **x** (red), **a'** (purple), **c** (grey).

Mutagenesis of proteins is a useful tool for better understanding the relationship between protein structure and function. Site-directed mutagenesis helps to understand the function and role of single amino acid residues in a protein of interest by comparing the wild type protein to the mutant protein carrying the desired amino acid change. There are now several methods that can be used to introduce a mutation at a specific residue and so produce the desired amino acid substitution, insertion or deletion (Higuchi, 1988, Ho et al., 1989, Senanayake, 1995). The method used in this project was site-directed mutagenesis by overlap extension first described in 1989 (Ho et al., 1989).

Site-directed mutagenesis by overlap extension uses complementary primers and the polymerase chain reaction (PCR) to generate two DNA fragments with overlapping ends in two separate PCR reactions. This method requires four primers in order to introduce a site-specific mutation. The nucleotide sequence of the primers used is shown in table 2.1 in the Materials and Methods section.

One pair of primers was used to amplify the DNA that contained the mutation site together with downstream sequences (see figure 2.3). The forward mutagenic primer encoded the mutations to be introduced into the wild type template DNA, whereas the reverse primer contained the wild type sequence. In this case, the reverse primer was the T7 terminator encoded by the plasmid. These two primers were used in an amplification reaction, PCR 1, to make the 3' fragment of the desired gene.

The other pair of primers was used to amplify the DNA that contained the mutation site and the upstream sequences. The reverse mutagenic primer of this second pair included the mutations to be introduced into the wild type DNA and the forward primer had a wild type sequence. In this case the forward primer was the T7 promoter encoded by the plasmid. At least 15 bases of the reverse primer in this pair were complementary to the forward primer in the first pair of primers. Amplification of this second pair of primers, in the PCR 2 reaction, generated the 5' end of the desired gene.

The PCR 1 and PCR 2 products were purified by gel electrophoresis in order to obtain a high purity of mutant DNA (Ho et al., 1989). In a third PCR reaction, these fragments were mixed, denatured and annealed to generate heteroduplexes which could then be amplified into full length DNA using the T7 promoter and T7 terminator primers (the two universal primers that bound to the extremes of the two initial fragments). Site-directed mutagenesis by overlap extension is also described in figure 2.3.

The newly generated full length DNA was then cut with the appropriate restriction endonucleases (*Bam*HI and *Nde*I) and gel purified so that the insert was ready to be ligated into the empty pHIA241 vector. Ligation mixture was then used to transform *E. coli* DH5 α cells. Colonies were grown and the DNA from these colonies was used to verify the sequence of the amplified DNA fragment in the vector and to ensure that no mutations other than those in the forward and reverse primers were introduced during these manipulations. Once the mutant sequences were confirmed, the DNA was used for transformation of each mutant into *E. coli* BL21 (DE3) pLysS to enable protein production.

Figure 2.3 illustrates the principle of site-directed mutagenesis by overlap extension.

Site-directed mutagenesis by overlap extension

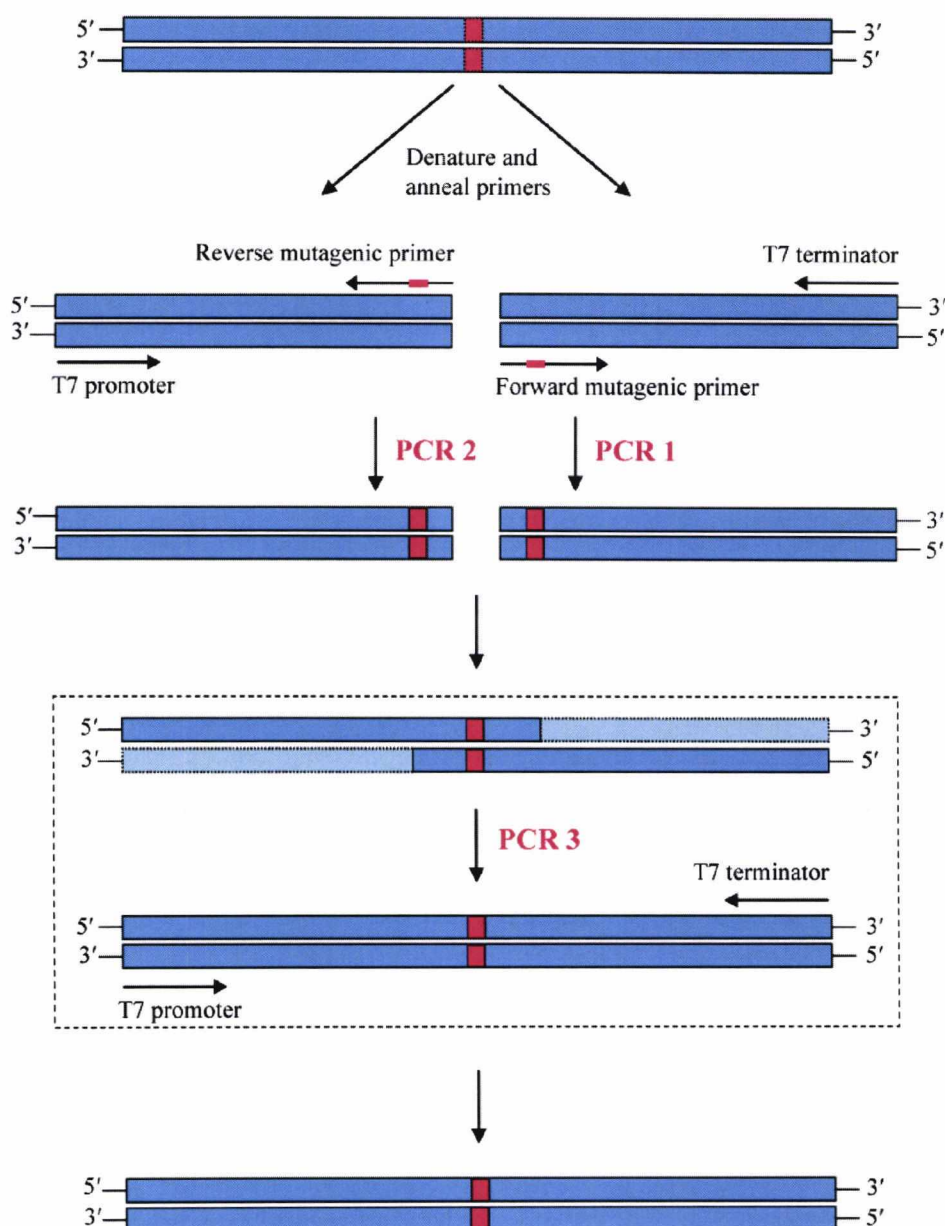


Figure 2.3 In separate amplifications, **1** and **2**, two fragments of the target gene are amplified. **PCR 1** uses the forward mutagenic primer and the T7 terminator, whereas **PCR 2** uses the reverse mutagenic primer and the T7 promoter. The mutation site in the template DNA is shown by the broken line red box. The part of the figure in the broken line box represents the intermediate steps that are likely to occur during **PCR 3**, in which the denatured fragments of the target gene anneal at the region of overlap and are extended as shown by the dashed lines to form full length double-stranded mutant DNA. Also in **PCR 3**, the full length mutant DNA is amplified using the T7 promoter and T7 terminator. (Adapted from Molecular Cloning a laboratory manual by Sambrook and Russell, 3rd Edition).

2.1.3 Protein expression and purification

All proteins discussed in this thesis were expressed and purified using the protocols described in the materials and methods section. Where ^{15}N and/or ^{13}C enrichment was required for NMR samples, cells were grown in media containing ^{15}N ammonium sulphate and/or ^{13}C glucose as sole nitrogen and carbon sources. Protein used for biophysical characterisation was initially expressed in cells grown in LB media, but when purified by gel filtration, it was evident that the majority of the protein was in the dimeric form. However, in minimal media, the proportions of monomer and dimer for WT **b'xa'c** were approximately equal. Therefore, all proteins needed for non-NMR experiments were expressed in minimal media containing laboratory standard ammonium sulphate and glucose as nitrogen and carbon sources.

Protein purity is very important for NMR studies where the protein sample needs to be at least 95% pure. Therefore, after expression, the *E. coli* BL21 (DE3) pLysS cells were lysed to release the cell contents, including the over-expressed PDI proteins, into solution. The lysis process was achieved by freeze-thawing of the *E. coli* cells containing the pLysS plasmid which encodes for T7 lysozyme disrupting the cell wall (Inouye et al., 1973, Studier and Moffatt, 1986). The lysate was then treated with DNase I to break down the DNA that was released during the cell lysis.

PDI proteins were purified from the soluble lysate using nickel affinity chromatography utilising the non-cleavable His tag at the N-terminus of the protein which bound tightly to the nickel ions bound to the chelating sepharose beads. Cellular proteins lacking the His tag should not bind to nickel and were washed away. The bound PDI proteins were eluted from the column with a buffer containing a high concentration of imidazole or EDTA. Imidazole chelated the nickel, removing it from the matrix and the protein. Further impurities were removed by anion exchange chromatography where the negatively charged protein, with a theoretical pI of 4.8, bound to a matrix of positively charged tertiary amine groups. The final purification step, gel filtration, was carried out to separate the different protein conformers according to size into monomer and dimer species. The monomer species were used in further characterisation and structural NMR experiments.

SDS- and native PAGE analysis was carried out throughout the purification process to characterise PDI proteins. Electrospray mass spectrometry was used to confirm the molecular weight of each PDI construct.

2.2 Materials and Methods

All laboratory reagents were supplied by Sigma-Aldrich (Gillingham, United Kingdom) unless otherwise stated.

2.2.1 Mutagenesis by overlap extension

2.2.1.1 Bacterial Strains

The *E. coli* strain DH5 α was used for DNA manipulation steps and recombinant protein expression was carried out in *E. coli* BL21 (DE3) pLysS.

2.2.1.2 Luria Bertani Media

Luria Bertani media was prepared using 10 g/L Tryptone, 5 g/L yeast extract and 10 g/L NaCl. 15 g/L agar was added when solid media was required. The contents were made up to the required amount with MilliQ[®] water in a Schott[®] bottle of approximately twice the liquid volume. The media was sterilised using a standard laboratory autoclave. Antibiotics; 100 mg/L ampicillin and 34 mg/L chloramphenicol (Melford Laboratories, Chelworth, UK) were added at the correct concentrations to the cooled (~50°C) media when required.

2.2.1.3 SOB Media

SOB media was prepared using 20 g/L Tryptone, 5 g/L Yeast extract, 2 g/L NaCl, 10 mL/L 250 mM KCl solution and 5 mL/L 2M MgCl₂. This was made up with the required amount of MilliQ[®] water, corrected to pH 7.0 and sterilised using a standard autoclave before use.

2.2.1.4 Oligonucleotides

HPLC purified forward and reverse oligonucleotides for mutants I272A, L343A, 2DA and W390F were synthesised and supplied by MWG (London, UK). The sequences and melting points of these are shown in Table 2.1.

Oligonucleotide T _m (°C)	Type	Sequence
I272A 62.1 °C	Forward	5' – GATCCTGTTTCGCCTTCATCGAC – 3'
	Reverse	5' – GTCGATGAAGGCGAACAGGATC – 3'
L343A 69.6 °C	Forward	5' – CCAGGAGGCGCCGGAGGACTG – 3'
	Reverse	5' – CAGTCCTCCGGCGCCTCCTGG – 3'
2DA 69.6 °C	Forward	5' – CGGAGGCCTGGGCCAAGCAGC – 3'
	Reverse	5' – GCTGCTTGGCCCAGGCCTCCG – 3'
W390F 62.7 °C	Forward	5' – GGCTCCATTTCGATAAACTGGG – 3'
	Reverse	5' – CCCAGTTTATCGAAAATGGGAGCC – 3'
T7 promoter 54 °C	Forward	5' – TTAATACGACTCACTATAGG – 3'
T7 terminator 58 °C	Reverse	5' – CTAGTTATTGCTCAGCGGT – 3'

Table 2.1 The forward and the reverse oligonucleotide sequences and their respective melting points of the primers used for the generation of the **b'xa'c** mutant proteins. The codons responsible for the mutants are highlighted in grey.

2.2.1.5 Preparation of *E. coli* competent cells for DNA transformation

E. coli DH5 α and BL21 (DE3) pLysS cells were made competent for transformation with DNA using an adapted method from Sambrooks and Russell (2001). Vector free cells from glycerol stocks were streaked out onto solid LB medium without antibiotics and incubated overnight at 37°C. A single colony was picked with a sterile loop and used to inoculate a 50 mL culture of sterile SOB medium. Cultures were incubated for growth at 37°C, 200 rpm until they reached an A_{600nm} of 0.5. Cells were transferred into a sterile 50 mL Falcon tube and harvested by centrifugation at 4,000 rpm at 4°C for 15 min.

Once the supernatant was discarded, the pellet was re-suspended in 25 mL ice-cold 50 mM CaCl₂ and incubated on ice for 30 min. Cells were centrifuged at 4,000 rpm at 4°C for 10 min and the supernatant was discarded. The pellet was re-suspended in 2 mL ice-cold CaCl₂. 50% (v/v) glycerol was added to a final concentration of 15% (v/v). The contents were mixed and 200 μ L aliquots were flash-frozen in liquid N₂ and stored at -80°C.

2.2.1.6 DNA transformation of *E. coli* cells

25 μ L of competent *E. coli* BL21 (DE3) pLysS or DH5 α cells were transformed with 1 μ L of pHIA241 (pET23-b) vector, gently mixed and incubated on ice for 10 min. The cells were heat-shocked at 42°C for 45 s before incubating on ice for a further 5 min. 200 μ L LB medium were added to the cells and shaken at 37°C for 1 h at 200 rpm before spreading on LB plates containing 100 mg/L ampicillin. 34 mg/L chloramphenicol was added to the plates for *E. coli* BL21 (DE3) pLysS cell growth. The cells were grown overnight at 37°C.

2.2.1.7 Preparation of *E. coli* glycerol stocks

Glycerol stocks of *E. coli* cells containing the expression plasmid were prepared when cells growing in LB media had reached an optical density of 0.5 at A_{600nm}. 800 μ L of cells were transferred into a cryovial and mixed with 200 μ L 50% (v/v) glycerol that was previously sterilised by autoclaving. The cells were flash-frozen in liquid N₂ and stored at -80°C for future use.

2.2.1.8 Preparation of template DNA

E. coli DH5 α cells from the WT **b'xa'c** glycerol stock were streaked on an agar/ampicillin plate for growth overnight. The next day, single colonies were picked and grown in 150 mL of LB media containing 100 mg/mL ampicillin at 37°C and 200 rpm for 16 h. After growth the cells were centrifuged at 3,000 rpm for 15 min. The vector DNA was isolated using a Qiagen HiSpeed plasmid Midi kit using the protocol recommended by the manufacturer. The DNA collected was stored at -20°C.

The concentration of DNA was estimated using the absorbance at 260 nm, assuming that 50 μ g/mL of double stranded DNA gives an absorbance of 1.0 at 260 nm.

2.2.1.9 Digestion of DNA with restriction endonucleases

PCR-amplified DNA and plasmid DNA (to obtain the empty vector) were restricted with *Bam*HI and *Nde*I (Promega, Southampton, UK) before proceeding to any cloning procedures. The protocol followed for the restriction reactions was as recommended by the restriction enzyme supplier (New England Biolabs/ Promega). The conditions used for DNA restriction digests were:

DNA	1 µg
<i>Bam</i> HI	1 U
<i>Nde</i> I	1 U
10x enzyme buffer (supplied with restriction enzyme)	2 µL
100x BSA (supplied with restriction enzyme)	0.02 µL
Volume made up to 20 µL with H ₂ O	

For larger scale digests, the reactions were scaled-up accordingly. The DNA was digested for 1-3 h at 37°C. Control experiments were carried out where the DNA was uncut and also cut with one enzyme at a time.

2.2.1.10 Agarose gel electrophoresis of DNA

1% (w/v) agarose gels were made up in TAE buffer (2 M Tris-HCl, 55.6 mL/L glacial acetic acid, 10% (v/v) 0.5 M EDTA, pH 8.0). 0.5 g of laboratory grade agarose (Bio-Rad, Hemel Hempsted, UK) was dissolved in 50 mL of 1x TAE buffer in an Erlenmeyer flask by heating in a microwave oven until the agarose was dissolved. When the solution had cooled to around 50°C, ethidium bromide was added to a final concentration of 0.5 µg/mL. The solution was mixed by swirling and poured into a gel former with combs to form wells for sample loading. The gel was allowed to set for 1 h before sample loading and addition of 1x TAE running buffer to cover the gel. DNA samples were mixed with 6x loading buffer (0.25% w/v bromophenol blue, 0.25% w/v xylene cyanol FF, 30% v/v glycerol) and loaded into the wells with a micropipette. Gels were run at 100 V for approximately 1 h until the fastest migrating dye front had travelled about $\frac{3}{4}$ of the gel length. The DNA was visualised using a short wave UV transilluminator.

A 1 kb DNA ladder (Promega or Fermentas) was run in the first lane of all agarose gels in order to estimate the size of any DNA fragments.

2.2.1.11 DNA purification from agarose gels

DNA bands of interest were cut out of agarose gels with a clean razor blade. DNA was extracted using the QIAquick gel extraction kit (QIAGEN) following the instructions recommended by the supplier. DNA was eluted in 50 µL of elution buffer.

2.2.1.12 PCR amplification

PCR reactions were set up in amplification tubes at 50 μ L final volume. All PCR reactions were carried out using the high fidelity enzyme *Pfu* polymerase (Promega), as specified by the manufacturer. Table 2.2 summarises the composition of the PCR mixture.

Reagent	Stock	Final Concentration
Template DNA	56.2 μ g/mL	0.1 μ g
Amplification buffer	10x	1x
Forward primer	100 pmol/ μ L	30 pmol/ μ L
Reverse primer	100 pmol/ μ L	30 pmol/ μ L
dNTPs	10 mM	200 μ M
<i>Pfu</i> polymerase	1 U/ μ L	1.25 U

Table 2.2 PCR mixture composition

PCR reactions consisted of 25 cycles of 1 min at 95°C (denaturation), 1 min at 50°C (annealing) and 2 min at 72°C (elongation). The first denaturation step was carried out for 2 min and the last elongation step for 5 min. The PCR parameters are summarised in table 2.3 below.

Step	Temperature (°C)	Time (min)	Number of cycles
Initial denaturation	95	2	1
Denaturation	95	1	25
Annealing	50	1	
Extension	72	2	
Final extension	72	5	1
Soak	4	Indefinite	1

Table 2.3 PCR thermocycling parameters used for DNA amplification

2.2.2.13 Splice by overlap extension PCR

PCR-generated DNA fragments (PCR 1 and PCR 2 products, see figure 2.3) were first purified to remove components of the PCR reaction and any contaminating PCR-amplified products and then used in a subsequent overlap extension reaction. Purification

of PCR products was performed by gel extraction using the gel extraction method described in section 2.2.1.11.

The following reagents were mixed in a sterile amplification tube in a final reaction volume of 100 μ L in order to join the 5' and 3' ends of the target gene and to amplify the final product:

Amplification product PCR 1	50 ng
Amplification product PCR 2	50 ng
dNTP mix (200 μ M)	2 μ L
10x amplification buffer	10 μ L
10 μ M T7 promoter	3 μ L
10 μ M T7 terminator	3 μ L
<i>Pfu</i> polymerase (1 u/ μ l)	2 μ L
H ₂ O	56 μ L

PCR reaction was carried out as shown in table 2.3 except the extension time was increased to 3 min to match the increased size of the product and the number of cycles was increased from 25 to 30 to increase the final DNA concentration. The newly generated uncut insert was purified as above before undergoing a double digest with restriction enzymes *Bam*HI and *Nde*I to create the insert ready for ligation into the empty vector.

2.2.1.14 Ligation of insert into empty vector

Ligation of the mutation-carrying insert into the empty pET23 derivative vector, pHIA241, was carried out overnight at 10°C or for 2 h at room temperature. The insert was present at approximately 5 times molar excess compared to the vector concentration. The following reagents were mixed in a sterile microcentrifuge tube to give a final reaction volume of 20 μ L:

NE Buffer4 (10x)	2 μ L
T4 Ligase	2 μ L
Vector	20 ng
Insert	100 ng

After ligation, the mixture was used to transform competent *E. coli* DH5 α cells. The same transformation protocol as described in section 2.2.1.6 was used. The transformed cells were spread on agar plates containing 100 mg/L ampicillin. The plates were incubated overnight at 37°C for growth. Controls were carried out where cells were transformed

with empty vector to determine the rate of self-ligation. Colonies from the plates were picked and used to inoculate 10 mL LB media containing ampicillin for growth overnight at 37°C and 200 rpm. Following growth, the cultures were centrifuged at 4,000 rpm, 4°C for 15 min. DNA was extracted from the pelleted cells using QIAprep[®] Miniprep kit as instructed by the manufacturer (Qiagen). The DNA was eluted in 50 µL elution buffer (10 mM Tris-Cl, pH 8.5) from the kit.

2.2.1.15 Screening

Colony screening was carried out in order to identify colonies containing the vector with the correct insert. Colonies were picked from the transformation plates and grown as described in 2.2.1.14. The vector DNA was purified using QIAprep[®] Miniprep kit as described above. The vector DNA was digested with restriction enzymes *Bam*HI and *Nde*I (Promega) for 1 h at 37°C and the digested DNA was run on agarose gels next to a control containing uncut vector.

2.2.1.16 DNA sequencing

Mutant DNA from *E. coli* DH5α transformants (extracted using QIAprep[®] Miniprep kit) was sent to Cogenics for sequencing.

Protein expression and purification techniques

2.2.2.1 Minimal Medium

For the preparation of isotopically enriched (^{15}N or $^{15}\text{N}/^{13}\text{C}$) recombinant proteins, the minimal medium, (MM) described in Table 2.4 was used. $(\text{Na}^+/\text{K}^+)\text{PO}_4$, Na_2SO_4 , EDTA trace elements and MilliQ[®] water were combined in an Erlenmeyer flask that was five times the volume of the media required and sterilized by autoclaving. All other components were sterilized using 0.2 μm single use syringe filters (Sartorius, Eastbourne, UK) and added aseptically to the flask prior to inoculation.

Component (stock concentration)	Stock solution	Volume /L	Method of sterilization
1. $(^{15}\text{NH}_4)_2\text{SO}_4$ (x100)	30 g/L	20 mL	Filter
2. PO_4/NaCl (x10)	Na_2HPO_4 68 g/L KH_2PO_4 30 g/L NaCl 5 g/L	100 mL	Autoclave
3. Na_2SO_4 (x1,000)	42.6 g/L	1 mL	Autoclave
4. EDTA trace (x100)	EDTA 10 g/L MnCl_2 3.2 g/L FeCl_3 1 g/L ZnCl_2 0.1 g/L CuCl_2 20 mg/L CoCl_2 20 mg/L H_3BO_3 20 mg/L	10 mL	Autoclave
5. MgSO_4 (x1,000)	246 g/L	1 mL	Filter
6. CaCl_2 (x1,000)	44.1 g/L	1 mL	Filter
7. Biotin (x1,000)	1 g/L	1 mL	Filter
8. Thiamine (x1,000)	1 g/L	1 mL	Filter
9. Glucose (x50)	150 g/L	20 mL	Filter
10. MilliQ water		845 mL	Autoclave
11. Ampicillin (x100)	100 g/L	1 mL	Filter
12. Chloramphenicol (x100)	34 g/L	1 mL	Filter

Table 2.4 Minimal Medium components and their appropriate concentrations. Components 2, 3 and 4 were autoclaved. All other components were filter sterilised into sterile containers and added to the flask aseptically, once the media was cooled prior to inoculation.

2.2.2.2 Recombinant protein expression in *E. coli*

A starter culture of 50 mL of growth media (LB or MM) containing 100 mg/L ampicillin and 34 mg/L chloramphenicol was inoculated with a single colony from a

streaked plate and grown at 37°C with shaking (200 rpm) until $A_{600} = 0.6-1.0$. The cells were sedimented (4,000 rpm, 15 min, 4°C) and re-suspended in 2 mL of media. This was used to inoculate a 1 L culture containing antibiotics, which was incubated at 37°C, 200 rpm for growth until an A_{600nm} of 0.6-0.8 was reached.

The cells were then induced with 0.45 mM IPTG for 3 h, cooled on ice for 5 min and sedimented by centrifugation in a JA 10 rotor at 6,000 rpm for 15 min at 4°C. The supernatant was discarded and the cells were re-suspended in 10 mL lysis buffer (20 mM NaH_2PO_4 , 50 mM NaCl, pH 7.3) before freezing at -20°C.

2.2.2.3 Lysis

Cells were thawed at room temperature before the addition of Triton[®] X-100 (200 μ L of 0.1% v/v stock) and then incubated for 20 min at room temperature. For DNase treatment, 50 μ L of 2 M MgCl_2 (10 mM final concentration) and 100 μ L DNaseI (20 μ g/mL final concentration) were added and the cells were incubated at room temperature for a further 20 min. The supernatant was separated from the insoluble cell debris by centrifugation at 13,000 rpm, for 20 min, at 4°C in a JA 20 rotor and used for further protein purification.

2.2.2.4 Metal Affinity Chromatography

This purification step was carried out using immobilised metal-affinity chromatography (IMAC), where nickel bound to chelating sepharose (GE Healthcare, Little Chalfont, UK) was used to bind to the hexa-histidine tag at the N-terminus of the protein. A simple column was produced by packing glass wool into the bottom of an empty 20 mL plastic syringe.

Chelating Sepharose Fast Flow (GE Healthcare) slurry was poured into the syringe body and the resin was allowed to settle at the bottom to get a final column volume of 5 mL. The resin was stored in 20% ethanol which was washed out of the column with 10 column volumes (CV) of water. The column was then charged with 1 CV of NiSO_4 in water (0.2 M), washed again with 5 CV of MilliQ[®] water and equilibrated with 10 CV of binding buffer (20 mM NaH_2PO_4 , 10 mM imidazole, 0.5 M NaCl, pH 7.3).

The supernatant containing the soluble protein was filtered with a 0.2 μ m single use syringe filter, loaded onto the column and allowed to flow through by gravity. The flow through was collected for SDS-PAGE analysis. The bound protein was washed with 10

CV of binding buffer followed by another 10 CV of wash buffer (20 mM NaH₂PO₄, 50 mM imidazole, 0.5 M NaCl, pH 7.3) to remove loosely bound impurities. Protein was eluted with 10 CV of elution buffer (20 mM NaH₂PO₄, 200 mM imidazole, 0.5 M NaCl, pH 7.3). The NiSO₄ was stripped from the column with 10 CV of strip buffer (20 mM NaH₂PO₄, 20 mM EDTA, 0.5 M NaCl, pH 7.3). After use, the column was cleaned with 10 CV of water and stored at 4°C in 20% ethanol for use at a later date. All wash, elution and strip fractions were collected for SDS-PAGE analysis.

2.2.2.5 Dialysis

The eluted protein was transferred into dialysis tubing (MWCO 12-14 kDa) to be dialysed overnight into 2 L of Buffer A (20 mM NaH₂PO₄, pH 7.3) at 4°C in order to desalt the sample ready for purification by anion-exchange.

2.2.2.6 Anion-exchange Chromatography

Anion-exchange chromatography was carried out using a Source 30Q column (5 mL CV with 30 µm beads) attached to the ÄKTA FPLC analyser (GE Healthcare). Buffers A (20 mM NaH₂PO₄, pH 7.3) and B (20 mM NaH₂PO₄, 0.5 M NaCl, pH 7.3) were filtered and degassed before use. Prior to loading the protein, the anion-exchange column was prepared by washing with 10 CV of Buffer A, 10 CV of Buffer B and equilibrated with 10 CV of Buffer A. 0.2 µm filtered protein was loaded onto the column via an injection superloop and eluted from the column by washing with a gradient of Buffer B from 0 to 100% over 20 CV at a flow rate of 5 mL/min. 2 mL fractions were collected throughout the gradient and the presence of protein was measured by an increase in absorbance at 280 nm. Fractions of interest were analysed by SDS and native PAGE.

2.2.2.7 Gel Filtration

Preparative and analytical gel filtration were carried out on Superdex 200 matrix 300 mL bed volume (GE Healthcare), poured into an XK 26 column with Buffer B as running buffer. Samples loaded onto the gel filtration column were pooled fractions from the anion-exchange chromatography which were concentrated to a maximum volume of 2 mL using Vivaspin 20 centrifugal filters (10 kDa MWCO; Sartorius, Eastbourne, UK). The sample was loaded onto the column by gravity flow via a three way tap and eluted

with 1.5 CV of Buffer B at a flow rate of 2 mL/min. 5 mL fractions were collected throughout the elution with in-line A_{280} monitoring using the ÄKTA FPLC analyser.

The column was calibrated using a low molecular weight gel filtration marker kit from GE Healthcare consisting of bovine lung aprotinin (6.5 kDa), bovine pancreas ribonuclease A (13.7 kDa), bovine erythrocyte carbonic anhydrase (29 kDa), hen egg ovalbumin (44 kDa), chicken egg white conalbumin (75 kDa) and Blue Dextran (2,000 kDa). Blue Dextran 2,000 was used to determine the void volume of the column which was equal to the retention volume for the dye. Aldolase (161 kDa) from rabbit muscle (Sigma, Gillingham, UK) was used as a high molecular weight marker in addition to the kit.

2.2.2.8 Mass Determination

The molecular mass of purified proteins were determined by Electrospray Mass Spectrometry carried out by Kevin Howland. In order to confirm that the correct proteins were expressed, 10 μ g of protein were analyzed using a Finnigan/Agilent 1100 MAT LQC ion-trap mass spectrometer. Mass spectra were recorded on a ThermoFinnigan LQC Classic ion trap mass spectrometer. Samples were desalted by in-line reverse-phase HPLC on a Phenomenex Jupiter C4 column running on an Agilent 1100 HPLC system. Mass spectra were deconvoluted with an expected accuracy of $\pm 0.01\%$ using the BioExplore package from ThermoFinnigan to get protein masses.

2.2.2.9 SDS-PAGE analysis

Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) was prepared using the Tris-Glycine buffer system to identify polypeptides under denaturing conditions (Laemmli, 1970). The discontinuous gel system consisted of a resolving gel at pH 8.8 and a stacking gel at pH 6.8. SDS-PAGE gels were produced with 10 or 12.5% (w/v) resolving gels depending upon the molecular weight of the protein to be resolved, overlaid with 4% (w/v) stacking gel using gel cassettes from Novagen. The composition of the resolving and stacking gel is given below in Table 2.5.

Gel Type	Reagent	10% Volume (mL)	12.5% Volume (mL)	Gel Type	4% Volume (mL)
Resolving gel	40% acrylamide (37.5:1)	6.25	7.81	Stacking gel	1
	1.5 M Tris, pH 8.8	6.25	6.25		2.5
	H ₂ O	11.99	10.43		6.29
	10% SDS	0.25	0.25		0.1
	10% APS	0.25	0.25		0.1
	TEMED	0.01	0.01		0.01

Table 2.5 Composition of 10 and 12.5% SDS-PAGE resolving and stacking gels.

25 mL of the resolving gel and 10 mL of the stacking gel premix were sufficient to make three gels after polymerization with 10% APS (ammonium persulphate) and TEMED (Bio-Rad, Hertfordshire, UK). Protein standards SDS7 and Precision Plus were purchased from Sigma and Bio-Rad, respectively. Gel running buffer (5x stock) consisted of 0.95 M glycine (pH 8.3), 0.125 M Tris and 0.5% (w/v) SDS. Non-reducing gel loading buffer (4x) consisted of 0.33 M Tris-HCl, 8% (w/v) SDS, 0.4% (w/v) bromophenol blue and glycerol to a final concentration of 40% (v/v) (Sambrook and Russell, 2001). DTT was added to a final concentration of 200 mM for a reducing buffer. The aliquoted buffer was stored at -20°C until required.

40 µL samples were prepared for analysis by mixing 30 µL of protein with 10 µL 4x reducing gel loading buffer and incubated at 100°C for 3 min. Once cooled, 10 µL of sample were loaded on the gel. Empty wells, if any, were loaded with 10 µL of 1x sample buffer. Gels were run at constant 125 V for approximately 1.5 h. They were then stained for 1 h in a 40% methanol, 10% acetic acid solution containing 1 mg/mL Brilliant Blue G and de-stained in a solution containing 10% methanol and 10% acetic acid. Once completely de-stained, gels were transferred in gel drying buffer (30% (v/v) ethanol, 3% (v/v) glycerol) for at least 1 h and then dried between two sheets of cellulose stretched over a frame overnight (Bio-Rad).

2.2.2.10 Native PAGE analysis

Continuous native PAGE analysis was carried out to determine the oligomeric state of the PDI proteins under non-denaturing conditions. 12.5% or 10% continuous resolving gels, lacking the stacking gel, were prepared using the Tris-Glycine buffer system described in 2.2.2.9 at pH 8.3 with the exception of SDS from the buffer mixture. Gel running buffer (5x

stock) consisted of 0.95 M glycine and 0.125 M Tris. Non-reducing gel loading buffer (2x) consisted of 25 mL 0.5 Tris, pH 6.8, 20 mL glycerol, 1 mg bromophenol blue and made up to 100 mL with water.

30 μ L samples were prepared for analysis by mixing 15 μ L of protein with 15 μ L of 2x gel loading buffer. 20 μ L of sample were loaded on the gel. Gels were run at constant 10 mA for approximately 3 h. They were then stained, de-stained and dried as the SDS-PAGE gels described in 2.2.2.9.

2.2.2.11 Determination of protein concentration

The absorbance of proteins at A_{280} nm was measured in an Ultropsec 2000 UV/Visible Spectrophotometer (Pharmacia Biotech). Where necessary, proteins were diluted 100x by adding 10 μ L of protein to 990 μ L of the appropriate buffer that was used to zero the spectrophotometer. The extinction coefficient was calculated by submitting the protein amino acid sequence into ProtParam in the ExpPASy website (<http://web.expasy.org/protparam/>). The extinction coefficient for WT **b'xa'c** and I272A, L343A and 2DA mutants was $24200 \text{ M}^{-1} \text{ cm}^{-1}$, whereas the extinction coefficient for the W390F mutants was $18700 \text{ M}^{-1} \text{ cm}^{-1}$. The protein concentrations were calculated in molar using the Beer-Lambert Law.

2.3 Results

2.3.1 Preparation of vector DNA

The purified vector containing **b'xa'c** was cut with restriction enzymes *Bam*HI and *Nde*I in order to obtain the empty pHIA241 vector to be used for the ligation of the inserts carrying the desired mutation. Controls were carried out in which the circular DNA was linearised with each of the restriction enzymes independently to ensure that the DNA was being digested by both enzymes. As seen from figure 2.4, the template **b'xa'c** DNA was cut successfully by both enzymes. The insert released could be seen between 750 and 1,000 bp, which matches with its expected size of 855 bp. The empty purified vector, lacking the insert, seen in figure 2.4, appeared between 3,000 and 4,000 bp, compatible with its expected size of 3633 bp.

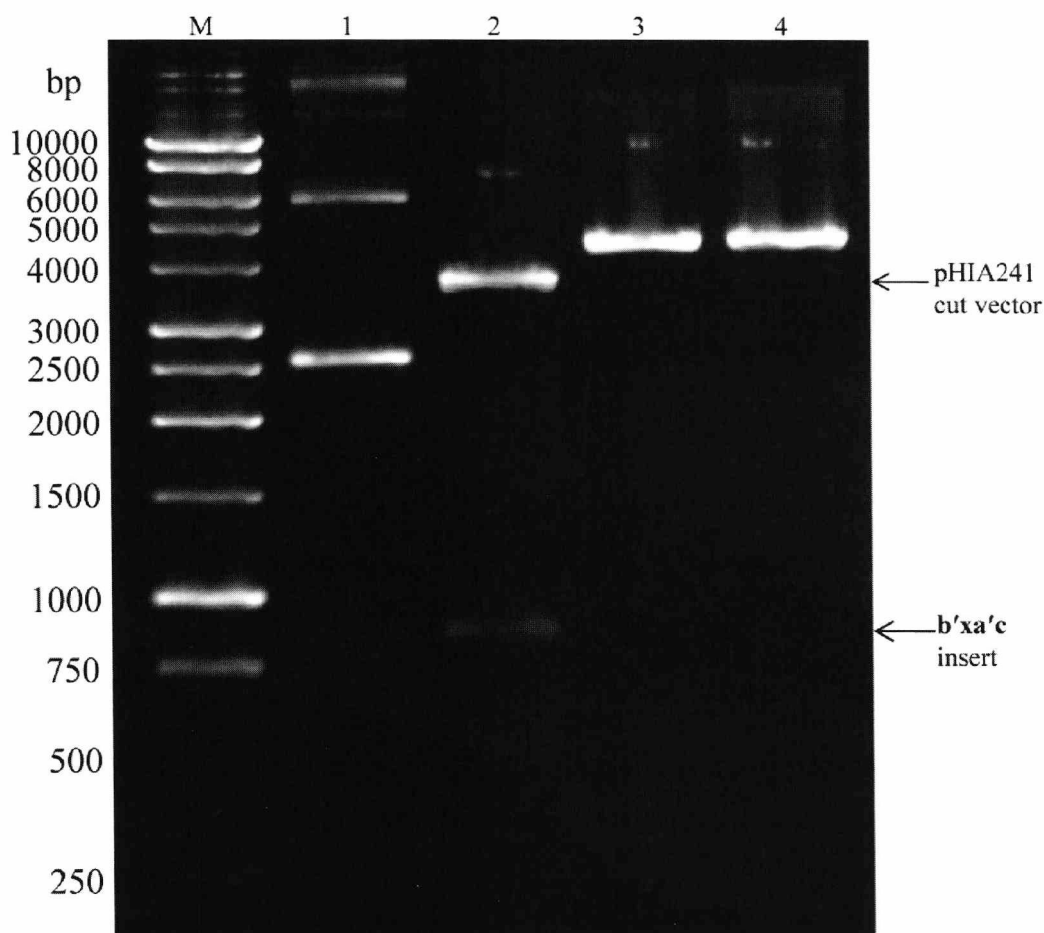


Figure 2.4 1% Agarose gels showing the digest of pHIA241 vector. Lane M: 1kb DNA ladder; Lane 1: intact vector; Lane 2: double digested vector; Lane 3: vector linearised with *Nde*I only; Lane 4; vector linearised with *Bam*HI only.

2.3.2 Generation of PCR products

The I272A, 2DA, L343A and W390F mutant inserts were generated by PCR overlap extension. The 5' and 3' ends of each insert were generated in two separate PCR reactions using two sets of primers. The products were then joined and amplified in a third PCR reaction as shown in figure 2.3. Table 2.6 shows the expected sizes of each of the fragments generated by PCR.

Mutants	PCR 1 product size (bp)	PCR 2 product size (bp)	PCR 3 product size (bp)	Size of <i>Bam</i> HI/ <i>Nde</i> I cut insert (bp)
I272A	811	282	1067	855
2DA	580	509	1067	855
L343A	591	497	1067	855
W390F	453	637	1067	855

Table 2.6 Expected sizes of the PCR-generated fragments and the cut inserts.

Agarose gel analysis of the PCR 1 and PCR 2 fragments for all three mutants showed that all fragments had the correct sizes, shown in figure 2.5. PCR 1 and PCR 2 fragments were gel-purified to remove components of the PCR reaction which might interfere with subsequent reactions. Joining the PCR 1 and PCR 2 fragments resulted in the formation of the uncut inserts carrying the desired mutations, which appeared to have the right sizes on agarose gels (figure 2.5). PCR 1 and PCR 2 products were successfully joined together in an amplification reaction which resulted in the b'xa'c insert carrying the desired mutation. The insert was effectively digested with restriction nucleases *Bam*HI and *Nde*I to generate the correct sticky ends for ligation into the empty vector.

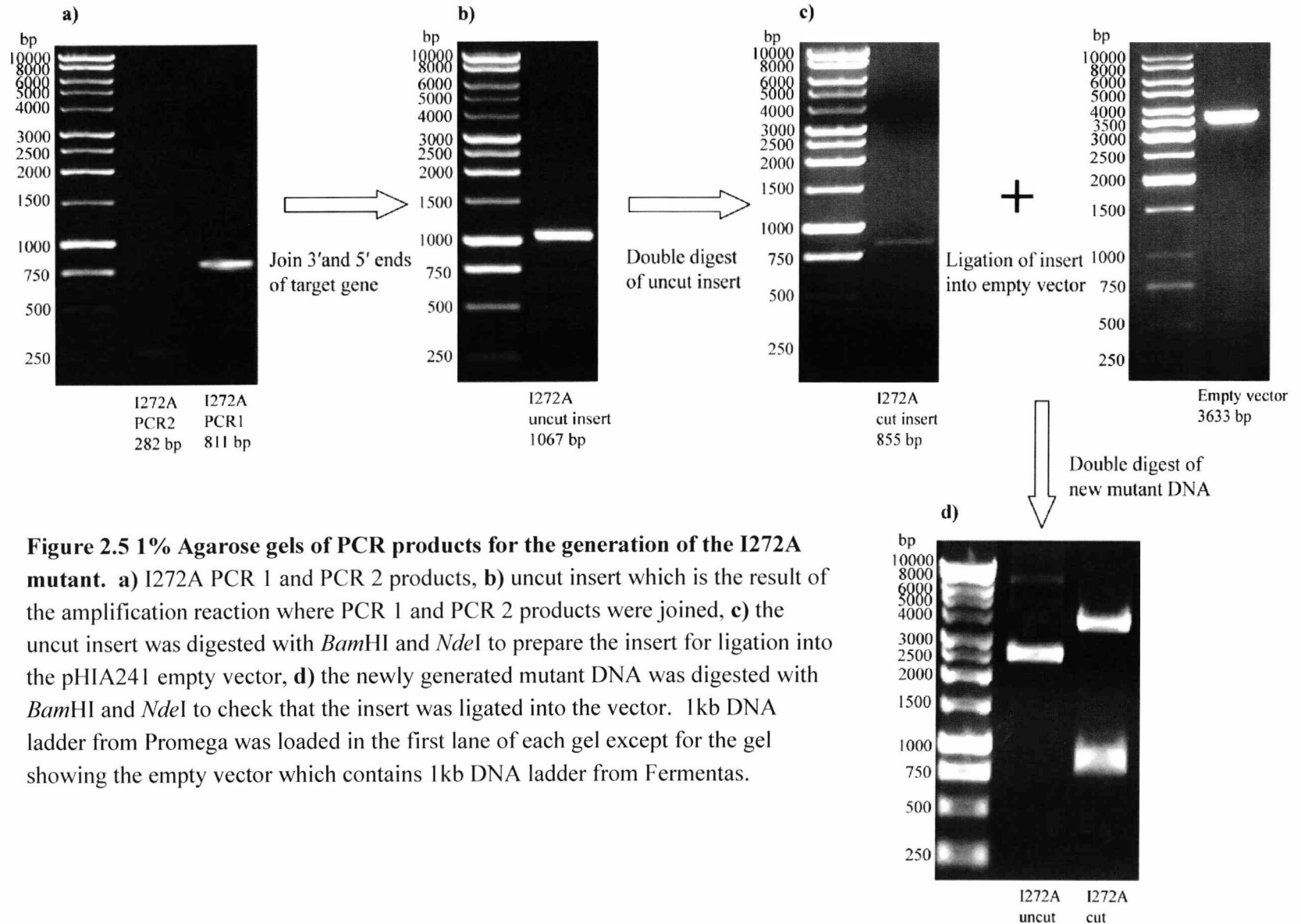


Figure 2.5 1% Agarose gels of PCR products for the generation of the I272A mutant. a) I272A PCR 1 and PCR 2 products, b) uncut insert which is the result of the amplification reaction where PCR 1 and PCR 2 products were joined, c) the uncut insert was digested with *Bam*HI and *Nde*I to prepare the insert for ligation into the pHIA241 empty vector, d) the newly generated mutant DNA was digested with *Bam*HI and *Nde*I to check that the insert was ligated into the vector. 1kb DNA ladder from Promega was loaded in the first lane of each gel except for the gel showing the empty vector which contains 1kb DNA ladder from Fermentas.

2.3.3 Generation of mutant DNA

Ligation mixture of the mutant insert and the empty vector was used to transform *E. coli* DH5 α cells. The presence of colonies after overnight incubation confirmed that the insert had been ligated into the vector. DNA extracted from these cells was sent for sequencing to establish that the newly generated DNA carried the desired mutation. Sequencing results confirmed that the I272A, L343A, 2DA as well as all of the W390F mutants were successfully engineered. The nucleotide and amino acid sequences for the **b'xa'c** mutants are shown in appendix 2.1.

2.3.4 Protein Expression

Expression of the construct in *E. coli* BL21 (DE3) pLysS produced round, white colonies on agar plates. Growth in LB media was normal with a 30 min generation time. Growth in minimal medium was slower, with a doubling time of approximately 60 min. The expression and purification results shown relate to WT **b'xa'c** and are a typical example of the analysis carried out at each stage. The same expression and purification protocol generated similar results for the mutant proteins.

All PDI constructs used in this project were expressed as soluble proteins as shown in figure 2.6. By comparison with the molecular weight markers, the over-expressed proteins were estimated to have a molecular weight between 25 and 37 kDa. The expected molecular weight of WT **b'xa'c** and its I272A, L343A and 2DA mutants is 32.5 kDa for the WT and 32.4 kDa for the mutant proteins.

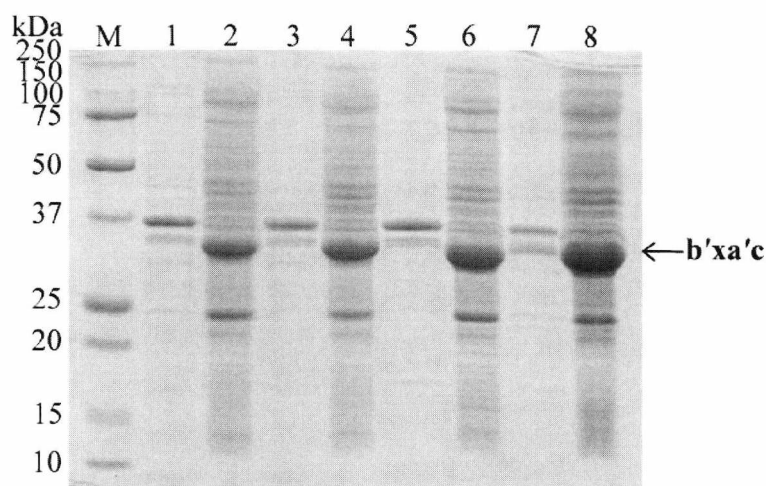


Figure 2.6 SDS-PAGE analysis of the soluble and insoluble fractions of **b'xa'c after cell lysis.** Lane 1: WT pellet; lane 2: WT supernatant; lane 3: I272A pellet; lane 4: I272A supernatant; lane 5: L343A pellet; lane 6: L343A supernatant; lane 7: 2DA pellet; lane 8: 2DA supernatant.

2.3.5 Metal Affinity Chromatography

Purification by IMAC proved to be successful in removing the bulk of contaminants from the protein sample as seen in the SDS-PAGE in figure 2.7. However, some high and low molecular weight impurities were still seen in the protein elution fractions.

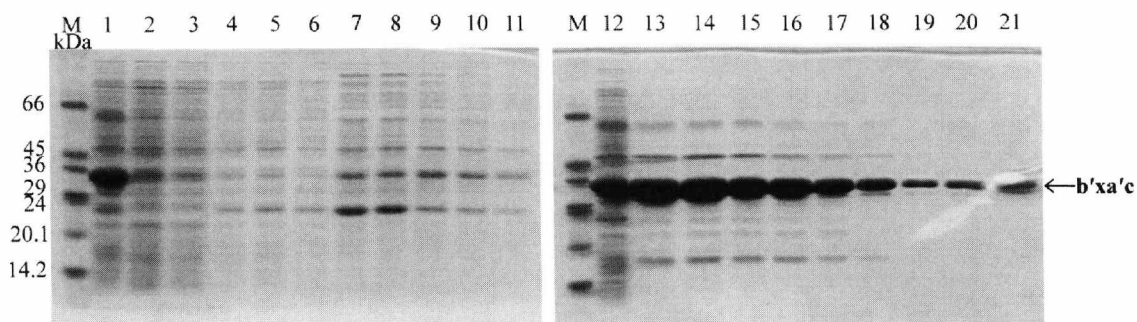


Figure 2.7 SDS-PAGE analysis of WT **b'xa'c from metal affinity chromatography.** Lanes 1 and 12: protein before metal affinity purification; lane 2: flow through; lanes 3-6: fractions from wash with binding buffer (10 mM imidazole); lanes 7-11: fractions from wash with wash buffer (50 mM imidazole); lanes 13-18: fractions from elution with 200 mM imidazole; lanes 19-21: fractions from EDTA strip.

To ensure that these impurities were removed, the fractions containing protein were pooled, buffer exchanged into Buffer A and further purified by anion exchange chromatography.

2.3.6 Anion-exchange Chromatography

Anion exchange was successful in removing the majority of the remaining impurities from the protein samples. To analyse the homogeneity of the sample, fractions across the peak were run on SDS and native PAGE.

The chromatogram in figure 2.8 shows that the ion exchange profile of WT **b'xa'c** presented with an asymmetric peak suggesting the presence of multiple protein species. Furthermore, even when the elution peak appeared symmetrical, suggesting a single species homogenous in charge and conformation, analysis by native PAGE revealed the protein comprised of at least two forms.

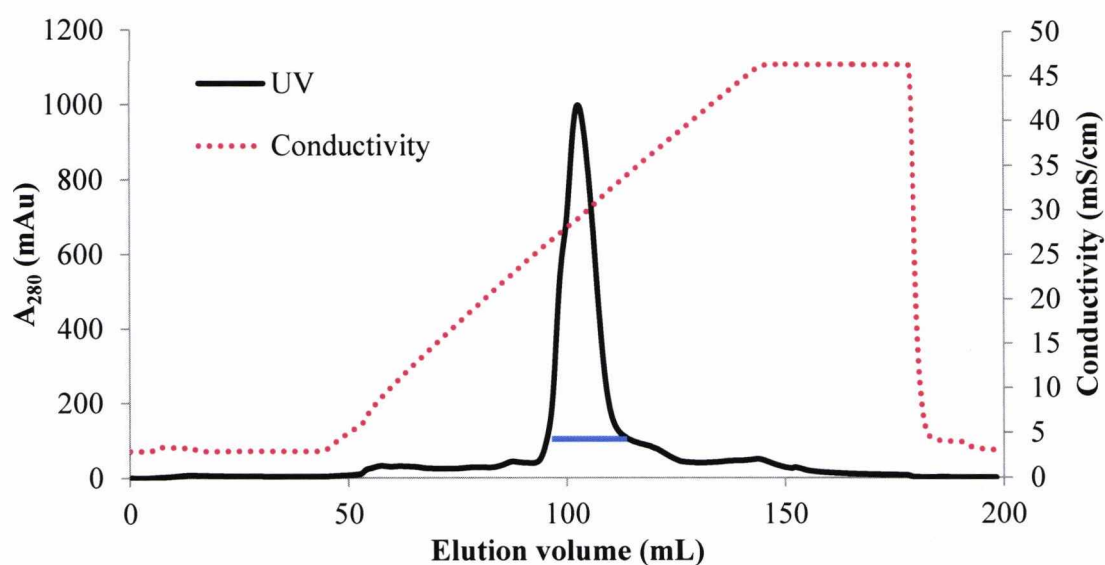


Figure 2.8 Anion-exchange chromatogram of WT **b'xa'c** using Source 30Q column. The absorbance at 280 nm is shown by the black solid line and the conductivity is shown by the red dotted line. The blue line across the peak represents fractions analysed in figure 2.9.

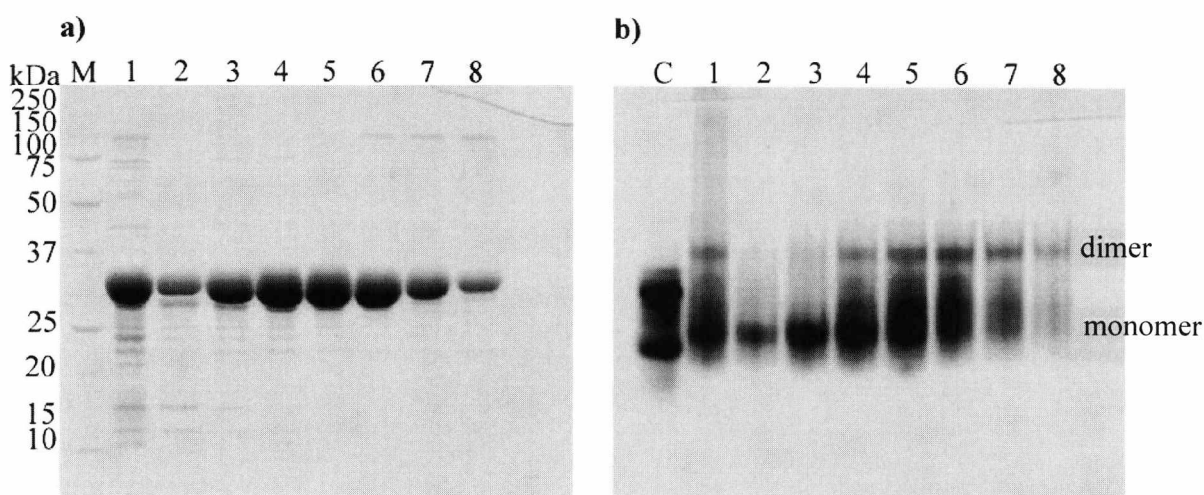


Figure 2.9 a) SDS-PAGE analysis of anion exchange fractions of WT b'xa'c. Lane 1: protein before anion exchange purification; lanes 2-8: fractions across the anion exchange peak. **b) native-PAGE of anion exchange fractions.** Control: b'x protein mixture of monomer and dimer; lane 1: protein before anion exchange purification; lanes 2-8: fractions across the peak.

Early fractions from the ion-exchange peak consisted of a species which had high mobility on native PAGE, whereas late fractions had low mobility. Fractions in the middle of the peak contained a mixture of the two species. However, fractions across the ion-exchange profile appeared homogenous on SDS-PAGE as well as electrospray mass spectrometry. Therefore, the forms observed on non-denaturing gels must reflect differences in conformation or oligomerization state. These forms are likely to be monomer and dimer.

2.3.7 Gel Filtration

Fractions across the ion-exchange elution profile were pooled and further purified on a Superdex 200 column. Preliminary gel filtration data from protein expression in LB showed that the majority of the protein appeared to be dimeric with very little monomer present (figure 2.10). This was a problem as monomer was needed preferentially for NMR experiments. But, protein expressed in minimal media, although presenting as a mixture of monomer and dimer contained a higher proportion of monomer than an LB expression.

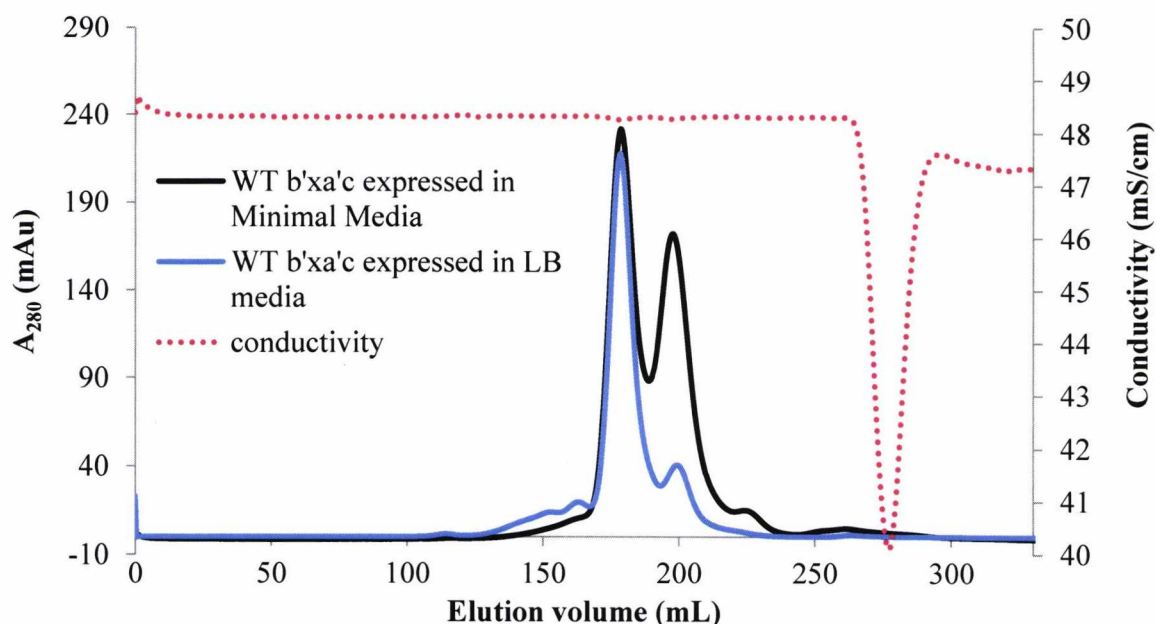


Figure 2.10 Gel filtration profile of WT **b'xa'c** using Superdex 200 media in a 300 mL column volume. Absorbance at 280 nm of protein expressed in LB (blue) and minimal media (black). The first peak represents the dimer and the second peak the monomer. Conductivity is shown in the red dotted line.

Although the exact proportions of monomer and dimer are dependent on expression conditions, it was evident from the beginning that in order to get enough monomer for the characterisation of **b'xa'c**, expression in minimal media had to be used instead of expression in LB. For this reason, all of the PDI proteins have been expressed in minimal media containing ^{14}N and ^{12}C when isotopic enrichment was not required.

2.3.8 Molecular weight confirmation by Mass Spectrometry

The predicted molecular weight of WT **b'xa'c** with the thiol groups of the **a'** active site in the reduced state is 32467.5 Da. Mass spectrometry analysis carried out on WT **b'xa'c** under reducing and non-reducing conditions confirmed that the gel filtration fractions of the monomer and dimer species contained protein that matched the exact molecular weight to that of WT **b'xa'c** in the oxidised state. The reduced protein was 2 Da heavier than protein without DTT which accounts for the addition of two hydrogen atoms upon reduction of the active site.

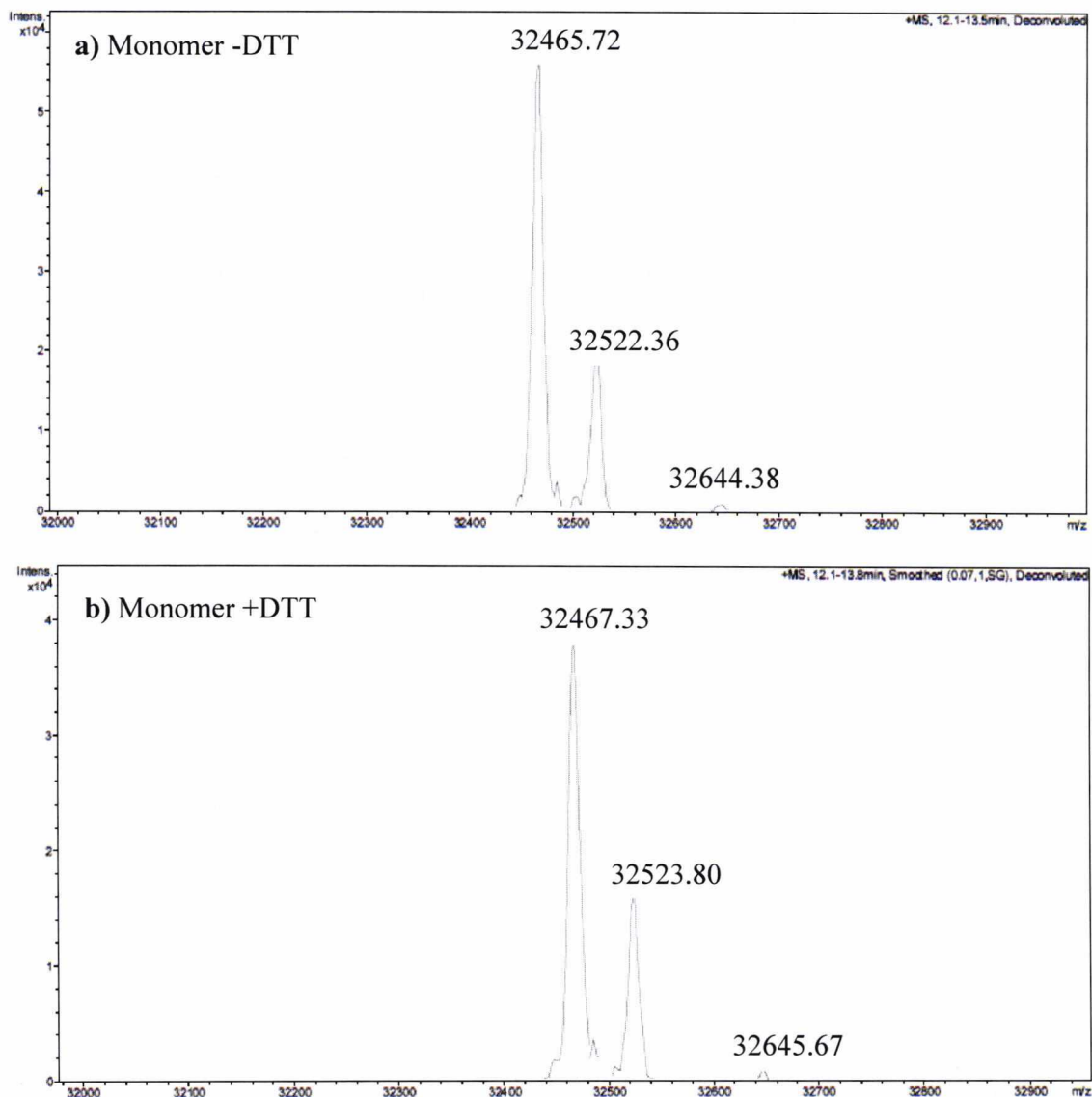


Figure 2.11 Mass Spectrometry analysis of WT **b'xa'c** monomer. a) Protein –DTT, b) protein +DTT. The expected mass of **b'xa'c** in the oxidised state is 32465 Da.

Mass spectrometry also showed that there were two additional species present in both of the protein samples. The first species were around 57 Da higher in mass than WT **b'xa'c** and the second around 178 Da higher. All of the **b'xa'c** mutants were subjected to mass spectrometry analysis to confirm their molecular weights. They all contained a small amount of the +57 Da modified species, but not the +178 Da species, in addition to the main species with the correct molecular weight in the oxidised state.

Mass spectrometry analysis of the dimer revealed that the dimerization was independent of the oxidation state of the protein as the WT **b'xa'c** dimer gave similar mass spectrometry data as the monomer.

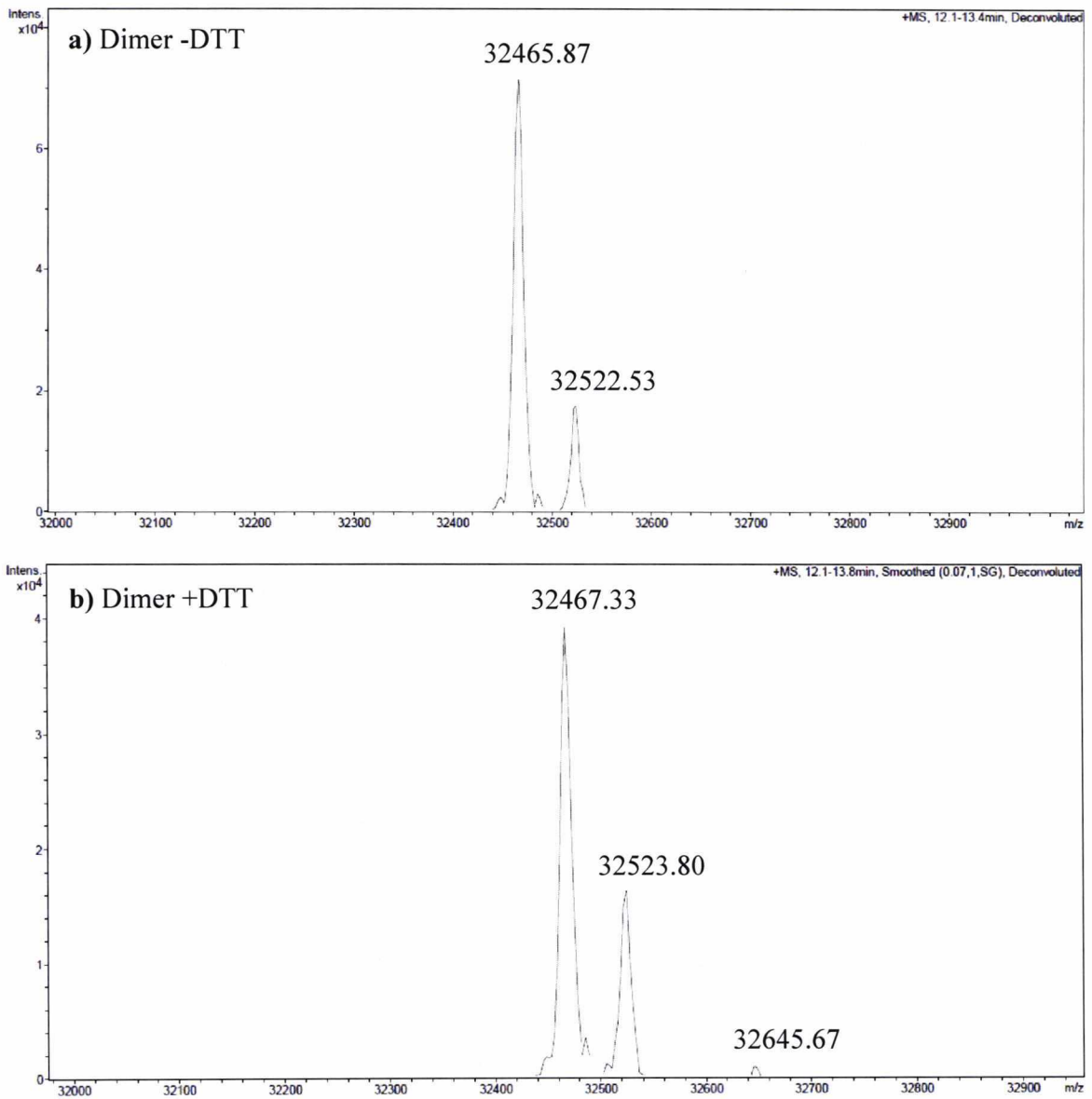


Figure 2.12 Mass Spectrometry analysis of WT **b'xa'c** dimer. a) Protein –DTT, b) protein +DTT. The expected mass of **b'xa'c** in the oxidised state is 32465 Da.

2.3.9 Characterisation of WT **b'xa'c** and mutants by gel filtration

Like WT **b'xa'c**, the I272A, L343A and 2DA mutants also presented as a mixture of monomer and dimer on gel filtration. Figure 2.13 shows a chromatogram of the gel filtration profiles of WT **b'xa'c** and its mutants. The mutants present with a very similar profile to the WT protein but with slightly different elution volumes suggesting different hydrodynamic volumes for each protein species. In terms of the monomer species for each protein 2DA was eluted first, followed by L343A, WT and lastly I272A. A similar pattern was followed by the dimer species with the exception of L343A and WT, which were eluted in exactly the same elution volume.

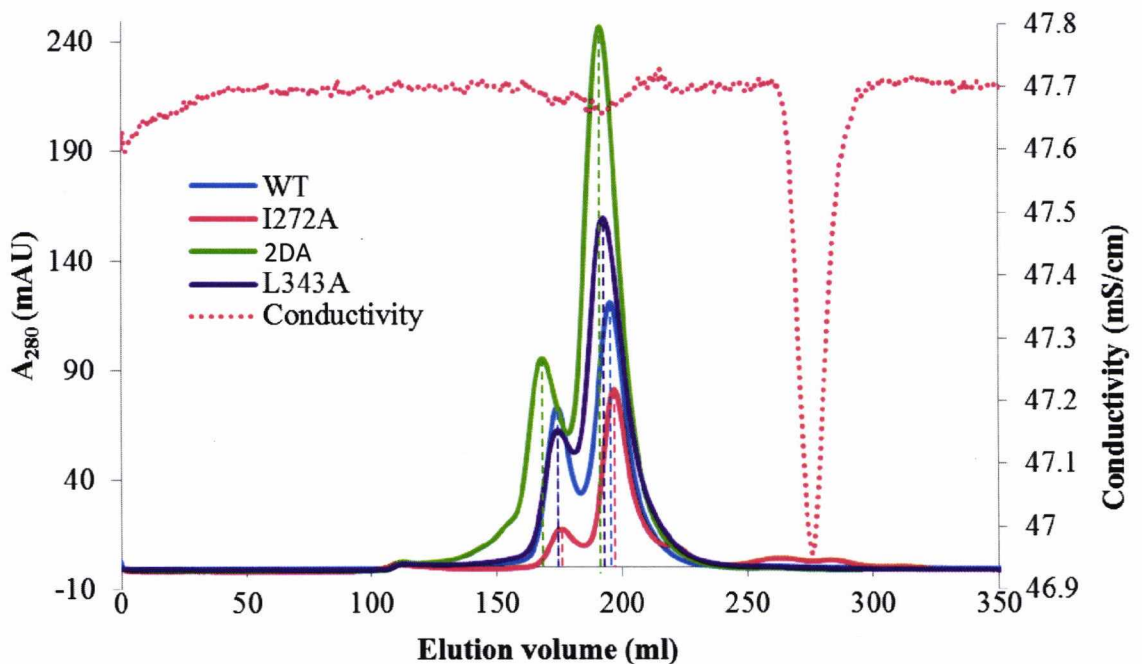


Figure 2.13 Size exclusion chromatography of **b'xa'c**. WT **b'xa'c** (blue), I272A (red), L343A (purple), 2DA (green). The absorbance at 280 nm is shown by the solid lines and the conductivity is shown by the red dotted line. The dotted lines represent the maxima of each elution peak in their respective colours.

SDS- and native PAGE analysis of fractions across the gel filtration peaks from WT **b'xa'c** was not very well resolved as there appeared to be a mixture of both species in all of the samples. Several attempts at running these proteins on native PAGE to separate the two conformers seen on gel filtration were unsuccessful. The SDS and native gels are shown in figure 2.14.

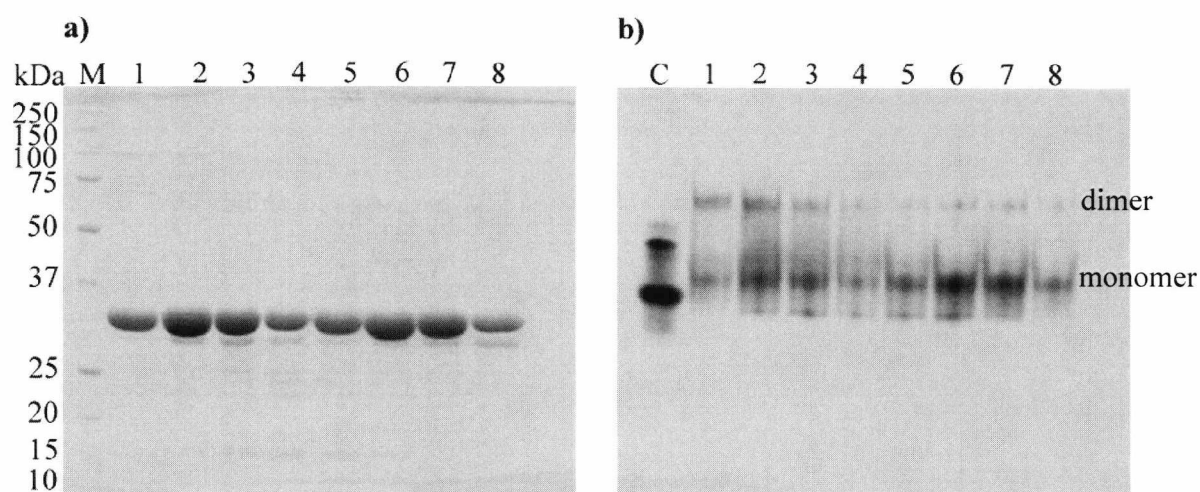


Figure 2.14 a) SDS-PAGE of WT **b'xa'c gel filtration fractions.** Lanes 1-8: fractions across the gel filtration peaks. **b) native PAGE of WT **b'xa'c** gel filtration fractions.** Control: **b'x** protein mixture of monomer and dimer; lanes 1-8: fractions across the gel filtration peaks.

As can be seen from figure 2.14, fractions across the WT **b'xa'c** gel filtration profile appear homogenous on SDS-PAGE with very few impurities. However, when the same fractions are run on native PAGE, two conformations appear. This agrees with the gel filtration chromatogram where two peaks are seen. But these conformers do not appear well separated on the native gel as there appears to be a mixture of monomer and dimer in the early fractions where there should be exclusively dimer. The late monomer fractions also appear to be eluted contaminated with some dimer although to a lesser extent. SDS and native gels for the **b'xa'c** mutants are not shown as they were very similar to the gels shown for WT **b'xa'c**.

The gel filtration column was calibrated using a set of standard proteins which allowed the determination of the molecular weights for the WT **b'xa'c** and the mutants. This was carried out by comparing an elution volume parameter, such as the gel phase distribution coefficient (K_{av}) of the protein of interest, with the values obtained from the known calibration standards. The molecular weight of the **b'xa'c** proteins could be determined from the calibration curve (plot of K_{av} versus $\log M_r$) once the K_{av} value had been calculated from the measured elution volume. K_{av} was calculated using equation 2.1.

$$2.1 \quad K_{av} = \frac{V_e - V_0}{V_c - V_0}$$

Where: V_e = Elution volume of the peak of interest

V_0 = Void volume as determined by the elution volume of Blue Dextran 2,000

V_c = Geometric column volume

Figure 2.15 shows the calibration curve with the standard proteins as well as the position of the monomer and dimer species for each of the **b'xa'c** proteins.

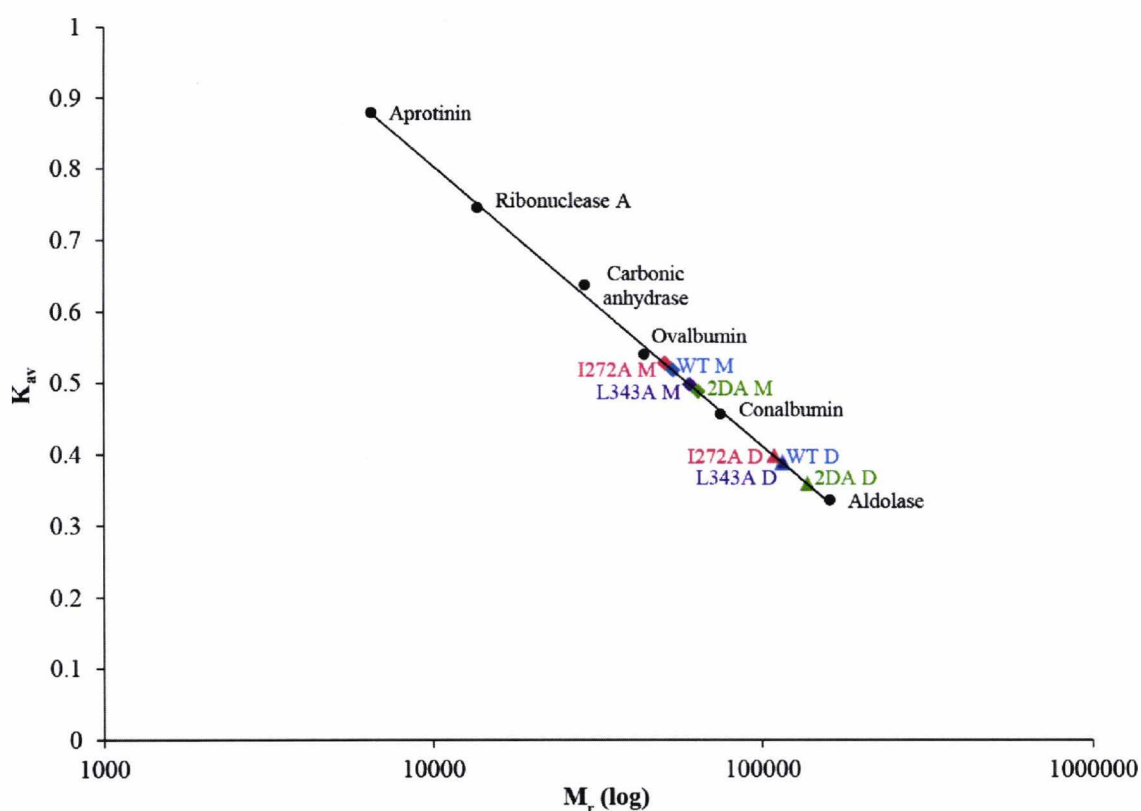


Figure 2.15 Gel filtration column calibration curve. The standard calibration proteins are represented by the black circles. The monomer (M) species of the **b'xa'c** proteins are displayed as squares and the dimers (D) as triangles with WT in blue, I272A in red, L343A in purple and 2DA in green.

The sizes calculated from the gel filtration column as well as the expected sizes are shown in table 2.7.

b'xa'c construct	Expected MW (Da)	Mass Spec data (Da)	Gel filtration data (kDa)	Oligomeric state
WT	32467.5	32465.7	54 116	Monomer Dimer
I272A	32425.4	32423.4	51 109	Monomer Dimer
L343A	32425.4	32423.8	61 116	Monomer Dimer
2DA	32379.5	32377.2	64 138	Monomer Dimer

Table 2.7 Expected and measured molecular weights of b'xa'c constructs. The expected molecular weight was calculated based on amino acid sequence. Mass spectrometry data was obtained by electrospray MS of the purified protein. The molecular weight obtained by gel filtration was calculated by comparison to protein standards as described above.

Table 2.7 shows that all of the **b'xa'c** constructs have been expressed in the oxidised state evident from the 2 Da difference of the mass spectrometry data compared to the expected molecular weight. Molecular mass calculations from gel filtration chromatography revealed that all of the **b'xa'c** constructs appear to have a larger hydrodynamic volume than expected for globular proteins of similar sizes. WT **b'xa'c** monomer was ~1.7 times bigger than expected whereas the dimer presented ~1.8 times bigger than its expected size. I272A appeared to have a smaller hydrodynamic volume than all of the other proteins, including WT **b'xa'c**, with the monomer ~1.6 and the dimer ~1.7 times bigger than their expected sizes. 2DA appeared larger in hydrodynamic volume than all of the proteins with the monomer and dimer species ~2 times bigger than expected. L343A monomer and dimer species were also ~1.9 and ~1.8 respectively bigger than expected. These calculated masses from gel filtration suggest that there is a possibility that the proteins may in fact be dimer-tetramer, rather than monomer-dimer. However, this will be discussed in the discussion section of this chapter.

2.3.10 Final protein yields

The overall protein yield for the expression of **b'xa'c** was determined using the A_{280} values from pooled ion-exchange fractions. The yield of material from expression in minimal media was calculated to be around 40 mg/L. The **b'xa'c** mutants had very similar expression yields to the WT. The proportions of monomer and dimer varied between expressions, especially for WT **b'xa'c**.

2.4 Discussion

Site-directed mutagenesis by overlap extension has been described to be 100% efficient as it is a simple method and eliminates the need for single-stranded DNA template and viral vector intermediates, therefore eliminating a cloning step (Ho et al., 1989). In this case, this was proved to be true as all the PCR steps leading to the ligation reaction were successful for all mutants, leading to the generation of inserts which could be cut with the appropriate restriction enzymes. All mutant inserts were successfully ligated into the empty vector and contained the correct nucleotide sequence, including the desired mutations. There were no other mutations in any of the DNA sequences. This was to be expected as a high fidelity enzyme, *Pfu* polymerase, was used in all PCR reactions, minimising the possibility of undesired mutations. The low error rate of *Pfu* DNA polymerase is about 1 in 10^6 base pairs duplicated as shown by direct measurement of PCR-generated point mutant DNA by denaturing constant capillary electrophoresis (Andre et al., 1997). This mutagenesis method was successful in introducing the desired point mutations in **b'xa'c** and generating the desired mutants.

All of the **b'xa'c** mutants were successfully expressed in *E. coli* BL21 (DE3) pLysS with expression yields similar to the WT protein. All **b'xa'c** proteins were soluble, as shown by their presence in the supernatant, and could be recovered by purification of the lysate by metal affinity chromatography using the N-terminal His tag. This first purification step proved to be very successful as it removed the majority of the impurities from the protein samples. Purification of the WT and mutant proteins by anion-exchange chromatography often gave rise to an asymmetric protein peak, indicating the presence of multiple species. Analysis of fractions from the anion-exchange profile by non-denaturing PAGE revealed that the proteins comprised of two forms which could be separated by gel filtration. After gel filtration, the two different conformers appeared as a single homogenous band on SDS-PAGE with the same molecular mass. This was true for all **b'xa'c** constructs. Mass spectrometry data revealed that both species had a molecular mass matching that of the WT **b'xa'c** of 32.47 kDa. The first peak from the gel filtration profile was believed to be the protein in the dimeric form, and the second was believed to be the monomer form. Another method, other than gel filtration, to confirm that the first peak was indeed a dimer would be to run the protein on native mass spectrometry so that the quaternary protein structure is preserved (Heck, 2008). Mass spectrometry data of WT **b'xa'c** also revealed the presence of two other species in the sample which had higher

molecular masses than WT **b'xa'c** at +57 and +178 Da. These are possibly due to post-translational modifications during expression inside the *E.coli* cells (Veniamin N. Lapko 2000). The +178 modification has been previously attributed to a His tag post-translational modification (Geoghegan et al., 1999). However, this modification did not interfere with the IMAC purification of the protein which utilised the His tag. However, the most likely explanation for the presence of these two extra species could be due to the presence of Ni^{2+} ions, with molecular mass ~59 Da, from the affinity purification. Therefore, the +57 Da species could be attributed to the presence of one Ni^{2+} ion, whereas the +178 Da species to the presence of three Ni^{2+} ions ($3 \times 59 = 177$). The Mass spectrometry analysis carried out in the presence of the reducing agent DTT revealed that the majority of the sample consisted of protein that was 2 Da heavier in mass than the protein without DTT. This 2 Da difference, from the non-reduced protein, accounts for the addition of 2 hydrogen atoms to the catalytic site cysteines upon reduction. The addition of DTT also confirmed that the species with the +57 Da modification was not a disulphide adduct.

The concentration of these modified species compared to that of **b'xa'c** is very small and therefore it is unlikely that they will have interfered with any characterisation experiments. Mass spectrometry analysis of the **b'xa'c** mutants confirmed that all of the mutants had the correct molecular weight for protein in the oxidised form. Just like the wild type protein they also included a very small amount of the +57 Da species, but less than WT, but none of the +178 Da modification.

The molecular sizes of the **b'xa'c** proteins could be calculated by comparison to a set of known standard proteins run on the same gel filtration. The calculated sizes revealed that the two conformers were monomer and dimer species. The molecular sizes of the two species were around 1.6 to 2 times higher than the actual size expected for the monomer and dimer, therefore one could assume that they are more likely to be dimer and tetramer. However, they are more likely to be monomer and dimer as a number of PDI constructs have been reported to give anomalously higher molecular sizes on gel filtration (Wallis et al., 2009). Full length PDI itself has been found as a mixture of monomer and dimer with anomalously higher molecular masses on gel filtration with the monomer with an apparent size of 116 kDa and the dimer eluting at over 200 kDa (Wallis et al., 2009, Li et al., 2006). This is thought to be due to the structure of PDI where the four domains adopt a non-spherical conformation. The unstructured C-terminal tail of PDI has been reported to contribute to the anomalously high dynamic volume and the elongated shape of the

molecule adds to this discrepancy (Li et al., 2006). Therefore, it is not surprising that **b'xa'c** also presents with these abnormal high masses on gel filtration. Other constructs such as **b'x** and **bb'x** also present with high molecular masses on gel filtration with **b'x** monomer and dimer having apparent masses of 23 kDa and 49 kDa, around 1.4 times larger than expected, and **bb'x** monomer and dimer species with apparent masses of 35 kDa and 77 kDa, 1.3 and 1.5 times larger than expected respectively (Wallis et al., 2009). Ultracentrifugation has been previously used to determine the correct masses of these proteins, but another method to confirm the oligomeric state of PDI constructs is small angle X-ray scattering, SAXS, which reveals information about the shape and size of molecules. PDI constructs containing the **b'** domain have been previously reported to form dimers. More specifically, **b'** and **bb'** present exclusively as homodimers, whereas **b'x** and **bb'x** form a mixture of monomer and dimer (Byrne et al., 2009). Therefore, the **b'** domain has been described to drive oligomerisation via hydrophobic interactions (Wallis et al., 2009).

The I272A mutant appeared to have a smaller hydrodynamic volume not only than L343A and 2DA, but also than WT **b'xa'c** as seen in table 2.7. WT **b'xa'c** presented with a larger hydrodynamic volume than I272A, followed by L343A and 2DA which was the largest of the proteins according to gel filtration data. This must be attributed to a difference in the conformation of the proteins. In **b'x**, the I272A and 2DA mutants have been shown to stabilise the capped conformation in which the **x** linker region is found to be bound in the hydrophobic binding pocket on the surface of the **b'** domain (Nguyen et al., 2008). Therefore, if the same capping event was taking place in the **b'xa'c** construct, we would expect these two mutants to behave in a very similar manner. However, the 2DA mutant has a larger hydrodynamic volume on gel filtration suggesting that it behaves differently to the I272A and L343A mutant. The L343A mutant in **b'x** has been found to favour the uncapping of the ligand binding site by **x**. So, if this mutant were to have the same effect in **b'xa'c** we would expect it to have be slightly larger in size than the capped mutants, which should be more compact in structure due to capping. As WT **b'xa'c** has a hydrodynamic volume that is larger than the capped I272A mutant but smaller than the L343A uncapped mutant, one could hypothesise that WT **b'xa'c** could be a mixture of capped and uncapped protein. However, it is not possible to determine the conformation of **x** just by comparing gel filtration data and further investigations into the WT **b'xa'c** and its mutants has been carried out. The monomeric state of **b'xa'c**, **b'x**, and **xa'c** proteins has

been confirmed by NMR in chapters 4 and 5, where the rotational correlation time, τ_m , of the proteins shows that they tumble in solution at a rate that is consistent with molecules of such sizes in the monomeric form.

In 2006, Li *et al.* reported that the **b'xa'c** construct eluted from gel filtration as two peaks with apparent molecular masses of around 700 kDa, presumably a higher order disulphide-bonded oligomer, and 62 kDa consistent with a dimer (Li *et al.*, 2006). However, growth conditions of the *E. coli* expressing the PDI proteins are unclear from the paper, but it is likely that expression was carried out in rich media rather than minimal media as no isotopic labelling was required for their experiments. This supports the gel filtration results obtained from growth in LB which show that the majority of WT **b'xa'c** is in the dimeric form. Studies have shown that cells grown in LB media with a good supply of carbon and energy, grow rapidly with a 20 min generation time. As a consequence their expression of genes involved in protein synthesis is elevated. Whereas cells in minimal media grow on a single carbon and energy source, in this case glucose, from which they have to synthesise metabolic components including all of the amino acids from scratch. Consequently, growth is significantly slower with a generation time of around 60 min. To deal with this burden, the genes responsible for stress tolerance are up-regulated and the cell is protected from the stress of living in a self-formed acidic environment due to the production of acetate (Tao *et al.*, 1999). With this in mind, it is likely that cells grown in minimal media expressed protein at a slower rate and therefore leading to diminished dimerization of **b'xa'c** compared to expression in LB. Also, it is likely that growth in minimal media reduces the rate of protein synthesis therefore the protein has the opportunity to fold without interference. In contrast, when **b'xa'c** is expressed rapidly in LB it is most likely to dimerise due to the close proximity to other folding molecules.

CHAPTER 3

Biophysical characterisation of WT b'xa'c and mutants by intrinsic fluorescence and limited proteolysis

3.1 Introduction

3.1.1 Intrinsic fluorescence

In fluorescence spectroscopy, molecules under investigation have a ground low energy electronic state and a higher energy excited electronic state. Vibrational states exist within each of these electronic states. Absorption of radiation occurs from the lowest vibrational level of the ground state to a higher vibrational level of the excited state. Then, there is a loss of vibrational energy to give the lowest vibrational level of the excited state. Fluorescence is the emission of radiation from the lowest vibrational energy level of the excited electronic state of a molecule as it decays to the ground state. The principle of fluorescence is shown in figure 3.1 (Price and Nairn, 2009).

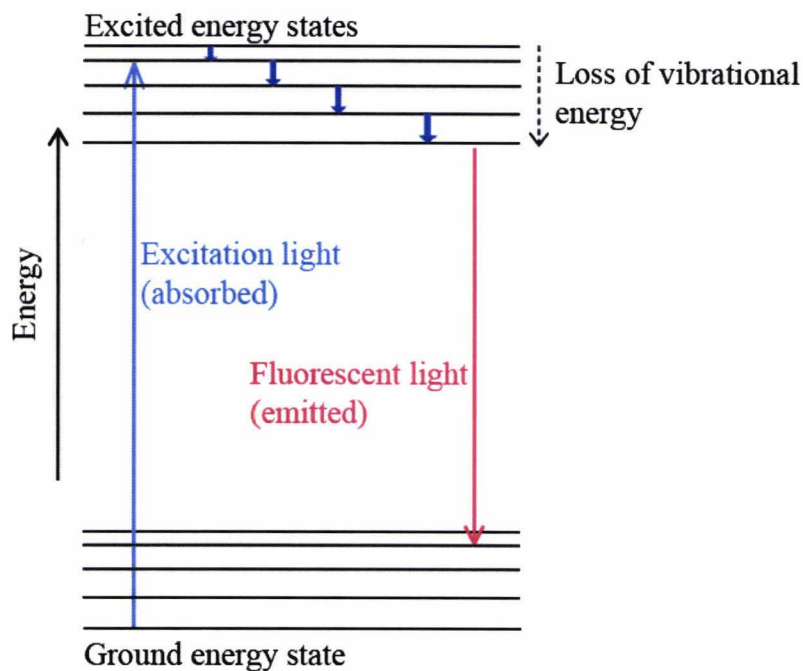


Figure 3.1 The process of absorption of energy and emission of fluorescence. Absorption of radiation, or excited light (blue arrow), promotes an electron from the ground state to a higher vibrational level of the excited state. This is followed by a loss of vibrational energy to the lowest vibrational level of the excited state. Fluorescence is the emission of this radiation to the ground energy state (red arrow).

Intrinsic fluorescence has been used for many years to monitor structural changes in proteins as well as measure their conformational stability. This is brought about by the chemical properties of aromatic or highly conjugated molecules or groups, and in the case of proteins those of tryptophan, tyrosine and phenylalanine side chains. The fluorescence emitted from tryptophan is much stronger than that from the other two aromatic amino acids and therefore dominates the emission spectra of proteins. Tyrosine can contribute to fluorescence at 280 nm as it is usually present in high numbers in proteins. However its fluorescence can be quenched through energy transfer to a nearby tryptophan when the tyrosine is in the excited state. Tyrosine quenching can also occur from loss of the proton on the aromatic hydroxyl group through ionisation in the excited state. Phenylalanine comprises of only one benzene ring and a methylene group, so it is only weakly fluorescent. Occasionally, tryptophan can become quenched by neighbouring protonated acidic groups, for instance aspartic acid or glutamic acid, resulting in a decrease in fluorescence intensity.

It has been known for many years that tryptophan fluorescence is very sensitive to the polarity of its local environment, with the wavelength of maximum emission, λ_{max} , ranging from 308 nm (azurin) to 355 nm (glucagon) (Vivian and Callis, 2001). The wavelength of maximal fluorescence, λ_{max} , is an excellent measure of polarity as it is responsive to the extent of which a tryptophan side chain is exposed on the surface of a protein so accessible to the surrounding aqueous solution, or buried in a protein's hydrophobic core. A blue-shift in fluorescence is characteristic of the tryptophan residue buried in the hydrophobic core of the protein and the spectrum shifts to shorter wavelength as the polarity of the solvent decreases, whereas a red-shift represents the exposure of the tryptophan side chain to the aqueous solvent on the surface of the protein. Excitation of a protein at 280 nm is dominated by the tryptophan fluorescence with very little contribution from tyrosine and essentially none from phenylalanine, so this wavelength is often used to follow changes in tryptophan emission spectra in denaturation studies. Tryptophan is a relatively rare amino acid and most proteins contain only one or few tryptophan residues in their sequence. In the case of a protein with multiple tryptophans, the emission spectrum is the sum of the individual tryptophan emissions. This extra complication can be circumvented by mutating all but one of the tryptophan residues so that the fluorescence originates from a single fluorophore. Conversely, a tryptophan residue can be mutated into the core of a protein lacking the fluorophore to allow fluorescence studies. Extrinsic probes can be attached as an alternative approach to introduce fluorescent properties in a protein.

3.1.2 Determination of conformations of PDI by intrinsic fluorescence

Intrinsic fluorescence has proved to be an invaluable tool in recent PDI studies owing to the presence of tryptophan residues in the catalytic **a** (W35 and W311) and **a'** (W379 and W390) domains, and most importantly in the **x** linker region (W347) which act as reporters on environment changes. Fluorescence, in conjunction with NMR studies, has assisted investigations into the conformational diversity of the **x** linker region in a number of constructs including full length PDI itself.

Studies of the PDI fragment containing the ligand binding domain and the 19 residue **x** linker, **b'x**, have shown that **x** can exist in at least two different conformations in

solution. In one of the conformations, **x** is found bound to a hydrophobic patch on the surface of the **b'** domain, shown to be the ligand binding site (Byrne et al., 2009, Nguyen et al., 2008). This has been termed the “capped” conformation. In the alternative “uncapped” conformation, **x** is found to be away from the ligand binding site and free in solution. The unique tryptophan residue in **x**, W347, allows tracking of this “capping” event by NMR and intrinsic fluorescence. Excitation at 280 nm of WT **b'x** in the capped conformation is characterised by a blue-shift in fluorescence emission spectra with maximum fluorescence around 334 nm. This is indicative of the hydrophobic environment of the **x** linker capping the ligand binding site. Conversely, emission spectra of **b'x** in the uncapped state are distinguished by a red-shift with maximum fluorescence around 355 nm suggestive of the hydrophilic environment of the tryptophan in solution. NMR is also a very efficient tool that has been used to track capping of the ligand binding site by **x**, but this will be described in more detail in subsequent chapters. Mutants of **b'x** that favour the capped and uncapped conformations were selected for further investigations into the mechanism of capping. A screen of **b'x** mutants revealed that mutations I272A in **b'** and D346A/D348A (2DA) in **x** favoured the capped conformation. This was further confirmed by the crystal structure of I272A **b'x** shown in figure 3.2., where **x** is seen capping the hydrophobic site on **b'**. Whereas the L343A mutation in the **x** linker region has been shown to favour uncapping of the ligand binding site by **x** (Nguyen et al., 2008).

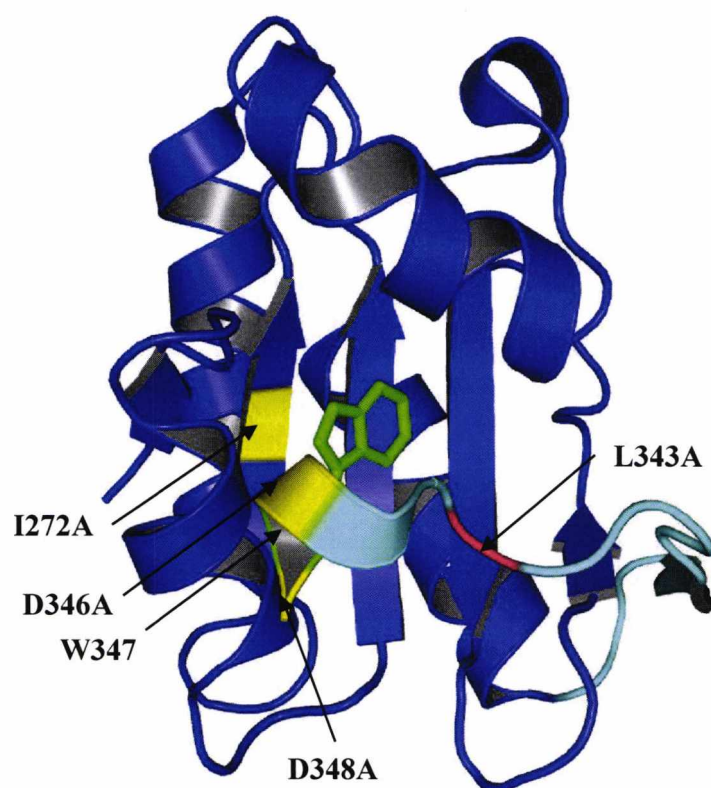


Figure 3.2 Crystal structure of I272A **b'x** (PDB accession code 3BJ5). The **x** linker region (cyan) is shown capping the **b'** domain (blue). The unique W347, which allows tracking of the **x** region by fluorescence, is shown in **green** with its side chain buried into the hydrophobic core of the **b'** domain and therefore displaying a blue-shift in fluorescence. The arrows point to the mutations that favour the capped (I272A and 2DA in yellow) and the uncapped (L343A in red) conformations.

The orientation of **x** in yeast PDI adopts a different conformation to its human homologue as neither of the two published structures for full length yeast PDI show capping of the ligand binding site by **x** (Tian et al., 2008, Tian et al., 2006). However, fluorescence emission spectra have been collected for full length human PDI, where W111F and W390F mutations were introduced so that W347 was the primary emitter of fluorescence. Excitation at 280 nm revealed that **x** was present in two conformations, represented by different fluorescence maxima in the spectra. The I272A, 2DA and L343A mutations appeared to have a similar effect on the conformation of the **x** linker in full length PDI as they did in **b'x**, though not to the same extent as the difference in fluorescence on blue and red-shifting is smaller in full length PDI than that observed in

b'x. Evidence of capping has also been seen in the fluorescence of **bb'x** and **abb'x** of PDI but there have been no fluorescence studies on **b'xa'c** to date to explore the capping event when **x** is tethered by both its neighbouring domains **b'** and **a'** (Byrne et al., 2009, Wang et al., 2010).

In this chapter, the fluorescence properties of various PDI constructs will be compared to those of WT **b'xa'c** and mutants in order to determine the effect of the **a'** domain on the conformation of the **x** linker region.

3.1.3 Determination of the conformational stability of b'xa'c

The three dimensional structure of a protein is essential for its biological function, but the unfolded state is equally as important as the native state when determining conformational stability. The denatured state is a form where the protein has lost its biological activity due to the disruption of its secondary and tertiary structure. It is well known that the most frequently used denaturants to study stability and unfolding pathways of proteins are urea and guanidine hydrochloride (GdmCl). Although they have been used extensively throughout many years, their mode of action on protein conformation is still poorly characterised. It is thought that they may bind directly to the biomolecule thereby weakening hydrogen bonds within the protein, or indirectly induce denaturation by altering the solvent environment and weakening the hydrophobic effect (Schellman, 2002, Vanzi et al., 1998).

Unfolding can be followed by a variety of techniques including biological activity, NMR, intrinsic fluorescence and circular dichroism (CD). Fluorescence is one the primary ways of following unfolding as it requires the least amount of material, whereas NMR and CD are more expensive techniques. For this reason, GdmCl denaturation of PDI constructs was followed by intrinsic fluorescence at 280 nm.

3.1.4 Determination of conformational flexibility of b'xa'c by proteolysis

Recently, the conformational plasticity of full length PDI as well as that of a number of PDI constructs has been studied by limited proteolysis. Susceptibility to proteases such as trypsin, chymotrypsin and proteinase K has been used as a measure of

flexibility, as flexible regions of the protein are more susceptible to proteolytic cleavage than those parts of the structure that are rigid (Wang et al., 2010). Contrary to the yPDI structure, where the N-terminal half of the protein is more flexible than the C-terminal half, in hPDI the C-terminal region has been shown to be more susceptible to proteolytic cleavage and therefore considered more flexible (Tian et al., 2008, Wang et al., 2010). The oxidation state of the **a** and **a'** active sites has also been shown to contribute to the flexibility of hPDI as reduction of the disulphide bond in the **a'** domain leads to a conformational change that reduces the protein's accessibility to proteases (Wang et al., 2012b).

The susceptibility of WT **b'xa'c** to proteinase K, in the presence and absence of DTT, was compared to that of the I272A, L343A and 2DA mutants to determine the effect of the mutations on the conformational flexibility of this protein.

3.2 Materials and Methods

3.2.1 Sample preparation

Protein production and purification was carried out as described in section 2.2.2. After gel filtration, fractions corresponding to the monomer and the dimer were pooled accordingly. The protein was concentrated and buffer-exchanged using 50 mM Tris-HCl, 150 mM NaCl, pH 7.6 to make a 0.5 mM stock which was used to make 5 μ M 1 mL samples for fluorescence.

3.2.2 Fluorescence studies to determine the conformation of **x**

Spectra of PDI constructs were collected on a Cary Eclipse fluorimeter, Scan Software Version: 1.1(132), using a 1 mL quartz cuvette at 25°C, in order to determine the effect of the mutations on the conformations of the **x** linker region in the **b'xa'c** construct. Spectra were collected for the monomer and dimer species of **b'xa'c**, I272A, L343A, 2DA in the W390F and WT backgrounds. For comparison with published data, proteins were excited at 290 nm and emission spectra (average of 5 scans) were collected from 310 – 400 nm at 0.5 nm intervals. Both emission and excitation slits were set at 5 nm. Blank samples, where the protein was omitted, were run and their fluorescence spectra was subtracted from that of the protein samples to ensure that the fluorescence measured was representative of the protein with no interference from the buffer. The wavelength of maximum fluorescence (λ_{\max}) was used as the measured parameter, but the barycentric mean emission wavelength (λ_m) was also calculated as an alternative to the λ_{\max} . The full emission data from 310 to 400 nm was used to calculate the barycentric mean emission wavelength (λ_m), the “centre of gravity” of the spectrum, which reduced the entire spectrum for each protein to a single value using equation 3.1. λ_m is not sensitive to changes in concentration and was used as an objective measure of spectral shift on the mutants (Chalton and Lakey, 2010).

$$3.1 \quad \lambda_m = \frac{\sum(F(\lambda) \times (\lambda))}{\sum F(\lambda)}$$

Where $F(\lambda)$ is the fluorescence intensity at wavelength λ .

3.2.2.1 Determination of conformational stability by GdmCl denaturation

For analysis of denaturation studies, the parameter examined was λ_{\max} , which is insensitive of protein concentration and enables comparison with other published fluorescence data (Nguyen et al., 2008, Wang et al., 2010). 5 μM protein samples in 50 mM Tris, 150 mM NaCl, pH 7.6 with varying concentrations of GdmCl, 0 to 6 M, were incubated at room temperature for 1 h prior to data collection. Spectra were collected as an average of 4 or 5 scans, with excitation at 280 nm and emission at 300 – 400 nm. Slit widths were set at 5 nm for excitation and 10 nm for emission. Data were collected at a scan rate of 300 nm/min at 0.5 nm intervals, or 600 nm/min at 1.0 nm intervals. Control samples were run for each GdmCl concentration and these spectra were subtracted from the protein emission spectra to minimise the interference of the buffer and GdmCl with the protein fluorescence data.

3.2.2.2 Determination of the free energy of denaturation ΔG_{app}^0

Changes in the fluorescence spectra of the proteins were used to monitor the unfolding where the wavelength of maximum fluorescence (λ_{\max}) was plotted for each denaturant concentration. Data analysis was carried out assuming a two-state model of denaturation where the stability of the fully folded native protein (N), relative to that of the denatured (D) is described in terms of the standard free energy, ΔG_{app}^0 , where K_D is the equilibrium constant between these two states:



The fraction of protein which was unfolded, f_D , was calculated at each concentration of the denaturing agent, GdmCl, using equation 3.3 assuming a two-state model where y is the measured parameter (λ_{\max}), y_0 is the baseline value for the parameter for folded protein and y_{\max} is the maximum value of λ_{\max} when the protein is fully unfolded (Kurtin and Lee, 2002).

$$3.3 \quad f_D = \frac{y - y_0}{y_{\max} - y_0}$$

Figure 3.3 shows a simple representation of a denaturation illustrating the terms used to calculate fraction unfolded, f_D .

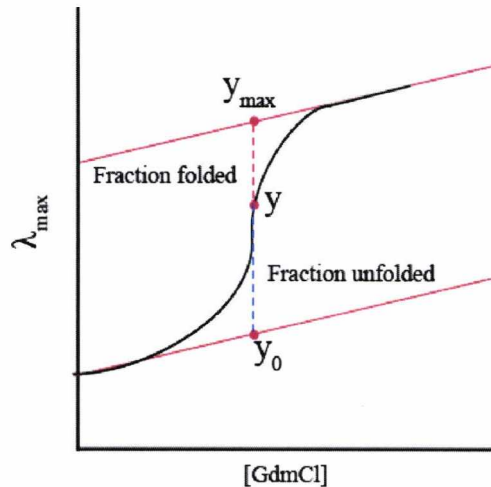


Figure 3.3 A simple schematic of a denaturation curve illustrating the terms used for the calculation of fraction unfolded.

Consequently this allowed the apparent equilibrium constant, K_D , to be calculated for each GdmCl concentration where:

$$3.4 \quad K_D = \frac{f_D}{1 - f_D}$$

Protein denaturation data usually exhibit a sigmoid dependence of the measured parameter on the concentration of denaturant. Assuming that there is an equilibrium constant between the native and the denatured state, the data (GdmCl concentration versus fraction unfolded) were fit with the Hill equation in which a is the difference between the maximum value of f_D when the protein is fully unfolded and the baseline value ($y_{\max} - y_0$), x is the denaturant concentration, b is the Hill coefficient necessary to fit the data and c is the interaction constant which is equal to the denaturant concentration where $\Delta G = 0$ and corresponds to the midpoint of denaturation (Kurtin and Lee, 2002).

$$3.5 \quad y = y_0 + \frac{ax^b}{c^b + x^b}$$

The free energy required for the unfolding of the proteins was determined by extrapolation of the line of best fit in the graph of ΔG_{app}^0 versus denaturant concentration and confirmed using the calculated K_D value and equation 3.6.

$$3.6 \quad \Delta G^0 = -RT \ln K_D$$

Where ΔG_{app}^0 is the change in standard Gibbs free energy at each denaturant concentration, R is the gas constant ($8.31 \text{ J K}^{-1} \text{ mol}^{-1}$), T is the temperature and ln is the natural logarithm. The ΔG value at 0 denaturant concentration was obtained using the linear extrapolation method (Ibarra-Molero and Sanchez-Ruiz, 1997). The midpoint of denaturation, where the proportion of the unfolded protein is equal to the fully folded state, was obtained from the plot of ΔG_{app}^0 against denaturant concentration where $\Delta G_{\text{app}}^0 = 0$.

3.2.3 Limited Proteolysis

Limited proteolysis was carried out using proteinase K. After testing a range of concentrations, 2 $\mu\text{g/mL}$ of enzyme was used on 2 mg/mL of protein. Reactions were allowed to take place for 1 h in a 25°C water bath. Proteolysis was stopped by flash-freezing samples in liquid nitrogen or adding reducing sample buffer and boiling immediately for SDS-PAGE analysis. Proteolysis was carried out in the presence and absence of DTT in order to determine the effect of the oxidation state on conformational flexibility of proteins.

3.3 Results

3.3.1 Determination of capping in **b'xa'c** by fluorescence

Previous studies on **b'x** and **bb'x** have exploited the fluorescence properties of the tryptophan residue in the **x** linker region as a sensitive probe to the alternative conformations of **x** with respect to **b'**, where **x** can be bound to the ligand binding site in the capped conformation or free in solution in the uncapped conformation (Nguyen et al., 2008). The capped conformation is characterised by a blue-shift in fluorescence whereas the uncapped form presents with a red-shift in fluorescence. Isolated **b'x** mutants, I272A and 2DA, favour the capped conformation and consequently their emission maxima are at a shorter wavelength than L343A in the uncapped form. In order to determine if such conformations occur in **b'xa'c**, fluorescence emission spectra were collected for the monomer and dimer species of the WT protein as well as the I272A, L343A and 2DA mutant proteins (figure 3.4). The measured parameter for all of the emission spectra was the wavelength of maximum fluorescence emission (λ_{\max}) as it is insensitive to small differences in protein concentration and has been shown to be an excellent indicator of the conformation of the **x** linker region (Nguyen et al., 2008). The excitation wavelength was set at 290 nm for comparison with published data (Wang et al., 2010).

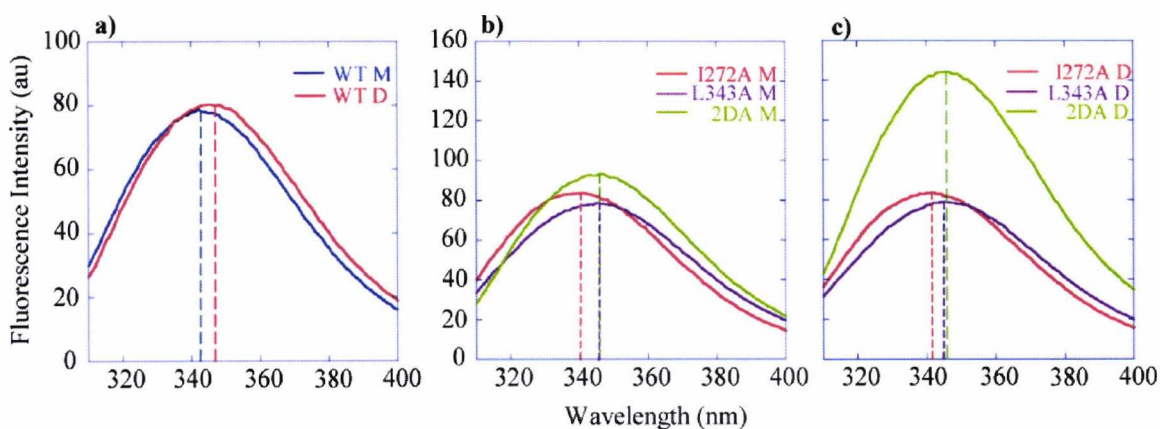


Figure 3.4 Intrinsic fluorescence emission spectra of WT **b'xa'c** and mutants with excitation at 290 nm. **a)** emission spectra of WT **b'xa'c** monomer and dimer, **b)** emission spectra of the monomers of **b'xa'c** mutants I272A, L343A and 2DA, **c)** emission spectra of the dimer species of **b'xa'c** mutants I272A, L343A and 2DA. The dashed lines indicate the maximal emission wavelength (λ_{\max}).

As shown in chapter 2, **b'xa'c** consisted of a mixture of monomer and dimer which could be separated by gel filtration. Excitation of the oligomeric states at 290 nm showed that they differ in their fluorescence properties with the dimer being more red-shifted than the monomer. This was also true for I272A, L343A and 2DA mutants as can be seen in figure 3.4. The fluorescence properties of the mutants also differed not only from the WT but also from each other. As expected, I272A was more blue-shifted than the WT and the other two mutants, suggesting that the tryptophan side chains in this mutant were in a more hydrophobic environment; consistent with the capped conformation. L343A appeared more red-shifted than the WT and I272A suggesting that this mutant was in the uncapped conformation. The 2DA mutant, which in **b'x** promotes capping of the ligand binding site by **x**, appeared more red-shifted than WT but very similar to the uncapped L343A mutant. This suggests that any mechanism of capping promoted by 2DA must be different to that promoted by I272A. As 2DA fluorescence spectra are very similar to those of L343A, the presence of the **a'** domain has disrupted capping of the ligand binding site by **x** in the 2DA mutant of **b'xa'c**.

However, there are three tryptophan residues in **b'xa'c**, W347 in the **x** region and W379 and W390 in the **a'** domain, which means that the emission spectra above showed the sum of the fluorescence from the environments of the tryptophans in **a'** and **x**, making it difficult to determine individual environments. In full length PDI, the fluorescence of W35 has been shown to be heavily quenched by the adjacent active site of the **a** domain in both the disulphide and dithiol states (Lappi et al., 2004). Therefore it was considered safe to assume that the fluorescence of W379, next to the **a'** active site, was also quenched and did not contribute significantly to the emission spectra of **b'xa'c**. Hence the W390F mutation was introduced into the WT **b'xa'c** and the I272A, L343A and 2DA mutants to ensure that emission spectra were dominated by the fluorescence of W347 and the conformation of the **x** linker region could be tracked. Emission spectra of WT **b'xa'c**, I272A, L343A and 2DA in the W390F background are shown in figure 3.5.

Fluorescence of the W390F mutants presented with a similar pattern to that of the proteins in the WT background, suggesting that even with contribution in fluorescence from the **a'** tryptophan, W347 was the main fluorophore in **b'xa'c**. However, L343A in the W390F background, both monomer and dimer species, appeared more blue-shifted than in the WT background, making it more similar to I272A than 2DA.

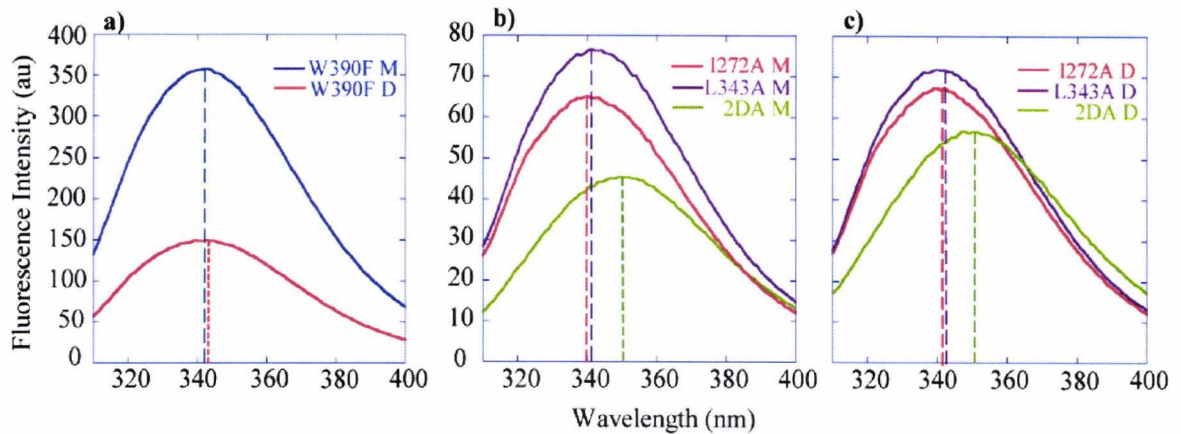


Figure 3.5 Intrinsic fluorescence emission spectra of WT **b'xa'c** and mutants in the **W390F** background with excitation at 290 nm. **a)** emission spectra of **W390F b'xa'c** monomer and dimer, **b)** emission spectra of the monomers of **W390F b'xa'c** mutants **I272A**, **L343A** and **2DA**, **c)** emission spectra of the dimer species of **W390F b'xa'c** mutants **I272A**, **L343A** and **2DA**. The dashed lines indicate the maximal emission wavelength (λ_{\max}).

There is a significant quenching of fluorescence when the dimer is formed as shown by the decrease in fluorescence intensity in figure 3.5a. This could be due to the displacement of the tryptophan from the hydrophobic patch in the monomer in order to form the dimer species.

The wavelength of maximum fluorescence (λ_{\max}) for each of the proteins has been summarised in table 3.1. The barycentric mean emission wavelength (λ_m) was also calculated as an alternative to the λ_{\max} as described in section 3.2.2.

b'xa'c protein	λ_{\max} (nm)	Shifts in λ_{\max} (nm)	Barycentric mean wavelength (λ_m)	Shift in λ_m (nm)
WT	M 342		349.30	
	D 345		350.89	
I272A	M 340	-2	347.35	-1.95
	D 342	-3	348.24	-2.65
L343A	M 345	+3	350.29	+0.99
	D 346	-1	350.77	-0.12
2DA	M 345	+3	350.96	+1.66
	D 346	-1	351.22	+0.33
W390F	M 342		348.80	
	D 343		348.70	
I272A/W390F	M 340	-2	348.39	-0.41
	D 342	-1	348.18	-0.52
L343A/W390F	M 341	-1	348.27	-0.53
	D 343	0	349.00	+0.30
2DA/W390F	M 351	+9	353.46	+4.66
	D 351	+8	353.45	+4.75

Table 3.1 Intrinsic fluorescence properties of WT **b'xa'c and mutants, monomer (M) and dimer (D) forms.** Shifts in λ_{\max} and λ_m for the mutants are the values subtracted from the λ_{\max} or λ_m of the corresponding background proteins. – and + correspond to blue-shift and red-shift, respectively.

The I272A mutant showed a blue-shift in fluorescence with maxima at 340 and 342 nm for the monomer and dimer species, respectively, whereas L343A appeared red-shifted with maxima at 345 and 346 nm, suggesting that the I272A is in the capped conformation and the L343A in the uncapped. This agrees with published **b'x** data where these mutations had similar effects on the fluorescence properties of W347. In contrast with **b'x**, the 2DA mutation in **b'xa'c** seemed to be red-shifted supporting an uncapped state. This was more accentuated when the W390F mutation was introduced and the emission maximum shifted from 345, in the WT background, to 351 nm in the W390F background. Overall, the W390F mutants confirmed the effect of the mutations in **b'xa'c** were the same as when the **a'** tryptophan was not mutated but smaller in magnitude (table 3.1). WT **b'xa'c** monomer presented with a λ_{\max} at 342, which when compared to I272A at 340 and L343A at 345, suggested that a mixture of the capped and uncapped conformations made up WT **b'xa'c**.

The barycentric mean emission wavelength also showed a similar pattern of shifts in fluorescence but moved to a higher wavelength than the λ_{\max} and the differences were even smaller between the mutants and the WT. Therefore this method was not considered the most suitable method to analyse this data. The barycentric mean wavelength is not objective and but it is dependent on the shape of the curve, therefore a broad asymmetric spectrum will skew the result.

3.3.2 Determination of conformational stability of PDI constructs

Denaturation was carried out by incubating the proteins in varying concentrations of GdmCl for 1 h at room temperature prior to data collection. Protein samples were excited at 280 nm and fluorescence emission spectra were used for the determination of the λ_{\max} at each concentration of denaturant in the presence and absence of the reducing agent DTT. Figure 3.6 demonstrates the shift in wavelength and fluorescence intensity that occurs in WT **b'xa'c** in the presence of 6 M GdmCl.

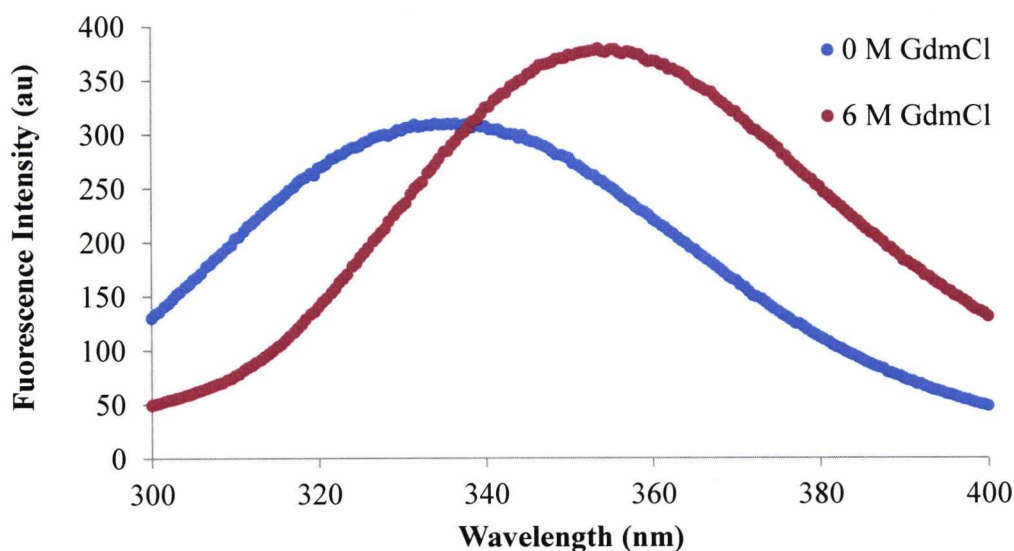


Figure 3.6 Reduced WT **b'xa'c** in the presence of 0 M (●) and 6 M (●) GdmCl.

The denaturation curves of WT and mutant **b'xa'c** proteins in the absence and presence of DTT are shown in figure 3.7.

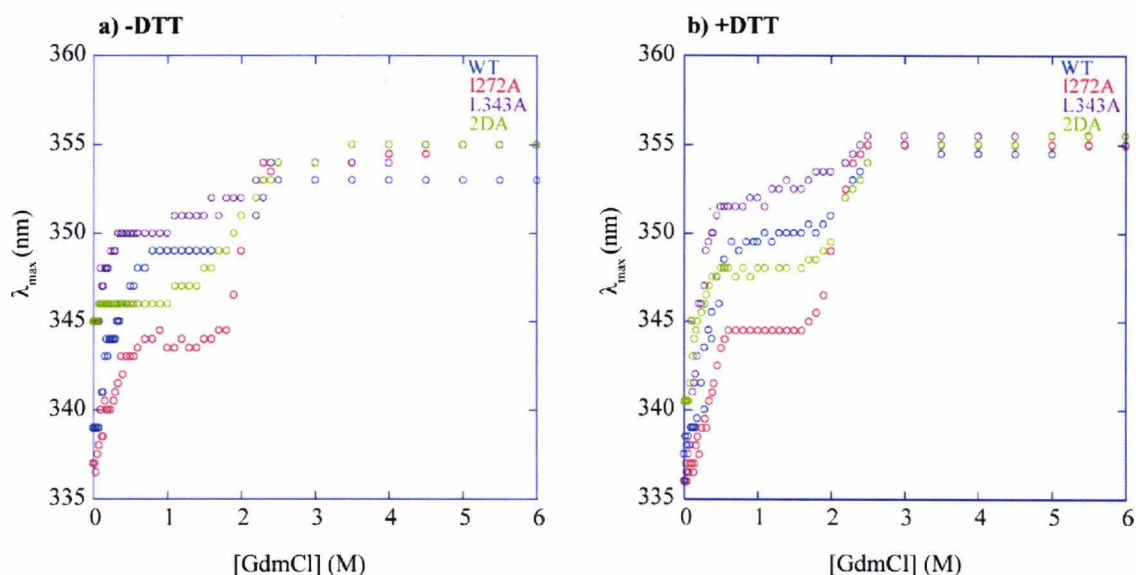


Figure 3.7 Denaturation of WT **b'xa'c (○) and mutants I272A (○), L343A (○) and 2DA (○).** The wavelength of maximal emission (λ_{\max}) plotted at each denaturant concentration shows a bi-phasing transition for all the **b'xa'c** proteins in **a)** the absence and **b)** presence of DTT.

Guanidine denaturation of WT and mutant **b'xa'c** showed two phases of transition, with the first phase at around 0 – 0.6 M and the second 1 – 2.5 M GdmCl concentration. All proteins were completely denatured from 2.5 – 3.5 M denaturant. As seen in figure 3.6, mutants of **b'xa'c** presented with different conformational stabilities not only from the WT but also from each other. The presence of DTT had a considerable effect on the first phase of denaturation, especially in the 2DA mutant where the first phase was essentially missing in the absence of DTT. Thermodynamic parameters were calculated separately for each transition phase assuming that they could be fitted to the two state denaturation model. The denatured state for the first phase was assumed to be the native state for the second denaturation phase when the fraction unfolded, f_D , was calculated using equation 3.3 (figure 3.8a and b). The free energy for each point in the transition was calculated using equation 3.6 and plotted against GdmCl concentration to determine the conformational stability in the absence of denaturant, $\Delta G^{(H_2O)}_U$, (figure 3.8d and e). The Hill equation was used to fit the transition region in phase 2 only but could not be fitted to phase 1.

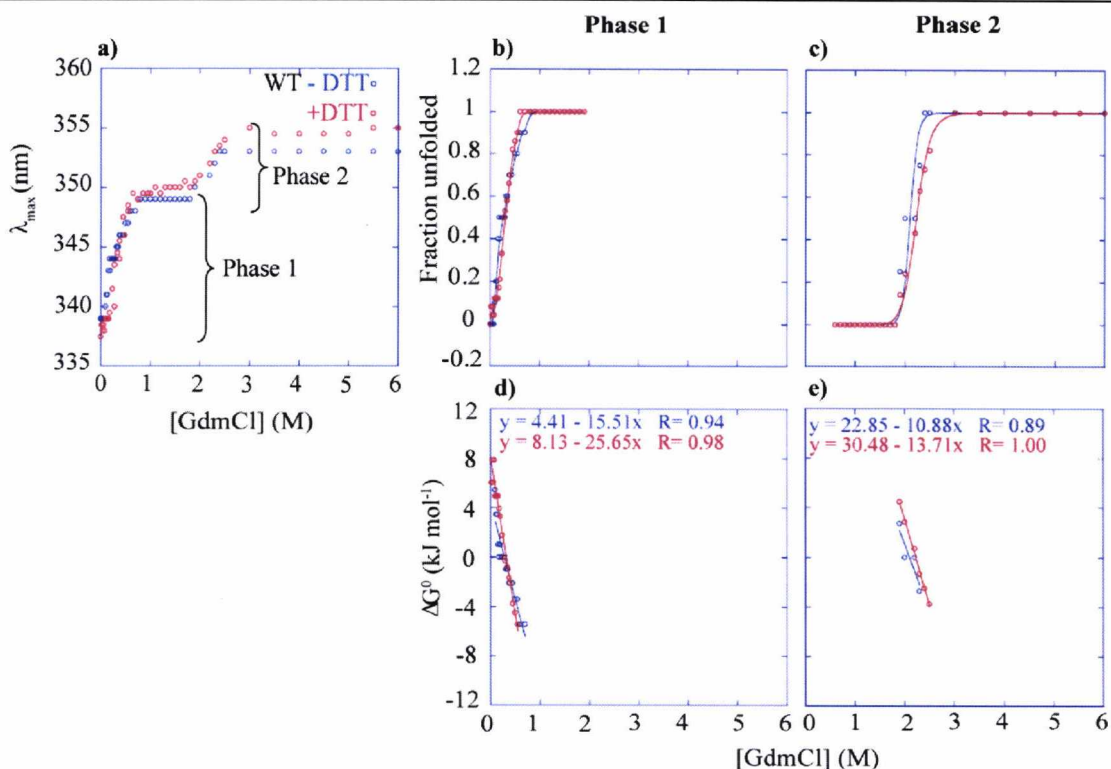


Figure 3.8 GdmCl denaturation of WT **b'xa'c** in the presence and absence of DTT. **a)** plot of λ_{\max} against denaturant concentration shows a bi-phasic transition. The fraction unfolded for phase 1 (**b**) and phase 2 (**c**) were used to calculate the equilibrium constant, K_D , which was applied to equation 3.6 to determine the energy required for the unfolding of each phase in the absence of denaturant $\Delta G^{(H_2O)}_U$, **d)** and **e)** respectively, in the presence (\circ) and absence (\circ) of DTT.

Mutants I272A, L343A and 2DA showed bi-phasic denaturation curves similar to WT **b'xa'c** (figure 3.9). The first denaturation step could not be fitted to the Hill equation by KaleidaGraph as it did not fit the 2-state denaturation system due to the lack of protein stability in the presence of low denaturant concentrations, but the K_D was calculated assuming that the native state of the protein was that at 0 M denaturant concentration. Just like in the WT protein, this first denaturation phase may represent a domain that is already partially unfolded, whose stability was significantly increased on addition of DTT suggesting that this phase could be due to the unfolding of the **a'** domain. This is also supported by **xa'c** and **b'x** data on page 20. This phase was considerably more stable in I272A and L343A than in 2DA.

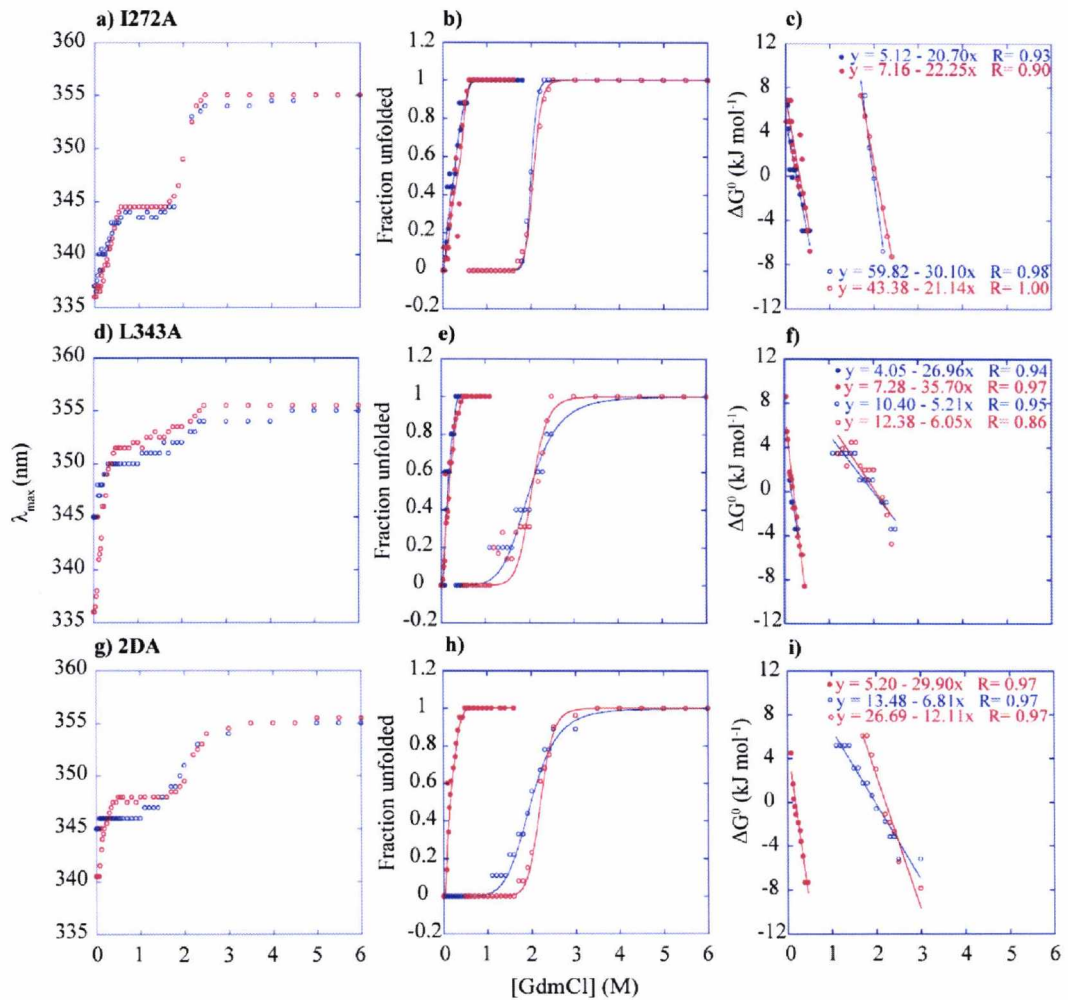


Figure 3.9 GdmCl denaturation of b'xa'c mutants I272A, L343A and 2DA. a) – c) I272A λ_{\max} , fraction unfolded and ΔG^0_{app} against GdmCl concentration, d) – f) L343A λ_{\max} , fraction unfolded and ΔG^0_{app} against GdmCl concentration, g) – i) 2DA λ_{\max} , fraction unfolded and ΔG^0_{app} against GdmCl concentration. (●) phase 1 –DTT, (●) phase 1 +DTT, (○) phase 2 –DTT, (○) phase 2 +DTT.

The second transition phase also showed a difference in stability between the mutants, where I272A was more cooperative than L343A and 2DA. Like the first transition phase, the second phase was also affected by the addition of DTT but to a smaller extent.

The energy required and the denaturant concentration for the unfolding of phase 1 was much lower than that of phase 2. The ΔG^0_{app} for phase 2 was very similar to the energy required for the unfolding of WT b'x with very similar midpoints as shown in published work as well as data collected from other members of the laboratory

(Pirneskoski et al., 2004). Therefore, assuming that phase 2 of denaturation was due to the unfolding of **b'x**, phase 1 must be attributable to the denaturation of the **a'** domain. To prove that the phases were indeed due to the unfolding of individual domains, denaturation data for **a'c**, **xa'c** and W390F **b'xa'c** were also collected.

Emission spectra at different GdmCl concentrations for **a'c** and **xa'c** were recorded with the **a'** active site in the reduced and non-reduced state to determine the effect of the oxidation state on the conformational stability of the **a'** domain (figure 3.10). Production and purification of these PDI constructs was carried out as described in chapter 2 for **b'xa'c**. W390F **b'xa'c** denaturation data were only collected in the absence of DTT as the sole fluorescence reporter in this mutant is W347 in **b'x** where there are no disulphide bond-forming cysteines. This was also carried out to confirm that the tryptophan to phenylalanine mutation did not interfere with the structural stability of **b'xa'c** and that the λ_{\max} shifts in fluorescence are indeed due to a change in the conformation of the **x** linker. As expected, the denaturation profile of this mutant was very similar to that of **b'x**. **a'c** appeared to be partially unfolded as it had very low stability even in very low GdmCl concentrations. Addition of the **x** linker region, in **xa'c**, improved the stability of **a'** somewhat but not enough to form a fully folded domain and produce a cooperative sigmoidal curve.

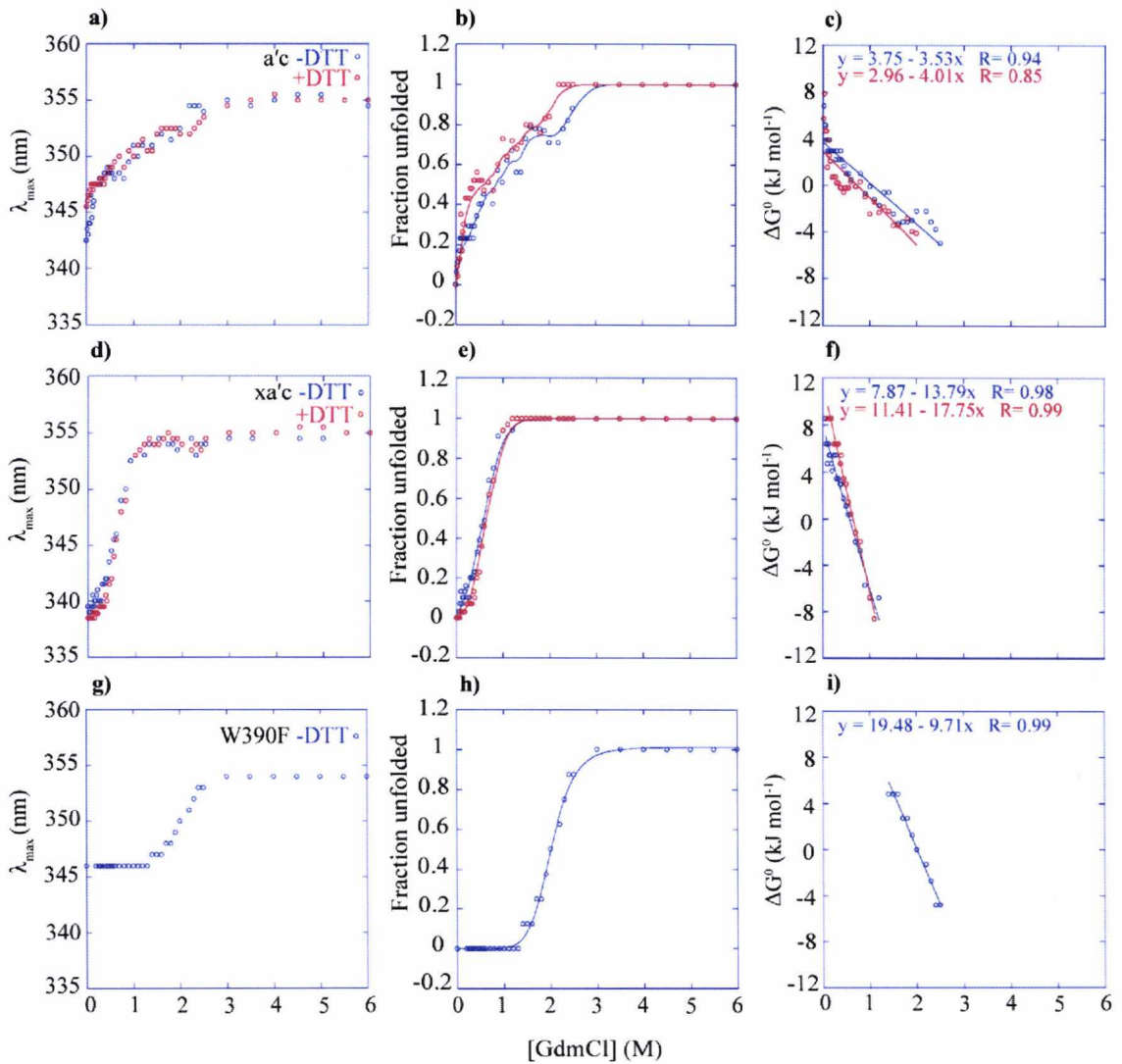


Figure 3.10 GdmCl denaturation of **a'c**, **xa'c** and **W390F b'xa'c**. **a) – c)** **a'c** λ_{\max} , fraction unfolded and ΔG^0 values against denaturant concentration, respectively. **d) – f)** **xa'c** λ_{\max} , fraction unfolded and ΔG^0_{app} dependence on GdmCl concentration. **g) – i)** **W390F b'xa'c** λ_{\max} , fraction unfolded and ΔG^0_{app} dependence on GdmCl concentration. **a'c** and **xa'c** were collected in the absence (\circ) and presence of DTT (\odot).

The conformational stability of the **a'** domain in **a'c** and **xa'c** was similar to that of phase 1 in WT **b'xa'c** and mutants, confirming that this phase was in fact due to the unfolding of **a'**. The denaturation curve for the W390F mutant was comparable to the second denaturation phase of **b'xa'c** proteins. As this mutant is mainly a **b'x** reporter, the second transition phase in **b'xa'c** must be caused by the unfolding of the ligand binding domain **b'**.

Table 3.2 summarises the thermodynamic properties of WT **b'xa'c** and its mutants as well as **xa'c** and **a'c** constructs without DTT for comparison with **b'x** proteins.

Protein -DTT	Phase 1 a'c			Phase 2 b'x		
	$\Delta G^{(H_2O)}_U$ (kJ mol ⁻¹)	[GdmCl] _{1/2} (M)	<i>m</i> value (kJ mol ⁻¹ M ⁻¹)	$\Delta G^{(H_2O)}_U$ (kJ mol ⁻¹)	[GdmCl] _{1/2} (M)	<i>m_U</i> value (kJ mol ⁻¹ M ⁻¹)
WT b'xa'c	4.41±0.3	0.28	-15.51	22.85±2.5	2.10	-10.88
I272A b'xa'c	5.12±0.4	0.50	-20.70	59.82±1.2	2.00	-30.10
L343A b'xa'c	4.05±0.2	0.15	-26.96	10.40±0.5	2.00	-5.21
2DA b'xa'c	-	-	-	13.48±0.4	1.98	-6.81
xa'c	7.87±0.2	0.57	-13.79	-	-	-
a'c	3.75±0.2	1.06	-3.53	-	-	-
W390F b'xa'c	-	-	-	19.48±0.2	2.00	-9.71
*WT b'x	-	-	-	29.11	2.04	-14.62
*I272A b'x	-	-	-	28.60	2.17	-13.18
*L343A b'x	-	-	-	19.09	1.67	-11.43
*2DA b'x	-	-	-	43.05	2.33	-18.48

Table 3.2 Thermodynamic properties of WT **b'xa'c and **b'x** and their respective mutants I272A, L343A and 2DA, **a'c** and **xa'c** in the non-reduced state.** ΔG^0_{app} was calculated as $\Delta G^0_{app} = -RT \ln K_D$. The midpoint of denaturation, [GdmCl]_{1/2}, was the denaturant concentration where $\Delta G^0_{app} = 0$ kJ mol⁻¹. The *m* value is the slope of the gradient of the dependence of ΔG^0 on denaturant concentration. ***b'x** data were collected by Holly Baldock and Rebecca Hayton.

The free energy of unfolding, $\Delta G^{(H_2O)}_U$, for the **a'** domain was much smaller than that for the **b'** domain. However, the ΔG^0_{app} for the first denaturation phase was calculated assuming that the **a'** domain was fully folded in the absence of denaturant when in fact it must have been partially unfolded due to the absence of a stable form at low denaturant concentrations. The extent of unfolding of **a'** must have been different for the WT protein and the mutants but it was not possible to determine the proportion of the unfolded population, therefore the denaturation midpoints for this domain are misleading. The midpoints of denaturation for the second transition phase of the **b'xa'c** proteins, the concentration of denaturant where $\Delta G^0_{app} = 0$, are similar, not only to each other, but also the **b'x** proteins. The differences between WT **b'xa'c** and its mutants were seen in the $\Delta G^{(H_2O)}_U$ values for the denaturation of **b'x**. The free energy of unfolding for the

denaturation of **b'x** in I272A **b'xa'c** was nearly three times larger than that of the WT protein. This is supported by previous data where I272A has been described as a **b'** stabilising mutant due to stabilisation of **x** capping the ligand binding site (Nguyen et al., 2008). Conversely, the uncapped L343A mutant presented with a $\Delta G^{(H_2O)}_U$ that was only half that of the WT, suggesting that this mutant has a destabilising effect on **b'xa'c**. Surprisingly, the free energy of unfolding of **b'x** in the 2DA mutant of **b'xa'c** was similar to that of L343A. 2DA in **b'x** promotes capping and has a greater $\Delta G^{(H_2O)}_U$ not only than WT **b'x** but also I272A **b'x**. The 2DA mutation in **b'xa'c** caused significant destabilisation of the **a'** domain, demonstrated by the absence of the first denaturation phase, which in turn must have affected the conformational stability of the **b'** domain. Denaturation of **a'c** showed that this domain is conformationally unstable and partly unfolded even in the absence of GdmCl. Its stability increased with the addition of the **x** linker region in **xa'c**, but it was still difficult to determine if the domain was fully folded in native conditions. This is supported by NMR data discussed in chapters 4 and 5. Overall, the **a'** domain is substantially less stable to chemical denaturation than the **b'** domain.

The *m* value is an experimental measure of the dependence of ΔG^0 on denaturant concentration, which is related to the chemical groups within a proteins that are exposed to the solvent in the denatured state (Myers et al., 1995). The *m* values for the denaturation of **b'x** in **b'xa'c** followed a similar trend to the ΔG^0 values. L343A and 2DA presented with higher *m* values than the WT and I272A, suggesting that they have more groups exposed to the solvent and as a result a different conformation. I272A was more conformationally stable than WT **b'xa'c** as it had a lower *m* value and as a consequence less exposed groups.

The W390F mutant had similar denaturation midpoint, ΔG^0_{app} and *m* value to WT **b'xa'c** confirming that this mutation did not affect **b'xa'c** structurally and that any shifts in fluorescence were due to the conformation of the **x** linker region.

Table 3.3 shows the effect of the addition of DTT on the conformational stability of **b'xa'c** proteins. Reduction of the **a'** active site caused an increase in stability of the proteins especially in the **a'** domain. This was more noticeable in the **a'** domain of the 2DA mutant, which in the non-reduced form appeared almost completely unfolded. Although addition of DTT increased the stability of **a'** and caused a blue-shift in fluorescence, this domain never appeared fully folded and therefore the thermodynamic

parameters shown in table 3.3 are still misleading. In summary, the **a'** domain denaturation data did not fit the 2-phase denaturation system and the cooperativity was low compared to **b'x**.

Protein + DTT	Phase 1 a'c			Phase 2 b'x		
	$\Delta G^{(H_2O)}_U$ (kJ mol ⁻¹)	[GdmCl] _{1/2} (M)	<i>m</i> value (kJ mol ⁻¹ M ⁻¹)	$\Delta G^{(H_2O)}_U$ (kJ mol ⁻¹)	[GdmCl] _{1/2} (M)	<i>m_U</i> value (kJ mol ⁻¹ M ⁻¹)
WT b'xa'c	8.13±0.2	0.32	-25.65	30.48±0	2.22	-13.71
I272A b'xa'c	7.16±0.7	0.32	-22.25	43.38±0	2.05	-21.14
L343A b'xa'c	7.28±0.2	0.21	-35.70	12.38±1.7	2.05	-6.05
2DA b'xa'c	5.20±0.2	0.18	-29.90	26.69±0.8	2.20	-12.11
xa'c	11.41±0.1	0.64	-17.75	-	-	-
a'c	2.96±0.4	0.74	-4.01	-	-	-

Table 3.3 Thermodynamic properties of WT **b'xa'c and its mutants I272A, L343A and 2DA, **a'c** and **xa'c** in the reduced state.**

Table 3.3 illustrates that reduced **b'xa'c** demonstrated an increase in the midpoint for the denaturation of the **a'** domain but had no significant effect on **b'**. However the ΔG^0 values for the unfolding of **b'x** were slightly lower compared to the non-reduced protein. Overall, the **a'** domain appeared to be less conformationally stable than **b'**. The above data shows that **x** has a stabilising effect not only on **b'**, as stated in the literature, but also on the **a'** domain and that mutations in **x**, like in L343A and 2DA, have a significant effect on the stability of **a'**. Reduction of the **a'** active site directly affects the stability of **b'** by lowering the energy required for its unfolding, whereas a more stable **b'** in I272A has a positive effect on the stability of the **a'** domain. Conformational stability data supports the shifts in λ_{max} of the WT and mutant **b'xa'c**, as red-shifted mutants have a lower conformational stability and the blue-shifted mutant is more stable than WT **b'xa'c**.

3.3.3 Determination of conformational flexibility by limited proteolysis

Limited proteolysis of WT **b'xa'c** and mutants I272A, L343A and 2DA, with proteinase K, was carried out in order to compare the effect of the mutations on the conformational flexibility of **b'xa'c**. Initial observations showed that WT and I272A

presented with similar proteolytic products on SDS-PAGE. L343A had a similar band pattern to WT and I272A, but proteolytic products were more intense on SDS-PAGE, suggesting increased flexibility of the L343A mutant compared to WT and I272A. 2DA was almost completely digested by proteinase K showing greater susceptibility to the protease than the WT and the other two mutants (figure 3.11). This must have been due to the fact that the **a'** domain was unfolded to a greater extent in this mutant and therefore was more prone to proteolytic cleavage. A thick band between 15 and 20 kDa in in lane 8 could be due to **b'x**.

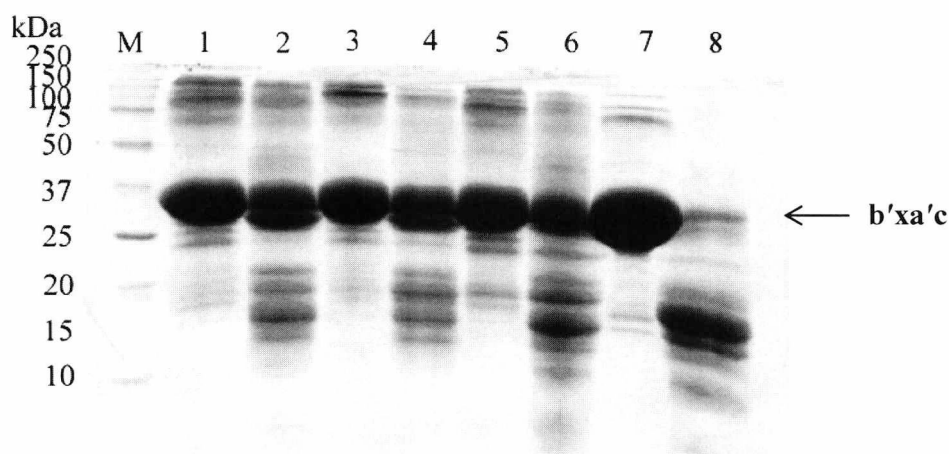


Figure 3.11 12.5% SDS-PAGE analysis of proteolysis products of WT **b'xa'c and mutants.** 2 mg/mL of each protein were digested with 2 μ g/mL of proteinase K for 30 minutes at 25°C. Controls consisted of undigested proteins. Lane 1: WT **b'xa'c** control; Lane 2: WT **b'xa'c** proteolysis products; Lane 3: I272A control; Lane 4: I272A proteolysis products; Lane 5: L343A control; Lane 6 L343A proteolysis products; Lane 7: 2DA control; Lane 8: 2DA proteolysis products.

The oxidation state of the **a'** domain has been shown to be important for the flexibility of PDI, therefore limited proteolysis was repeated with the proteins in the reduced and non-reduced states. Addition of DTT decreased the susceptibility of the proteins to proteinase K suggesting that the proteins are less conformationally flexible when the **a'** active site is in the reduced state. Figure 3.12 shows the susceptibility of **b'xa'c** to proteinase K in the presence of DTT.

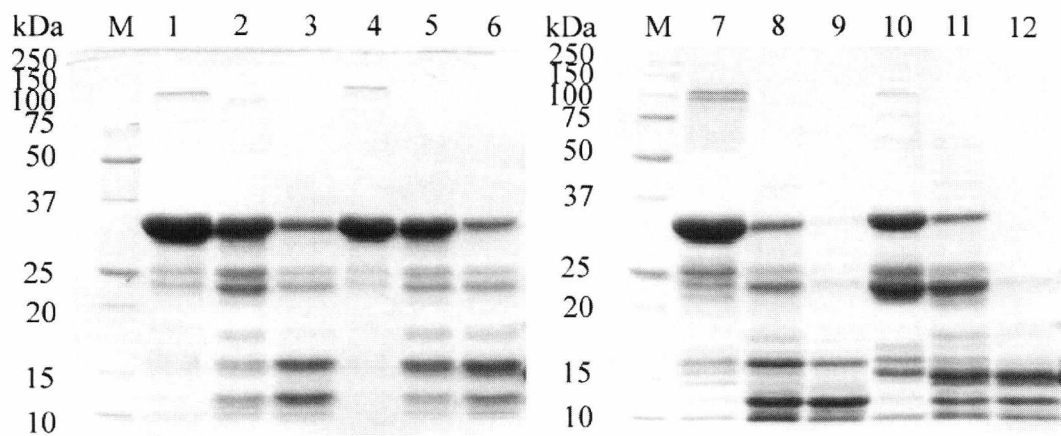


Figure 3.12 12.5% SDS-PAGE analysis of WT **b'xa'c** and mutants in the presence and absence of DTT. Digestion was allowed to occur for 1 hour at 25°C. Lanes 1 – 3 : WT **b'xa'c** control, proteolysis products + DTT and – DTT respectively; Lanes 4 – 6: I272A control, proteolysis products + DTT and – DTT respectively; Lanes 7 – 9 L343A control, proteolysis products + DTT and – DTT respectively; Lanes 10 – 12: 2DA control, proteolysis products + DTT and – DTT respectively.

Figure 3.11 shows that WT and I272A **b'xa'c** were less susceptible to proteolytic cleavage by proteinase K than L343A and 2DA, both in the presence and absence of DTT. L343A and 2DA were completely digested in the absence of DTT, whereas WT and I272A in lanes 3 and 6 respectively still contained some undigested protein. Even in the reduced state, L343A and 2DA were more susceptible to cleavage by the protease than WT and I272A **b'xa'c**.

3.4 Discussion

3.4.1 Determination of capping in **b'xa'c** by fluorescence

Capping of the ligand binding site on the **b'** domain by the **x** linker region has been shown to occur in a number of PDI constructs and there has even been evidence presented that full length PDI itself exhibits this conformational change (Wang et al., 2010, Nguyen et al., 2008). The presence of tryptophan residues in PDI has assisted in the determination of the alternative conformations of PDI. W347 in the **x** linker region has proved an excellent tool in tracking the capped and uncapped conformations by intrinsic fluorescence. I272A and 2DA mutants in **b'x** have been shown to favour the capped conformation, where the **x** linker region caps the ligand binding site on the surface of the **b'** domain. During this conformational change the side chain of W347 is buried in the hydrophobic environment of the binding site and as a consequence displays a blue-shift in fluorescence. Conversely, the L343A mutant of **b'x** promotes uncapping of the ligand binding site by **x**, therefore W347 is found exposed to the surrounding aqueous solution. A red-shift in fluorescence is observed as a consequence of the hydrophilic environment of the W347 side chain. The I272A and L343A mutants have been shown to have a similar effect on the fluorescence of properties of full length PDI as in **b'x**, but the shifts in the wavelength of maximal intensity between the mutants and the wild type have been at a smaller magnitude.

Fluorescence studies on **b'xa'c** of PDI revealed that the I272A and L343A mutants had a similar effect on the conformation of the **x** linker region as they do in **b'x**, with I272A being more blue-shifted and L343A more red-shifted than WT **b'xa'c** (figure 3.4). This suggests that these mutants also stabilise capping and uncapping, respectively, and as the λ_{\max} of the WT protein fell in between the values of these two mutants, WT **b'xa'c** must be present as a mixture of the two conformations. However, the differences in λ_{\max} shifts between the mutants and the WT are smaller than those seen in **b'x**. So, it is likely that the I272A mutation is promoting capping, but as **x** has another domain on its side, the **b'** binding site is not fully capped but W347 is close enough to report on its hydrophobic environment. Therefore capping in **b'xa'c** may not be as complete as in **b'x**, but **x** is still in close proximity with hydrophobic patch on **b'**. Also, the structural arrangement of capping is most likely different in **b'xa'c** compared to the crystal structure of I272A **b'x**. It is not surprising that I272A and L343A have these effects on **b'xa'c**. L343 is one of the residues in **x** that interacts with the ligand binding site and when mutated to an alanine, the

interaction is disrupted and consequently capping is inhibited (Nguyen et al., 2008). The side chain of I272 is found behind the hydrophobic pocket on **b'** and mutating this residue to an alanine has been shown to decrease peptide binding; most likely due to competition from the **x** linker region (Pirneskoski et al., 2004). 2DA was surprisingly red-shifted in **b'xa'c** when the same mutant promotes capping in **b'x** and displays a blue-shift in fluorescence. As the 2DA mutations are in the **x** linker region and this change could detrimentally inhibit **x – a'** interactions as well as affect **x – b'** interactions and the capping event.

Dimers of WT **b'xa'c** and mutants were more red-shifted than the monomers. This was not surprising as dimerisation is thought to occur through interaction of the hydrophobic ligand binding patch on the surface of each monomer therefore displacing **x** from the ligand binding site and setting it free in solution giving more red-shifted fluorescence than the monomers.

3.4.2 Determination of conformational stability of PDI constructs

Very little is known about the conformational stability of PDI, possibly due to its complicated fluorescence profile and conformational flexibility. Previous fluorescence-based denaturation studies using GdmCl have shown a single transition phase from native to the denatured form of **b'** and **b'x**, with midpoints for denaturation at 1.65 and 2.32 M respectively. The denaturation midpoint for full length PDI has been reported as 1.35 M (Pirneskoski et al., 2004).

GdmCl denaturation of WT **b'xa'c** and mutants I272A, L343A and 2DA presented as a bi-phasic denaturation curve, with the first phase owing to the unfolding of the **a'** domain and the second phase as a result of the unfolding of **b'x**. The free energy of unfolding of **a'** was much lower than that for **b'**, as the **a'** domain appeared to be partially unfolded even in the absence of denaturant. Therefore the denaturation midpoints for the **a'** domain were misleading as the data did not fit to the two-state denaturation system and it was difficult to measure the folded base-line. Denaturation of the **a'c** and **xa'c** constructs as well as the W390F mutant confirmed that the bi-phasic denaturation curve was indeed as a result of the unfolding of the individual domains suggesting that the domains in **b'xa'c** behave like two domains on a string rather than as a globular protein, as reported by the gel filtration data in chapter 2.

Mutants I272A, L343A and 2DA appeared to have different conformational stabilities compared to WT **b'xa'c**. The I272A was the most stable **b'xa'c** protein, confirming its role as a **b'** stabilising mutant. 2DA **b'xa'c** was the mutant with the lowest conformational stability as the **a'** domain appeared nearly completely unfolded and **b'** stability was also reduced compared to the WT protein. Therefore, the 2DA mutant does not have the same stabilising effect in **b'xa'c** as it does in **b'x**. The L343A mutation also decreased the conformational stability of **b'xa'c**, but this was to be expected as L343A in **b'x** is considered a **b'** destabilising mutant as it favours uncapping of the ligand binding site by the **x** linker region (Nguyen et al., 2008). The **a'c** construct in isolation showed very low conformational stability and did not show a simple two state reaction for a folded protein. Addition of the **x** linker region in the **xa'c** construct significantly increased the conformational stability of the **a'** domain suggesting that the **x** linker confers stability to the **a'** domain, possibly through interactions.

Addition of DTT significantly increased the conformational stability of the **a'** domain in all of the proteins, especially in the 2DA mutant where the first denaturation phase in the absence of DTT was nearly non-existent, but on addition of DTT the free energy of denaturation could be calculated. DTT increased the conformational stability of the **a'** domain in all of the PDI proteins. This is in agreement with recently published data which shows the **a'** domain form interactions with the **x** linker and the **b'** domain forming a compact structure upon reduction of the **a'** active site.

The m values are a measure of the change in accessible surface area on unfolding of the protein. Therefore, a high m value denotes a large number of exposed amino acid groups and a greater extent of unfolding than a lower m value. The m values for WT and mutant **b'x** proteins are similar to each other, suggesting that the folded structures of **b'x** proteins are similar. This is not the case for **b'xa'c** proteins, which show a variation in the change of accessible surface area, therefore suggesting structural differences in the folded form. This difference is seen in the presence and absence of DTT, giving confidence in the data. As the change in accessible surface area for the **b'x** domain in I272A **b'xa'c** is smaller than that of WT **b'xa'c**, this may suggest that I272A has a smaller accessible surface area than the WT protein in the native state. Whereas, the accessible surface area is larger for L343A and even larger for 2DA, suggesting an uncapped conformation. This is reasonable if the accessible surface area of the unfolded forms of the proteins is the same and a two state system is assumed.

Interestingly, the I272A and 2DA mutants appear to behave differently across the two data sets, i.e **b'x** and **b'xa'c**. In the case of I272A, the *m* value is much higher than the WT in the **b'xa'c** background compared to **b'x** only, whereas in the case of 2DA, *m* is higher than the WT **b'x** background but lower in **b'xa'c**. Presumably these differences are due to the behaviour of **x** when it is tethered by unfolded **a'c**. The 2DA mutation in **x** is able to stabilise **b'** when **x** is 19 residues, as in **b'x**, but not when it is lengthened by the unfolded **a'c** in **b'xa'c**.

3.4.3 Determination of conformational flexibility by limited proteolysis

The C-terminal half of PDI, **b'xa'c** has been shown to be more susceptible to proteolytic cleavage than the N-terminal **ab** fragment (Wang et al., 2010). This has been attributed to the more flexible conformation of **b'xa'c** compared to **ab**. Mutants of **b'xa'c** also showed different susceptibility to proteinase K, implying that they are in a different conformation to the WT protein. Limited proteolysis supported the denaturation data in highlighting the negative effect of the 2DA mutant on the conformational stability and flexibility of **b'xa'c**. WT and I272A **b'xa'c** showed similar susceptibility to proteinase K, whereas L343A was slightly more susceptible. 2DA was completely digested by proteinase K due to the low conformational stability of the **a'** domain. This data also shows that the 2DA mutant is definitely not capping the ligand binding site in **b'xa'c**.

Addition of DTT decreased the susceptibility of the proteins to proteinase K suggesting that the proteins were less conformationally flexible when the **a'** active site is in the reduced state. This was supported by the GdmCl data which showed that the proteins exhibited increased conformational stability upon reduction of the **a'** active site. Both reduced **bb'xa'** and **abb'xa'** have been shown to be less susceptible to proteolytic cleavage. This is due to a redox-dependent conformational change upon reduction of the **a'** active site, in which the **b'xa'** fragment assumes a more compact structural unit (Wang et al., 2012a, Wang et al., 2012b)

CHAPTER 4

Characterisation of **b'x** and **xa'c** by NMR Spectroscopy

4.1 Introduction

NMR assignments allow identification of individual resonances from observed NMR active nuclei within a protein. For many years, sequential assignment has been the primary method of amino acid backbone resonance assignment of proteins (Wütherich, 1986). This approach, used for the assignment of **xa'c**, exploits a combination of through-bond triple resonance experiments to match a distinctive spin system to the type of amino acid to which it belongs, followed by the identification of the adjacent spin systems and their position in the amino acid sequence of the protein. This method has enabled the assignment of large complex molecules in a short period of time (Kay, 2005). Once completed, NMR assignments could be used for chemical shift mapping in pH and temperature titrations, ligand binding and structural NMR relaxation dynamics studies.

Characterisation of the single domain constructs **b'x** and **xa'c** of PDI has been carried out as a stepping stone to assist in the investigations of the double domain construct **b'xa'c** in the next chapter. **b'x** has already been the subject of extensive studies as the X-ray crystal structure of the capping I272A mutant and NMR assignments of WT **b'x** have already been solved (Nguyen et al., 2008, Byrne et al., 2009). Conversely, very little

research has been carried out on the **a'** domain of human PDI, possibly due to the challenging nature of the NMR data from this domain. Initial studies showed that expression levels, in *E. coli* of **a'**, without the C-terminal tail and the **x** linker region were very low and insufficient for NMR studies. Protein levels increased when the **c** tail was part of the construct in **a'c**, but $^{15}\text{N}/^1\text{H}$ HSQC spectra appeared line broadened and of poor quality, suggestive of conformational exchange. However, the addition of the **x** linker region in **xa'c** significantly improved the quality of the NMR data as peaks appeared better resolved and less line broadened. For this reason, **xa'c** was used in triple resonance experiments for the assignment of the **a'** domain as well as relaxation and ligand binding studies.

4.1.1 $^{15}\text{N}/^1\text{H}$ HSQC experiments

Heteronuclear Single Quantum Correlation (HSQC) is the standard 2D NMR experiment that shows the correlations between ^1H and the directly bonded NMR active ^{15}N or ^{13}C heteroatoms. These are mainly the backbone amide groups but indole groups of Trp ($\text{N}\epsilon\text{-H}\epsilon$) and Asn/Gln side chain ($\text{N}\delta\text{-H}\delta 2/\text{N}\epsilon\text{-H}\epsilon 2$) groups are also visible. Therefore the HSQC spectrum contains a peak for each unique proton attached to the labelled heteroatom, except Pro residues which do not contain an amide proton in their structure so are not seen in $^{15}\text{N}/^1\text{H}$ HSQC spectra.

4.1.2 Sequential backbone assignment by Triple Resonance

Triple resonance assignment experiments for **xa'c** required protein that was $^{15}\text{N}/^{13}\text{C}$ isotopically enriched to enable the assignment of $^1\text{H}_\text{N}$, ^{15}N , $^{13}\text{C}_\alpha$ and $^{13}\text{C}_\beta$. Each resonance in the NMR spectrum is associated with a specific nucleus in the molecule being investigated and is assigned to a spin system in a particular amino acid residue in the protein sequence. So, each resonance must not only be precisely assigned to the amino acid type, but each assignment must be specific for individual nuclei in each amino acid in the protein sequence. These assignments are enabled by through-bond triple resonance experiments which are typically run in pairs. The first experiment of the pair correlates the amide ^1H and ^{15}N with the ^{13}C nuclei in a particular amino acid (designated “i”) and the ^{13}C nuclei of the preceding amino acid (designated “i-1”). The second experiment

correlates the same amide ^1H and ^{15}N with the ^{13}C nuclei of the preceding amino acid only. These experiments are named after the correlations that they detect. For example, in the CBCA(CO)NH and CBCANH pair of experiments that were used to assign **xa'c**, CBCANH correlates the ^1H and ^{15}N with the $^{13}\text{C}_\alpha$ and $^{13}\text{C}_\beta$ nuclei of the same and preceding amino acid, whereas the CBCA(CO)NH correlates the ^1H and ^{15}N with only the $^{13}\text{C}_\alpha$ and $^{13}\text{C}_\beta$ of the preceding amino acid. Figure 4.1 shows how these experiments are used to link one NH group to the next into a long chain.

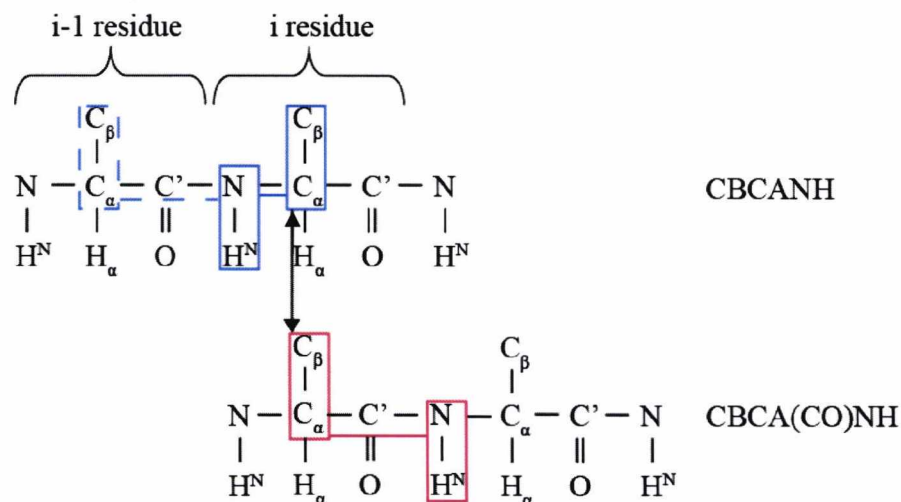


Figure 4.1 Dipeptide segment of a protein backbone showing correlations obtained in **CBCANH** and **CBCA(CO)NH** experiments. CBCANH correlates each NH group with the C_α and C_β chemical shifts of its own residue and of the residue preceding. CBCA(CO)NH only correlates the NH group of the preceding C_α and C_β chemical shifts.

This approach utilises the distinctive C_α and C_β chemical shifts which are characteristic of the amino acid type. The chemical shifts of certain amino acids, such as threonine and serine, are very different to those of the other amino acids as their C_β present at a higher field than the C_α . Conversely, the C_β of alanine lies at a lower field than the C_β of all the other amino acids making it easy to identify. Glycine is also easy to spot as it does not contain C_β , therefore it is only presents as one peak at around 45 ppm which represents its C_α . Valine, isoleucine and proline are also likely to stand out as they have higher than normal C_α chemical shifts. The random coil NMR chemical shift values for all of the amino acids obtained by Sykes and co-workers is shown in the figure 4.2 (Wishart et al., 1995).

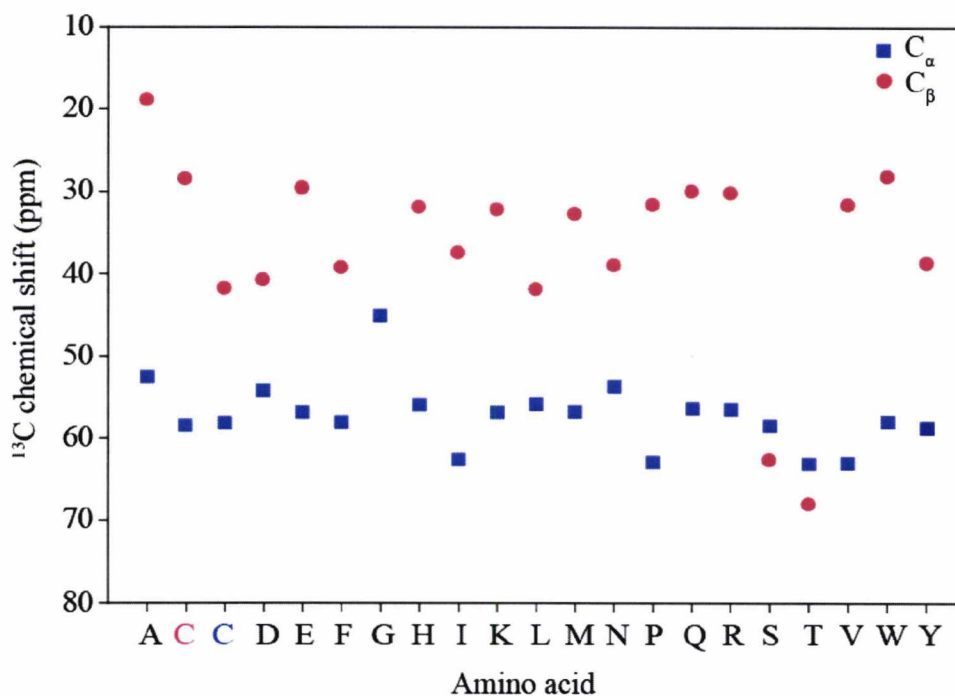


Figure 4.2 Chemical shift patterns of $^{13}\text{C}_\alpha$ (■) and $^{13}\text{C}_\beta$ (●) resonances for amino acids in random coil conformation (Wishart et al., 1995). The distinctive patterns of glycine, alanine, serine and threonine can be seen compared to the other amino acids. The chemical shifts for cysteine in the reduced (C) and oxidized (C) form are also shown.

The data from the triple resonance experiments was loaded into CcpNmr Analysis software which placed each data pair into clusters and interrogated the cluster lists to provide matches in rank order of best fit. Matching clusters were then linked together to form a sequential assignment. $^{13}\text{C}_\alpha$, $^{13}\text{C}_\beta$, ^{15}N and $^1\text{H}_\text{N}$ resonances could be unambiguously assigned using the unique chemical shift pattern for glycine, alanine, serine and threonine.

4.1.3 Chemical Shift Mapping

Frequency and wavelength are the common units used in spectroscopy. However, NMR spectroscopy uses chemical shifts instead, as frequencies of NMR lines are directly proportional to the magnetic field strength. This means that doubling the field strength doubles the frequency so it is not possible to compare absorption frequencies between spectrometers which operate at different field strengths. Whereas on the chemical shift scale, the positions of the peaks are independent of field strength (Keeler, 2005).

Chemical shift mapping is routinely used to monitor changes in chemical shifts when a protein is altered in solution. Mapping uses two-dimensional experiments like the $^{15}\text{N}/^1\text{H}$ HSQC to follow changes in the backbone amide chemical shift which can then be mapped onto the proteins structure or amino acid sequence. Chemical shifts are sensitive to variations in chemical environment which could be due to oxidation status, temperature, pH or addition of a ligand that interacts with the protein. All of these changes are reported to have an effect on the PDI constructs discussed in this thesis.

4.1.4 Secondary Structure Prediction from chemical shifts

DANGLE (Dihedral Angles from Global Likelihood Estimates) is a program built into CcpNmr Analysis Version 2 used for secondary structure prediction (Cheung et al., 2010, Vranken et al., 2005). DANGLE is an algorithm which uses a Bayesian inference statistical approach to predict the secondary structure of a protein based on the amino acid type and its chemical shift, as well as taking into account the adjacent amino acids and interrogating a chemical shift/structure database for related sequences. This information is then used to predict the most likely secondary structure. The chemical shifts obtained from the triple resonance assignment for $^1\text{H}_\text{N}$, ^{15}N , $^{13}\text{C}_\alpha$ and $^{13}\text{C}_\beta$ were used for the secondary structure prediction in DANGLE.

Secondary structure defines specific dihedral angles across the protein backbone as they rotate about the C-C $_\alpha$ and C $_\alpha$ -N bonds, Ψ (Psi) and Φ (Phi) respectively. These allow the backbone to adopt regular structure arrangements which are also affected by the type of amino acids and their side chains. Typical backbone secondary structures and their angles are represented by the Ramachandran plot in figure 4.3 (Ramachandran et al., 1963). DANGLE generates Ramachandran plots for all of the amino acids in the protein sequence for which the chemical shifts have been previously assigned.

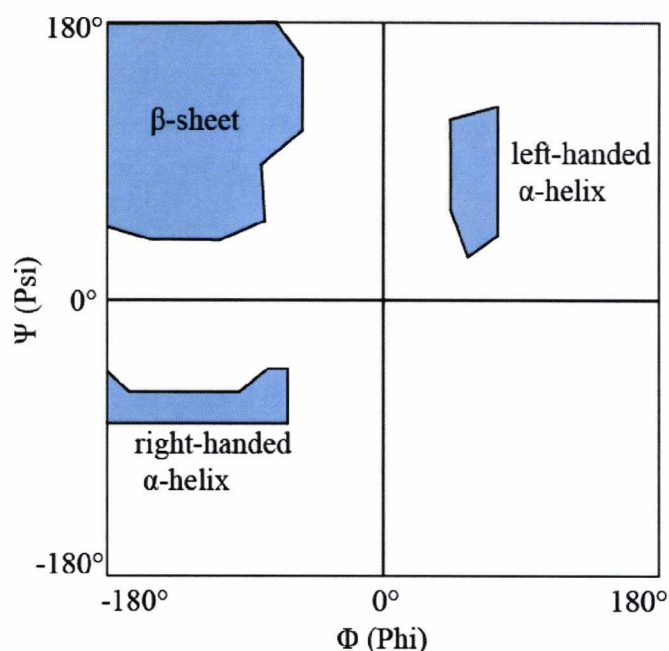


Figure 4.3 Simple representation of a Ramachandran plot. Angles Φ (Phi) between C_{α} and N bond and Ψ (Psi) between the C and C_{α} bond characterise the different types of secondary structure. Typical angles and secondary structure types of the protein backbone are shown by the blue shaded regions.

4.1.5 ^{15}N Backbone NMR Dynamics

NMR relaxation dynamics consist of a suite of experiments that provide detailed information related to the global and internal structural motions of the protein under investigation, as well as conformational exchange between two or more states. Molecular dynamics can be calculated using in depth quantum mechanics and mathematical modelling in order to quantify the degree of motion that exists within a protein molecule as proteins are flexible dynamic molecules that can assume a number of different conformations in solution. The dynamic motion of a protein in solution is defined by the global correlation time τ_m (sometimes referred to as τ_c). The global correlation time is defined as the time taken for a molecule to rotate through one radian or 57.3° (Abragam, 1961). The correlation time is related to the diffusion coefficient for the protein in solution and so it is dependent not only on the size of the molecule but also the temperature and viscosity of the solution. Equation 4.1 shows a generalized relationship between the number of residues in a protein and the global correlation time, τ_m (Daragan and Mayo, 1997).

$$4.1 \quad \tau_m = \frac{9.18 \times 10^{-3}}{T} \exp\left(\frac{2416}{T}\right) N^{0.93}$$

Where T is the temperature in Kelvin and N is the number of residues in the protein under investigation, which is assumed to be isotropic in shape. The global correlation time, τ_m , reports on the overall shape of the protein and increases with size (therefore molecular weight). As few proteins are truly spherical, molecular tumbling anisotropy provides different effective correlation times for each nucleus in the protein (Palmer, 2001).

Molecular motions occur in different timescales and this allows for the analysis of individual sites and sequence-specific flexibility in a protein. Proteins primarily tumble in solution in the nanosecond timescale, whereas parts of the structures like loops and termini of the protein exhibit much faster internal motion, usually with motions in the picosecond timescale. Side chains and regions of secondary structure can display additional slower motions in the milli to nanosecond timescale. Therefore over the years, several NMR experiments have been designed to interrogate protein motions occurring in a broad range of timescales as shown in figure 4.4.

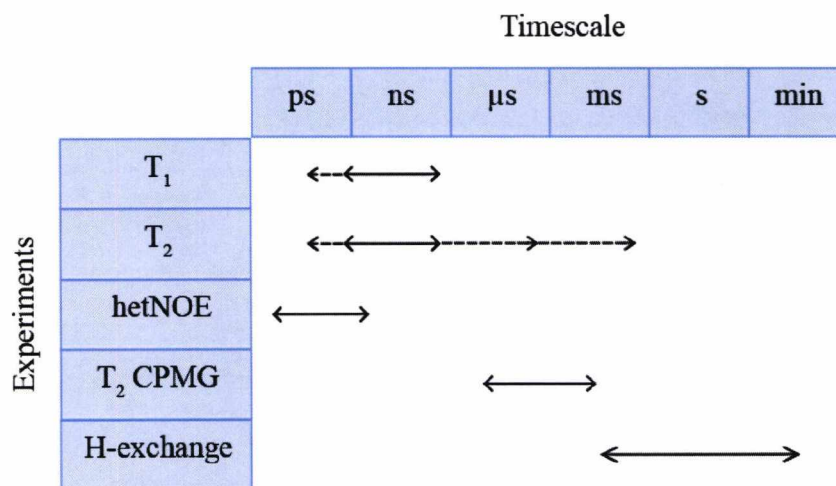


Figure 4.4 Timescales available for interrogation by different NMR experiments. Measurements of T_1 and T_2 show motions mainly in the nanosecond timescale, but can also display movement in the picosecond timescale. T_2 experiments can furthermore detect micro and millisecond motions. HetNOE experiments detect motions in the sub-nanosecond timescale. Experiments such as T_2 CPMG (Carr-Purcell-Meiboom-Gill) and Hydrogen exchange monitor longer timescale motions.

Protein backbone dynamics have been investigated using three types of through-bond experiments that measure NMR relaxation rates for each assigned amino acid residue in the protein: ^{15}N T_1 , T_2 and heteronuclear Nuclear Overhauser Effect (hetNOE). These experiments reveal different aspects of molecular dynamics and provide insight into the motion of individual nuclei at both the residue and bond level.

Longitudinal relaxation time (T_1), also known as spin-lattice relaxation, is influenced by molecular motions within the nano and sub-nanosecond timescale. Its classical interpretation is that nuclear spins relax to equilibrium with energy lost to the surroundings. It enables the recovery of bulk magnetization required for the detection of NMR signals, revealing information about molecular anisotropy and overall tumbling motion representative of molecular size. The inverse of T_1 gives the relaxation rate R_1 (s^{-1}) (Barbato et al., 1992, Rule and Kevin Hitchens, 2006).

Transverse relaxation time (T_2), also known as spin-spin relaxation, is not only influenced by the molecular motion in the nano to sub-nanosecond timescale, but is also sensitive to events in the micro and millisecond time scale as well as the overall size of the molecule and any chemical exchange. Its classical interpretation is that nuclear spins relax with a loss of phase coherence to adjacent spins. During T_2 relaxation, the detected signal is reduced when the bulk magnetization is perturbed by the phase coherence. The inverse of T_2 gives the relaxation rate R_2 (s^{-1}) (Palmer, 1997).

Heteronuclear Nuclear Overhauser Effect (hetNOE) measures the cross-relaxation that occurs between the ^1H and ^{15}N nuclei of the amide bond and therefore reports on the internal motion of proteins in the sub-nanosecond timescale. The hetNOE is observed as a difference in cross peak intensity when the saturation of the amide protons is selectively switched on and off. This is calculated using equation 4.2 where I is the saturated and I_0 is the reference experiment:

$$4.2 \quad \text{hetNOE} = \frac{I - I_0}{I_0}$$

The I and I_0 hetNOE experiments give rise to negative values for each cross peak where more negative values suggest faster internal motions in the sub-nanosecond timescale.

4.1.6 Model-free Formalism

Although plots of T_1 , T_2 and hetNOE against residue number provide vital information about the molecular motions of a protein, NMR relaxation data can be further analysed using basic models, such as the Model-free formalism (Lipari and Szabo, 1982a, Lipari and Szabo, 1982b). Longitudinal and transverse relaxation can be theoretically defined from the global correlation time, τ_m , and are described by a series of equations containing the spectral density function.

NMR relaxation is associated with the frequencies exhibited during the global or local motions of a protein, and the spectral density function, $J(\omega)$, describes the distribution of these frequencies displayed by the protein under investigation. The spectral density function is shown in equation 4.3:

$$4.3 \quad J(\omega) = \frac{2}{5} \frac{\tau_m}{1+(\omega\tau_m)^2}$$

Where $J(\omega)$ represents spectral density, τ_m global molecular correlation time and ω is the frequency of the NMR relaxation. Therefore, the spectral density is a function of motion from the correlation time and the frequency of the resonance. NMR relaxation is perturbed by anisotropic tumbling and exchange contributions. More specifically, T_1 is modified by the anisotropic shape of the protein described as the overall rotational diffusion term, D . Whereas T_2 is modified using terms associated with the rate of exchange, R_{ex} , and correlation time of exchange, τ_m . This results in T_1 containing information about the global τ_m based motion, whereas T_2 containing information about the global nanosecond τ_m and slower micro to nanosecond τ_e based motions. Consequently this will also perturb the spectral density function, which lacks the necessary terms to describe individual nuclei in a protein. Therefore, Lipari and Szabo proposed the model-free formalism which uses the spectral density function, including additional terms, to

describe NMR relaxation data of macromolecules as shown in equation 4.9 (Lipari and Szabo, 1982a, Lipari and Szabo, 1982b):

$$4.4 \quad J(\omega) = \frac{2}{5} \left[\frac{S^2 \tau_m}{1 + (\tau_m \omega)^2} + \frac{(1 - S^2) \tau}{1 + (\tau \omega)^2} \right]$$

Global Internal
motion motion

Where definitions are as stated in equation 4.3 with the addition of S^2 and τ' . S^2 is known as the order parameter and can take a value between 0 and 1; τ is a combination of the global correlation time τ_m and the internal correlation time, τ_e , so that $\tau' = \tau_m^{-1} + \tau_e^{-1}$. As well as defined as the internal correlation time, τ' is also used to describe fast internal motions and is known as the effective correlation time. The order parameter S^2 is a crucial component of the model-free formalism that describes the rigidity of the protein backbone. A rigid structural environment is represented by $S^2 = 1$ whereas a completely random flexible backbone is represented by $S^2 = 0$. Therefore the dynamic nature of a protein and that of each backbone amide can be quantified through the derivation of S^2 and τ' where global motion is represented by τ_m and internal motion by τ_e . Additional parameters used to fit NMR relaxation data include a general exchange rate, R_{ex} , and the expansion of the spectral density function into S^2 , S_f^2 and S_s^2 with the τ_e modified into τ_f and τ_s where “f” and “s” represent fast and slow internal motions respectively. Therefore, it is through the derivation of S^2 and τ_m for each backbone amide that the dynamic nature of a protein can be quantified in a simple way.

^{15}N T_1 and T_2 plots are a simplified version of the model-free formalism as they include the order parameter, S^2 , and global correlation time, τ_m , as contour lines to allow a visual comparison and analysis of experimental data. The plots also provide immediate comparison between different proteins as well as estimation of the S^2 and τ_m for each residue. However these plots do not include experimental hetNOE values whereas the model-free formalism incorporates all three relaxation parameters in order to define the backbone flexibility of the protein.

The Modelfree4 program developed by Palmer and colleagues (free to download at <http://www.palmer.hs.columbia.edu/software/modelfree.html>) can provide molecular

dynamics optimization from T_1 , T_2 and hetNOE data (Palmer et al., 1991, Mandel et al., 1995). If the three relaxation parameters have been measured at a single magnetic field, then no more than three model-free parameters can be fit to this data, in addition to the rotational correlation time or the diffusion tensor. There are five possible sets of model-free parameters that can be fit to three experimental data points:

Model 1: S_s^2

Model 2: S_s^2 and τ_e

Model 3: S_s^2 and R_{ex}

Model 4: S_s^2 , τ_e and R_{ex}

Model 5: S_s^2 , S_f^2 and τ_e

Here the term S_s^2 is the same as S^2 and $S_f^2 = 1$ for models 1-4. Whereas in model 5, S_s^2 and S_f^2 refer to internal motions on slow and fast timescales respectively. The order parameter S^2 can also be defined as $S^2 = S_f^2 \times S_s^2$ when fast/slow motions are active. The term R_{ex} is a numerical definition used to describe the chemical exchange contribution to the relaxation values.

The ModelFree4 program is widely used in the analysis of NMR relaxation data, for both peptides and proteins. One of hundreds of publications using ModelFree4 is the study of the dynamics of ribosomal protein S6 and its mutant P⁵⁴⁻⁵⁵ (Ohman et al., 2010). Model-free analysis of the relaxation data for S6 revealed a correlation time of 7.3 ns with residues in secondary structure elements displaying high ^1H - ^{15}N order parameter S^2 , with an average of 0.91. Whereas model-free analysis of the relaxation data for the P⁵⁴⁻⁵⁵ permutant yielded a correlation time of 6.9 ns and an average S^2 of 0.92. The relaxation studies combined with model-free analysis revealed flexible regions of the protein that were not apparent from the crystal structure of S6. Overall, the structure and dynamic properties of S6 in solution revealed that the overall stability of the protein was not sensitive to mutations in the flexible region.

4.1.7 Spectral Density and Spectral Density Mapping

Spectral density is a description of motion that occurs at specific frequencies and is related to the field strength of the NMR spectrometer. All relaxation experiments shown in this thesis were collected at 600 MHz, which means that for ^{15}N relaxation, 5 frequencies can be examined: $J(0)$ with a frequency of 0 Hz, $J(\omega_{\text{N}})$ at 60 MHz, $J(\omega_{\text{H}})$ at 600 MHz, $J(\omega_{\text{H}}+\omega_{\text{N}})$ at 540 MHz and $J(\omega_{\text{H}}-\omega_{\text{N}})$ at 660 MHz. Although five different frequencies are observed, the spectral density values are derived from the same three sets of relaxation data, ^{15}N T_1 , T_2 and hetNOE. True spectral density requires relaxation data collected at more than one magnetic field strength, so when only one field is used the resulting motion data is referred to as reduced spectral density and provides a reduced set of parameters $J(0_{\text{eff}})$, $J(\omega_{\text{N}})$ and $J(\omega_{\text{H}})$.

4.2 Materials and Methods

4.2.1 NMR sample preparation

Protein samples for NMR experiments were expressed in *E. coli* BL21 (DE3) pLysS grown in Minimal Media enriched with ^{15}N or $^{15}\text{N}/^{13}\text{C}$ as required. The proteins were purified as described in chapter 2 and concentrated using Vivaspin 4 concentrators (MWCO 10 kDa) to a final concentration between 0.4 and 2.0 mM as determined by absorbance at A_{280} nm. At the same time, proteins were buffer exchanged into the required buffers containing 10% (v/v) D_2O (Goss Scientific Ltd.).

4.2.2 NMR Data acquisition and processing

NMR data was acquired with the assistance of Dr. Mark Howard and Dr. Michelle Rowe. Most of the experiments were carried out on a Varian UnityINOVA spectrometer operating at 14.1 Tesla (^1H resonance frequency of 600 MHz) equipped with a 5 mm HCN z-pulse field gradient probe (except where stated). Experiments were collected at various temperatures (15, 25, 35 and 40°C). ^1H chemical shift referencing was based on the position of the water resonance with the exact value being related to the known $^1\text{H}_2\text{O}$ resonance with temperature, whereas ^{15}N and ^{13}C dimensions were referenced using a spectrometer based macro that utilises the gyromagnetic ratios relationship (Wishart and Sykes, 1994). All NMR experiments were solvent suppressed using WATERGATE (Piotto et al., 1992) in order to reduce the intensity of the signal from water.

NMR relaxation experiments for **b'x**, both WT and I272A, were collected on a Bruker spectrometer operating at the same magnetic field, whereas all **xa'c** data were collected on the Varian UnityINOVA spectrometer.

All NMR data were processed on a Linux workstation running SuSE 10.2, using NMRpipe software (Cornilescu et al., 1999) with Gaussian processing function.

4.2.2.1 $^{15}\text{N}/^1\text{H}$ HSQC NMR experiments of PDI constructs

$^{15}\text{N}/^1\text{H}$ heteronuclear single quantum correlation (HSQC) experiments (Cavanagh et al., 2007) were typically acquired with 2048 points (9000 Hz) in the direct F2 dimension (^1H) and 256 points (2100 Hz) in F1 (^{15}N). For ^{13}C enriched samples, a modified $^{15}\text{N}/^1\text{H}$ HSQC was used to incorporate ^{13}C decoupling during both T_1 and T_2 acquisition periods.

4.2.2.2 Amino acid sequential backbone resonance assignment

Backbone resonance assignments were obtained from a 2.0 mM reduced $^{13}\text{C}/^{15}\text{N}$ enriched **xa'c** sample in a pH 7.0 buffer containing 20 mM MOPS, 50 mM NaCl and 20 mM DTT. CBCANH and CBCA(CO)NH triple resonance data sets were acquired, at 25°C, with 1024 points (9000 Hz) in the direct F3 dimension (^1H), 60 points (10000 Hz) in F2 (^{13}C) and 20 points (2100 Hz) in F1 (^{15}N) indirect dimensions. Carrier frequencies for the triple resonance experiments were set to 4.766 ppm, 45.919 ppm and 119.454 ppm for ^1H , ^{13}C and ^{15}N respectively.

4.2.2.3 ^{15}N NMR Relaxation

^{15}N Relaxation experiments were carried out for **xa'c** and **b'x** (WT and I272A). Each T_1 and T_2 experiment was run with 2048 points (9000 Hz) in the direct F2 dimension and 256 points (2100 Hz) in the F1 dimension. All **xa'c** data were run on the Varian UnityINOVA spectrometer at 25°C with the carrier frequencies set to 4.766 ppm, 59.143 ppm and 118.477 ppm for ^1H , ^{13}C and ^{15}N respectively. T_1 relaxation delays for **xa'c** were set to 128, 256 (x2), 385, 513, 641 (x2), 769 and 894 ms. T_2 relaxation delays were set to 20, 40, 60 (x2), 80, 100, 120 (x2), 140 and 160 ms.

WT and I272A **b'x** relaxation data were run on a 600 MHz Bruker AV3 spectrometer with QCI-F cryoprobe at 25°C and the carrier frequencies were adjusted accordingly at 4.766 ppm, 59.143 ppm and 119.08 ppm for ^1H , ^{13}C and ^{15}N respectively. T_1 relaxation delays were set to 256, 384, 512, 640 (x2), 768, 894 and 1024 ms. T_2 relaxation delays were set to 33.92, 50.88, 67.84, 84.80 (x2), 101.76, 118.72 and 135.68 ms.

^{15}N HetNOE data were obtained by observing the intensity of the NH peaks with and without saturation of amide protons. HetNOE experiments were run with the same spectral widths as the T_1 and T_2 experiments.

4.2.3 NMR Data analysis

4.2.3.1 Triple Resonance assignment

Sequential assignment of the NMR backbone resonances for **xa'c** was carried out in software package CcpNmr Analysis Version 2 (Vranken et al., 2005). This was achieved

using triple resonance data from experiments CBCANH and CBCA(CO)NH and the distinctive patterns of the C_α and C_β of amino acids such as alanine, glycine, threonine and serine as well as the random coil shifts obtained by Sykes and co-workers shown in figure 4.2 (Wishart et al., 1995). The principle of triple resonance assignment using these experiments is described in the introduction section of this chapter.

4.2.3.2 Chemical Shift Mapping

Chemical shift mapping was used to show changes in chemical shifts ascertained from $^{15}\text{N}/^1\text{H}$ HSQC data obtained for proteins under identical conditions as an indication of change in the chemical environment for each residue. This type of analysis was used to generate a graph showing the minimal chemical shift difference with residue number. This type of analysis is carried out when comparing spectra when only one of the spectra is fully assigned and the distances from signals in the assigned spectrum to the nearest peak of the unassigned spectrum is measured. The minimal chemical shift difference was calculated using equation 4.5:

$$4.5 \quad \text{shift difference} = \sqrt{(\Delta^1H_N)^2 + \left(\frac{1}{6}\Delta^{15}N\right)^2}$$

Where Δ^1H_N is the chemical shift change in the 1H_N dimension and $\Delta^{15}N$ is the chemical shift change in the ^{15}N dimension, both in ppm. As each value is squared, the shift difference is always positive regardless of the Δ^1H_N or $\Delta^{15}N$ value. The ^{15}N difference is factored down to 1/6 to compensate for the chemical shift range of ^{15}N being ~30 ppm whereas that of 1H_N is ~5 ppm.

4.2.3.3 Secondary Structure Prediction using DANGLE

Secondary structure prediction for **xa'c** was completed using the DANGLE (Dihedral Angles from Global Likelihood Estimates) program built into the CcpNmr Analysis software. ^{15}N , 1H_N , $^{13}C_\alpha$ and $^{13}C_\beta$ chemical shifts obtained from the triple resonance assignment of **xa'c** were used for DANGLE to generate a Ramachandran plot for each residue in the sequence. During the analysis, prediction limits were set on the number of acceptable islands (regions within the Ramachandran plot) of possible dihedral angles that are implied by the chemical shifts: all predictions used in this study set limits to two islands. Any residue with more than two predicted dihedral angle islands was

excluded from the secondary structure prediction. In addition to the Ramachandran plots, DANGLE also provided predictions in a form on an individual residue basis where H represented a helix, C = coil and E = strand.

4.2.3.4 NMR Relaxation Analysis

All processed NMR spectra were assigned using the Analysis software package by copying peaks from spectra assigned using triple resonance experiments. ^{15}N T_1 and T_2 experiments measure the intensity of each ^{15}N - ^1H correlation over time, t . Peak heights were fitted to a decaying exponential in Analysis using equation 4.6.

$$4.6 \quad A \cdot \exp(-Bx)$$

Where A is the peak height for each relaxation delay, x is the relaxation delay and B is the calculated ^{15}N T_1 or T_2 time.

HetNOE values were calculated using $(I-I_0)/I_0$ where I is the peak height with saturation of amide protons and I_0 is the peak height of the reference spectrum.

^{15}N T_1 , T_2 and hetNOE data was entered into rplot, an internally developed program which accurately measures the τ_m of the protein under investigation and from which data can be exported directly into model-free. Model-free analysis of the relaxation data was carried out using ModelFree version 4.0 program (Lipari and Szabo, 1982a, Lipari and Szabo, 1982b). The model type used for each residue was based on the recommendation from the program developers starting with the simplest model (model 1: S^2 only), then working through to the more complex models whilst observing the improvement in the error of the calculated results. For T_1 versus T_2 plots, model-free order parameter contours, S^2 and τ_m , were generated in KaleidaGraph4.1 from a spreadsheet created by Dr. Mark Howard using the equations 4.7 – 4.10 with τ_e set at 50 ps (Farrow et al., 1995).

$$4.7 \quad \frac{1}{T_1} = \frac{d^2}{4} [J(\omega_H - \omega_N) + 3J(\omega_N) + 6J(\omega_H + \omega_N)] + \frac{c^2 J(\omega_N)}{3}$$

$$4.8 \quad \frac{1}{T_2} = \frac{d^2}{8} [4J(0) + J(\omega_H - \omega_N) + 3J(\omega_N) + 6J(\omega_H) + 6J(\omega_H + \omega_N)] + \left(\frac{c^2}{6}\right) [3J(\omega_N) + 4J(0)]$$

$$4.9 \quad d^2 = \left(\frac{\gamma_H \gamma_N \hbar \frac{\mu_0}{4\pi}}{r_{NH}^3} \right)$$

$$4.10 \quad c^2 = (\omega_N \Delta_N)^2$$

Where γ_H and γ_N are the gyromagnetic ratios of 1H and ^{15}N nuclei, μ_0 is the permeability of free space, r_{NH} is the length of the amide bond, Δ_N is the difference between the parallel and perpendicular components of the chemical shift tensor taken from Broadhurst and co-workers (1995), \hbar is reduced Planck's constant and ω_N and ω_H are the Larmor frequencies of 1H and ^{15}N at 14.1 Tesla (Broadhurst et al., 1995).

4.2.4 NMR temperature and pH titrations

$^{15}N/^1H$ HSQC NMR experiments were used to observe the effect of temperature and pH on **b'x** and **xa'c** as these conditions played a significant role in stabilising conformational exchange in **b'xa'c**, the data for which will be shown in the next chapter. Temperature experiments, ranging from 15°C to 40°C, were carried out on 1 mM protein samples containing 10 mM DTT, 20 mM NaH_2PO_4 buffer and 50 mM NaCl at pH 6.5. pH experiments were carried out on the same protein samples that were used for the temperature studies (1 mM protein containing 10 mM DTT). The proteins were buffer exchanged into the appropriate buffers for the pH required and $^{15}N/^1H$ HSQC spectra were recorded at 40°C. Table 4.1 shows the buffers used for each pH.

pH	Buffer composition
6.0	20 mM 2-(<i>N</i> -morpholino)ethanesulfonic acid (MES), 50 mM NaCl
6.5	20 mM NaH_2PO_4 , 50 mM NaCl
7.0	20 mM 3-(<i>N</i> -morpholino)propanesulfonic acid (MOPS), 50 mM NaCl
8.0	20 mM tris(hydroxymethyl)aminomethane (Tris), 50 mM NaCl

Table 4.1 Composition of buffers used in the NMR pH experiments.

NMR experiments for temperature and pH titrations were run in the following order:

Temperature experiments at pH 6.5 25°C → 15°C → 35°C → 40°C
 pH experiments at 40°C pH 6.5 → pH 6.0 → pH 7.0 → pH 8.0

4.2.5 Ligand binding using Δ -somatostatin

Ligand binding data of **b'x** to Δ -somatostatin have been published in the literature therefore these experiments were not repeated in this study (Byrne et al., 2009). ^{15}N -labelled **xa'c** was prepared to give a final concentration of 0.4 mM in a pH 7.0 buffer containing 20 mM MOPS and 50 mM NaCl. The protein was reduced with 4.0 mM DTT before the addition of the peptide ligand. Unlabelled Δ -somatostatin, present as a lyophilised powder, was prepared by making a stock solution of 1.2 mM stock in the MOPS buffer. The ligand was added to the protein in equimolar concentration. The sample was incubated for 30 min at room temperature before the addition of D_2O for NMR $^{15}\text{N}/^1\text{H}$ HSQC experiments. NMR spectra were run at 40°C on the Varian UnityINOVA 600 MHz spectrometer as described in section 4.2.2.1. A reference **xa'c** sample was also run without ligand as a control.

Δ -somatostatin was synthetically prepared by Dr. Kevin Howland and consisted of 14 amino acids (AGSKNFFWKTFTSS) where S3 and S14 have replaced C3 and C14 in the original somatostatin amino acid sequence. The ligand concentration was determined by measuring the absorbance at A_{280} nm due to the presence of the Trp residue in the peptide. The extinction coefficient for Δ -somatostatin is $5500 \text{ M}^{-1} \text{ cm}^{-1}$.

4.3 Results

4.3.1 Expression and purification of **b'x**, WT and I272A, and **xa'c**

Expression and purification of **b'x**, WT and I272A, and **xa'c** were carried out as described in chapter 2. All protein samples were run on gel filtration to separate the monomer from the dimer species as shown in figure 4.5. This was not only carried out as part of the purification protocol, but it was also used as part of the characterisation of PDI constructs.

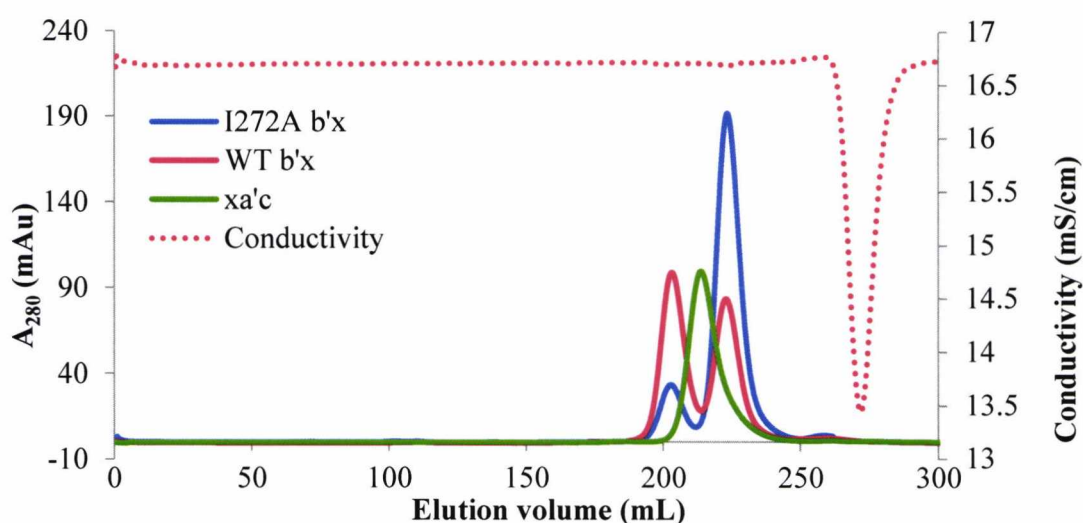


Figure 4.5 Gel filtration of **b'x** and **xa'c** using Superdex 200 media. The absorbance at A_{280} nm of WT **b'x** (red), I272A **b'x** (blue) and **xa'c** (green) is shown by the solid lines and the conductivity is displayed as the red dotted line.

Table 4.2 shows the expected molecular sizes for WT and I272A **b'x** and **xa'c** compared to those calculated from gel filtration. WT and I272A **b'x** appeared ~ 1.1 times larger in hydrodynamic volume than expected, whereas **xa'c** was ~ 1.4 times bigger.

Protein	Expected MW (kDa)	Gel filtration MW (kDa)	Oligomeric state
WT b'x	16.8	19	Monomer
		38	Dimer
I272A b'x	16.8	19	Monomer
		38	Dimer
xa'c	19.3	27	Monomer

Table 4.2 Expected and calculated molecular weights of WT and I272A **b'x** and **xa'c**. Molecular weight calculations were carried out as described for table 2.7 in chapter 2.

As seen in previous publications, WT **b'x** was a mixture of monomer and dimer (Wallis et al., 2009). I272A presented primarily as a monomer with a very small proportion of dimeric protein, whereas **xa'c** was exclusively in the monomeric form. Both monomer **b'x** proteins, WT and I272A, were eluted with the same volume of elution buffer, suggesting a similar hydrodynamic volume for both proteins. Conversely, **xa'c**, was eluted in fractions prior to the monomer of **b'x**, 16.8 kDa, and after the dimer, 33.6 kDa, in agreement with its molecular weight of 19.3 kDa. The monomer species of the proteins were carried forward for further analysis by NMR.

4.3.2 The effect of temperature and pH on PDI constructs

NMR experiments have shown the **b'xa'c** fragment of PDI to be in conformational exchange, demonstrated by the line broadening and poor quality of the spectra. This set of data will be discussed in more detail in the next chapter, but it is important to mention here that studies showed that reduction of the **a'** active site, temperature and pH contributed to lowering the degree of conformational exchange and hence improved peak dispersion and resolution of NMR $^{15}\text{N}/^1\text{H}$ HSQC spectra. For this reason, the effect of temperature and pH on **b'x** and **xa'c** were also investigated. As **b'x** has previously been shown to be a good target for NMR studies (Nguyen et al., 2008, Byrne et al., 2009), it was decided that a change in temperature and pH was not necessary to improve the already good quality of NMR spectra. However, the isolated **a'** domain expressed very low levels of protein, less than 1 mg/L, therefore the sample prepared for NMR analysis was only 60 μM in concentration, compared to the standard 1.0 mM samples prepared for the other constructs. The reasons for the poor expression levels of **a'** were not thoroughly investigated, however the results imply that the acidic C-terminal extension, **c**, plays an important part in increasing the stability and hence the expression levels of the **a'** domain.

Initial $^{15}\text{N}/^1\text{H}$ HSQC spectra showed the **a'** domain in isolation was not a suitable target for NMR studies due to the poor peak dispersion and resolution. Low concentration of the **a'** protein contributed to the reduced quality of the $^{15}\text{N}/^1\text{H}$ HSQC spectra due to the long acquisition time and water suppression issues. Figure 4.6 shows that **a'** in the absence of the reducing agent DTT appeared to be line broadened beyond detection, the peak intensities were very weak and the spectrum was plotted closer to the noise in **a' –DTT** than in the other spectra in order to distinguish some of the peaks that were from the protein.

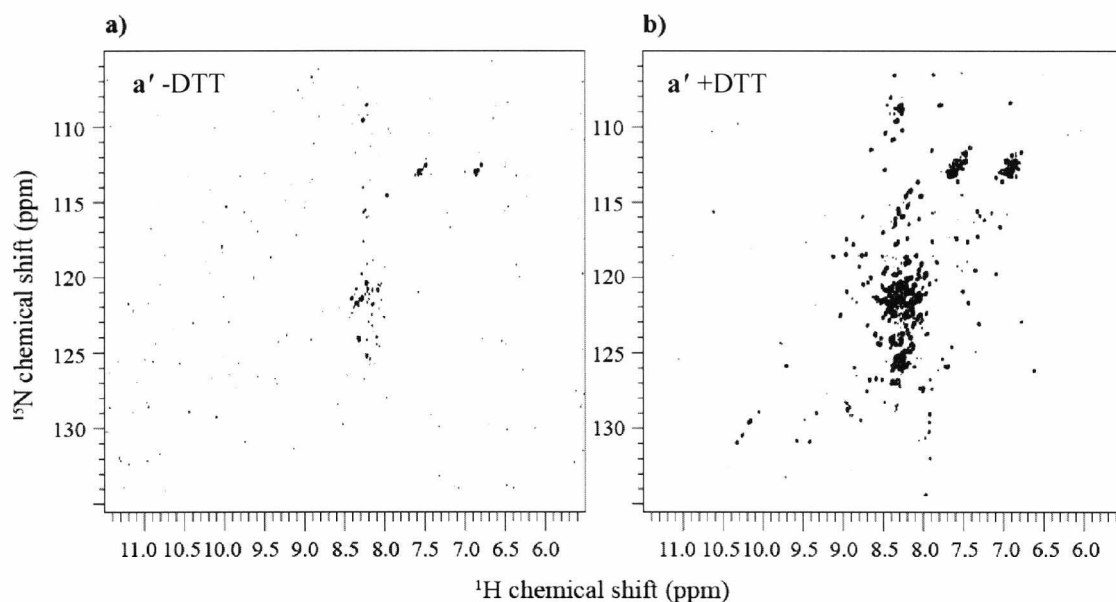


Figure 4.6 $^{15}\text{N}/^1\text{H}$ HSQC spectra of the **a'** domain of PDI. **a'** in the absence of DTT (a) shows an absence of NMR peaks which improve on addition of DTT (b).

Addition of DTT improved peak dispersion but due to the poor expression levels of the isolated **a'** domain, **a'c** was consequently investigated as a suitable target for NMR studies of the **a'** domain of PDI.

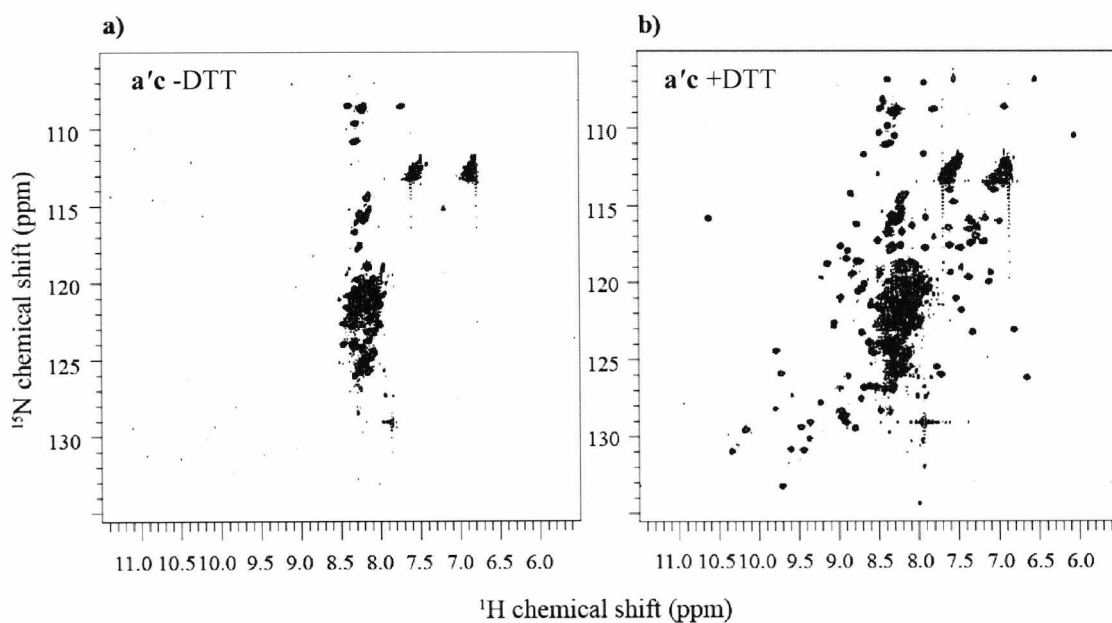


Figure 4.7 $^{15}\text{N}/^1\text{H}$ HSQC spectra of the **a'** domain of PDI with the C-terminal extension, **c**. **a'c** in the absence of DTT (a) shows the absence of NMR peaks, except for the **c** extension, improve on addition of DTT (b).

Although **a'c** expressed high levels of protein and upon addition of DTT the quality of the NMR spectra was significantly improved, there were sharp resonances localised between 8.0 and 8.5 ppm in the random coil region of the spectrum due to the unstructured C-terminal extension, **c**. As the fluorescence data in chapter 3 showed that **a'c** was in fact folded, NMR data indicates that this protein must be experiencing a high degree of conformational exchange leading to extensive line broadening. Consequently, the **xa'c** construct was considered as a better candidate for the characterisation of the **a'** domain of PDI. An added advantage to studying this construct was that the behaviour of the **x** linker in **xa'c** could be compared to that of **x** in the **b'x** construct. $^1\text{H}/^{15}\text{N}$ HSQC spectra of **xa'c** showed a major improvement in peak resolution and dispersion compared to **a'c**, as can be seen in figure 4.8.

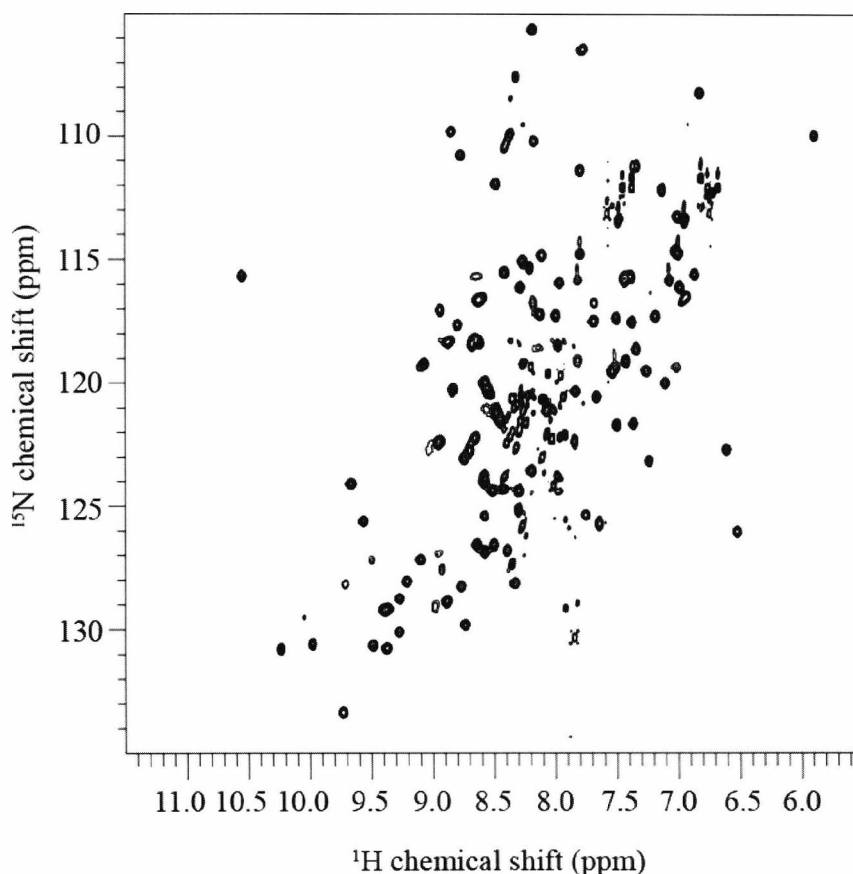


Figure 4.8 $^{15}\text{N}/^1\text{H}$ HSQC spectra of reduced **xa'c** at pH 6.5 run at 25°C.

The minimal chemical shift perturbation between **xa'c** and **a'c** was calculated using peaks from an **xa'c** assigned spectrum, peaks from an **a'c** unassigned spectrum and

equation 4.5. The minimal shift map is shown in figure 4.9. The data is shown in tabulated form in appendix 4.1.

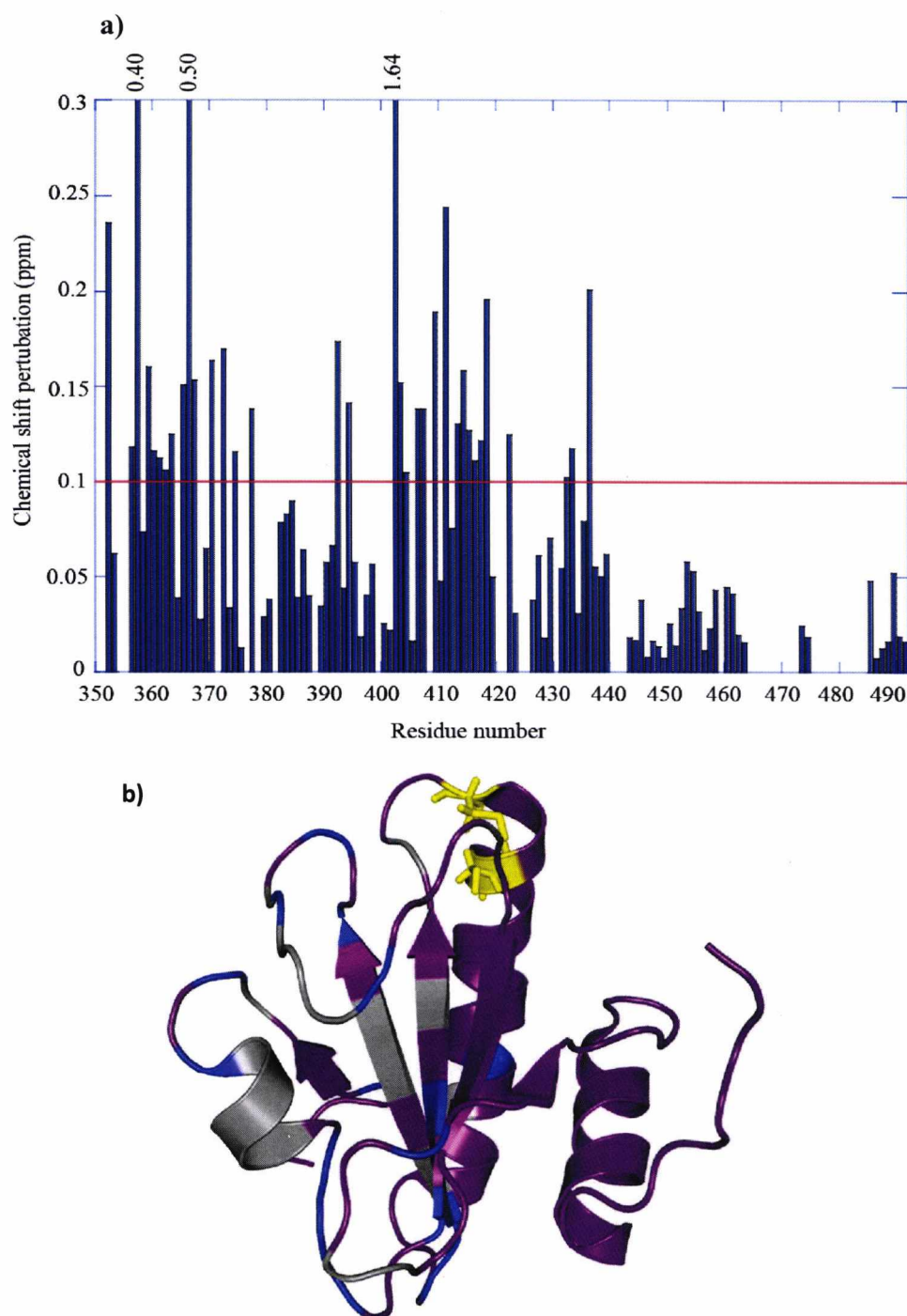


Figure 4.9 Minimal chemical shift map to the nearest peak for backbone amide ^1H and ^{15}N resonances of **xa'c** with **a'c** and mapped on the NMR structure of the **a'** domain (**1X5C.pdb**). Chemical shift differences greater than 0.1 ppm are mapped on the **a'** structure in grey and shifts greater than 0.15 ppm are coloured in blue. The cysteines of the active site are shown as yellow sticks.

Chemical shift perturbations were seen throughout the chemical shift plot with major differences (shifts above 0.15 ppm) seen for residues V352, G357, N359, F365, D366, E367, N370, F372, K392, N402, I403, D409, T411, E414, V418 and T436.

Once **xa'c** was established as a suitable target for NMR studies, experiments were run at different temperatures with the protein in buffers with varying pH in order to find the optimal conditions for this protein. Change in temperature and pH showed no major improvement in the quality of the $^{15}\text{N}/^1\text{H}$ HSQC spectra, but peak height was increased with the rise in temperature and pH as shown in figure 4.10. The data used for the graphs is shown in appendices 4.2 and 4.3 for the temperature and pH experiments respectively.

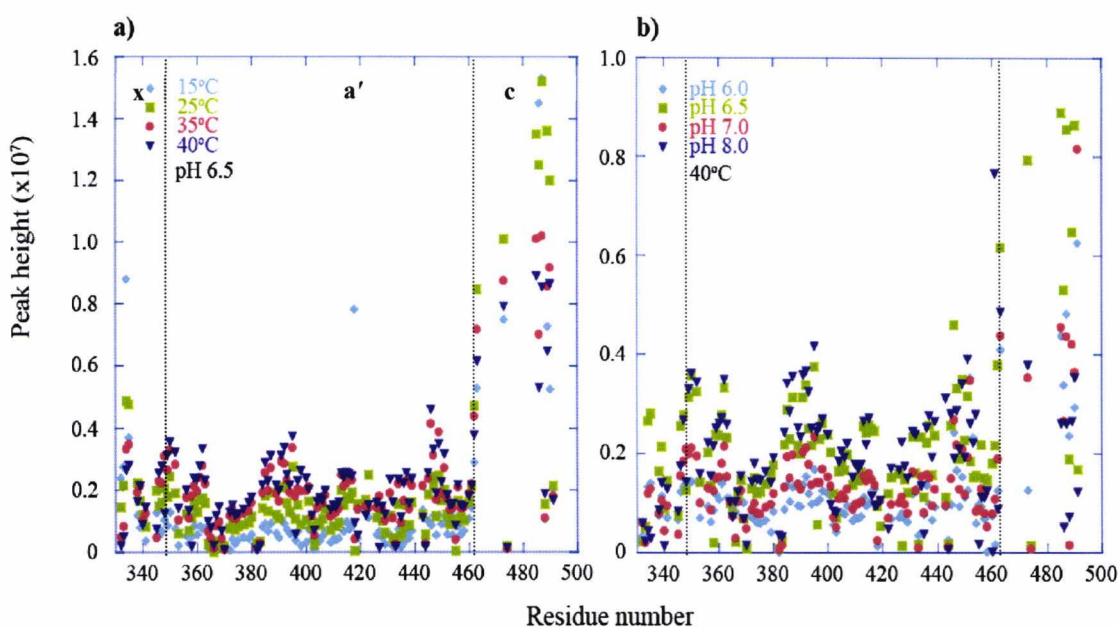


Figure 4.10 The effect of temperature and pH on the peak height of reduced **xa'c** $^{15}\text{N}/^1\text{H}$ HSQC spectra. a) peak height increases with a rise in temperature as well as b) an increase in pH. The dotted lines represent the domain boundaries.

The highest peak heights are observed at 40°C and pH 8.0. However as there was no significant difference in spectral quality, it was decided for future NMR experiments to be conducted at 25°C and pH 7.0. In addition to this, it was noted that **b'xa'c** was unstable at pH 8.0 and as comparable data was required, pH 7.0 was used in further experiments. This set of data also showed that the sharp peaks between 8.0 and 8.5 ppm were due to residues in the C-terminal extension, **c**, as they presented with the highest peaks in the spectra, suggesting that this region of the protein was unstructured.

The same experiments were also conducted for WT **b'x**, but as this construct already gave good quality NMR spectra at 25°C and pH 6.5, and as previously published data for this protein was collected under the same conditions, it was decided that a change in temperature and pH was unnecessary (Nguyen et al., 2008, Byrne et al., 2009).

In order to further characterise **xa'c** and **b'x** by NMR spectroscopy, assignment of the proteins was essential. WT **b'x** was previously assigned by other members of the group and those assignments could be mapped onto I272A **b'x** (Byrne et al., 2009). Therefore, assignments for **xa'c** were completed as shown in section 4.3.3.

4.3.3 Sequential backbone assignment of reduced **xa'c**

Sequential backbone assignments for $^1\text{H}_\text{N}$, $^{15}\text{N}_\text{H}$, $^{13}\text{C}_\alpha$ and $^{13}\text{C}_\beta$ nuclei for reduced **xa'c** were made from the pair of triple resonance experiments CBCA(CO)NH and CBCANH. These are shown in Appendix 4.4. Ignoring the His tag, backbone assignments for **xa'c** were obtained for 77% of amide ^1H and ^{15}N , 83% of $^{13}\text{C}_\alpha$ and 78% of $^{13}\text{C}_\beta$. Specific residues such as Ser/Thr and Gly were identified from the triple resonance experiments and linked sequentially as the example shown in figure 4.11 for the amino acid stretch from Glu454 to Gln458.

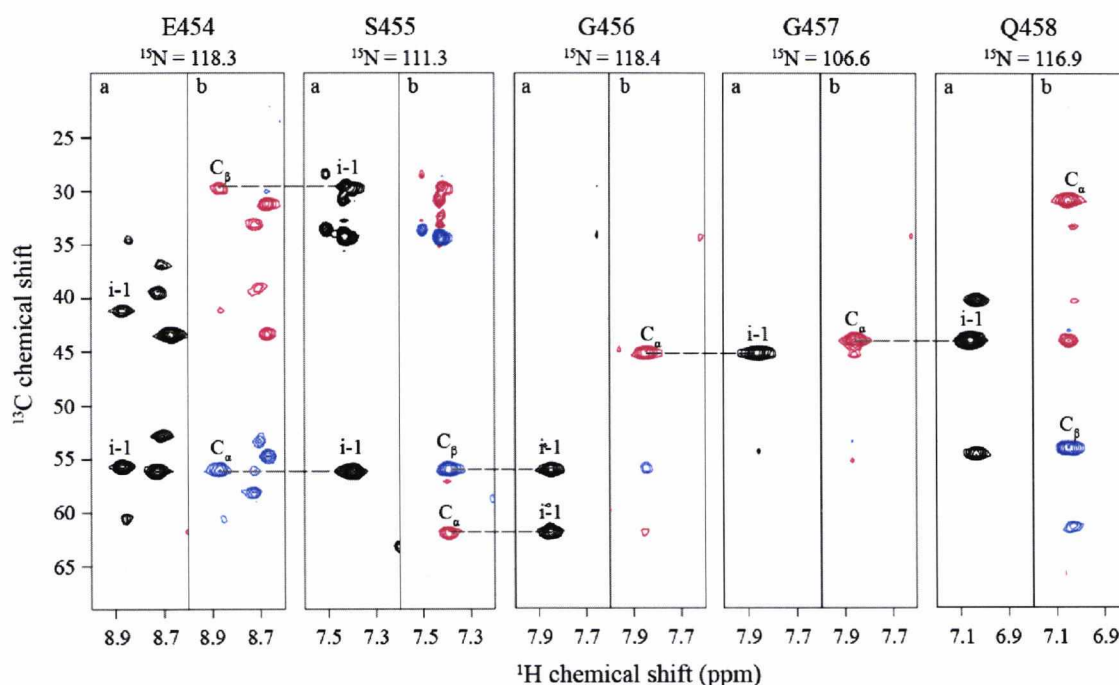


Figure 4.11 Backbone triple resonance sequential assignment example from **xa'c experiments.** CBCA(CO)NH and CBCANH strips, **a** and **b** respectively, show the assignment stretch for residues E454 to Q458. Sequential matches are highlighted by the dotted line. The CBCA(CO)NH experiment shows the C_α and C_β (black) atoms of the preceding amino acid. The CBCANH strips show the C_α as positive peaks in blue and the C_β as negative peaks in red. The relevant ^{15}N chemical shift position of each plane is shown above the strip.

Residue-labelled $^{15}\text{N}/^1\text{H}$ HSQC spectrum of reduced **xa'c** is shown in figure 4.12. The assignments were obtained as part of the triple resonance sequential assignments shown in figure 4.11 and listed in Appendix 4.4.

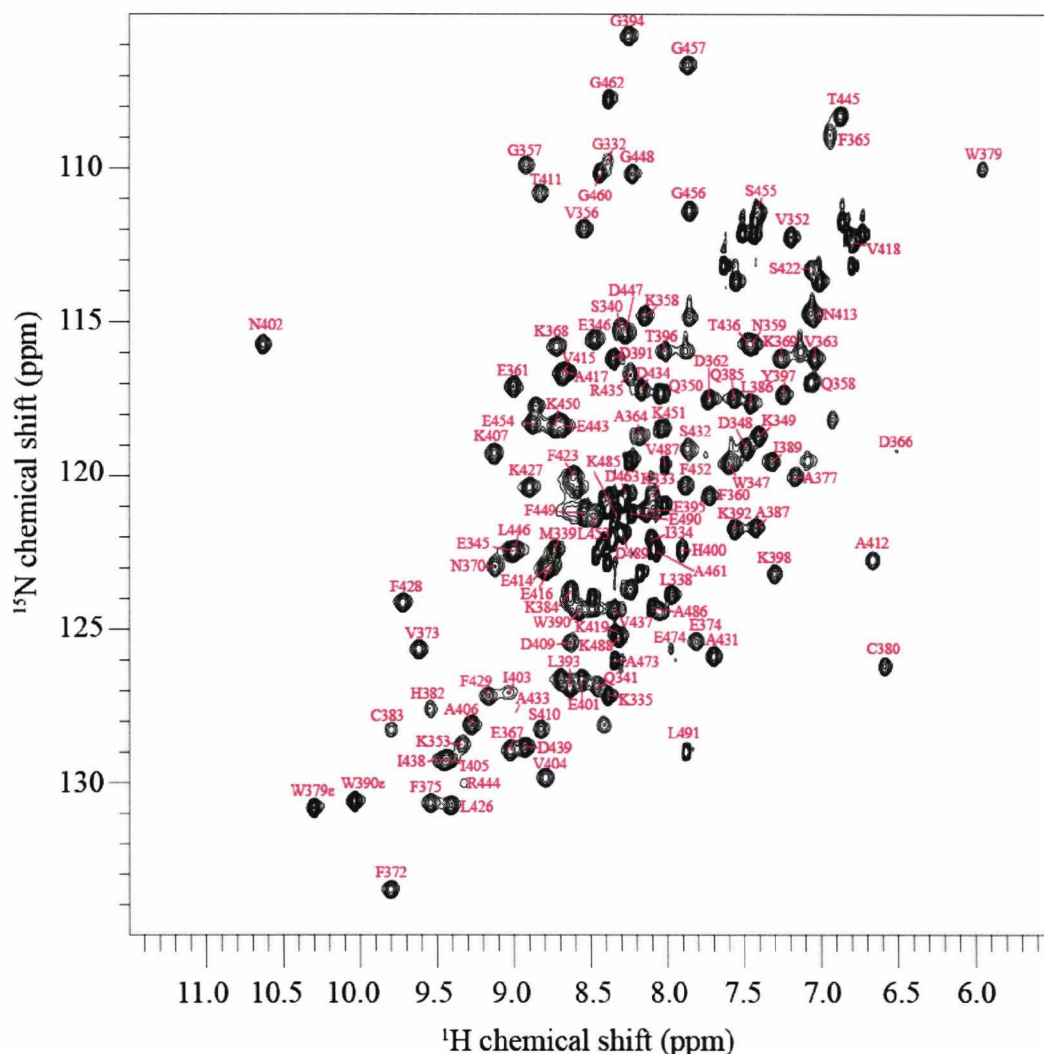


Figure 4.12 $^{15}\text{N}/^1\text{H}$ HSQC spectrum for reduced **xa'c**. All known amide resonance assignments are labelled in red on the spectrum. Experiments were run at 25°C with the protein sample at pH 7.0.

Once sequential backbone assignments were completed, further characterisation of **xa'c** and **b'x** could be carried out using NMR spectroscopy.

4.3.4 Assignment of WT and I272A **b'x**

Triple resonance assignment of WT **b'x**, which have been previously carried out by other members of the group (Byrne et al., 2009), are displayed on the $^{15}\text{N}/^1\text{H}$ HSQC of WT **b'x** in figure 4.13.

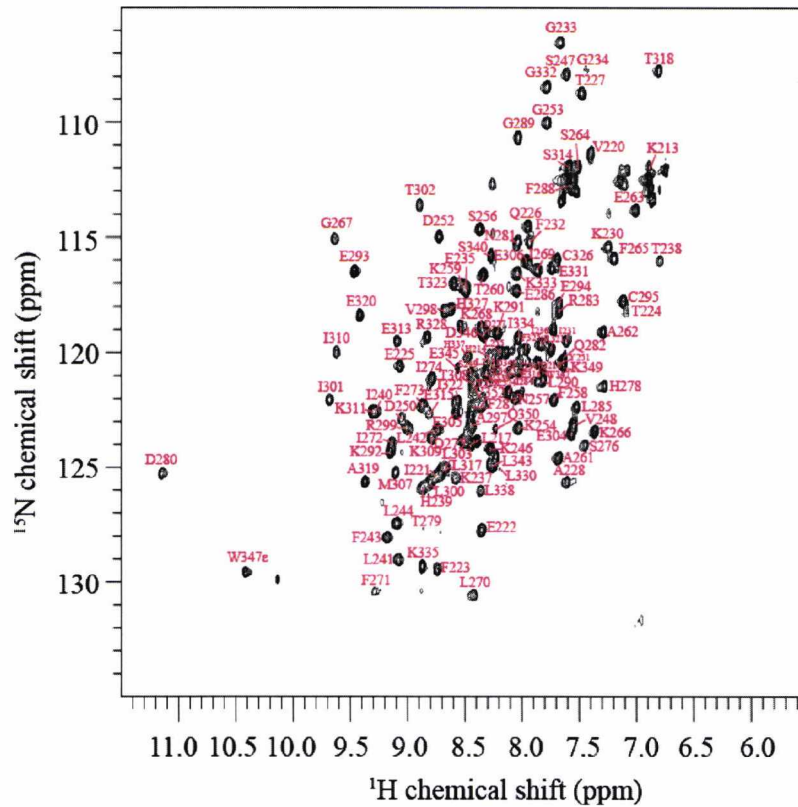


Figure 4.13 $^{15}\text{N}/^1\text{H}$ HSQC spectrum for WT **b'x**. All known amide resonance assignments are labelled in red on the spectrum. Experiments were run at 25°C with the protein sample at pH 6.5.

As the $^{15}\text{N}/^1\text{H}$ HSQC spectra of the I272A **b'x** mutant could be overlaid on WT **b'x**, the majority of the assignments could be mapped with confidence from the WT to the mutant protein peaks. Ambiguous assignments for peaks that had shifted as a result of the mutation were omitted from the I272A **b'x** spectra. Unmapped assignments were K213, V220, F222, F223, E225, Q226, L241, S264, I274, N281, I284, I285, F288, L290, I301, S314, E316, R321, K333, H337, M339, Q341, E342, W347 and D348. An assigned $^{15}\text{N}/^1\text{H}$ HSQC spectrum of I272A **b'x** is shown in figure 4.14.

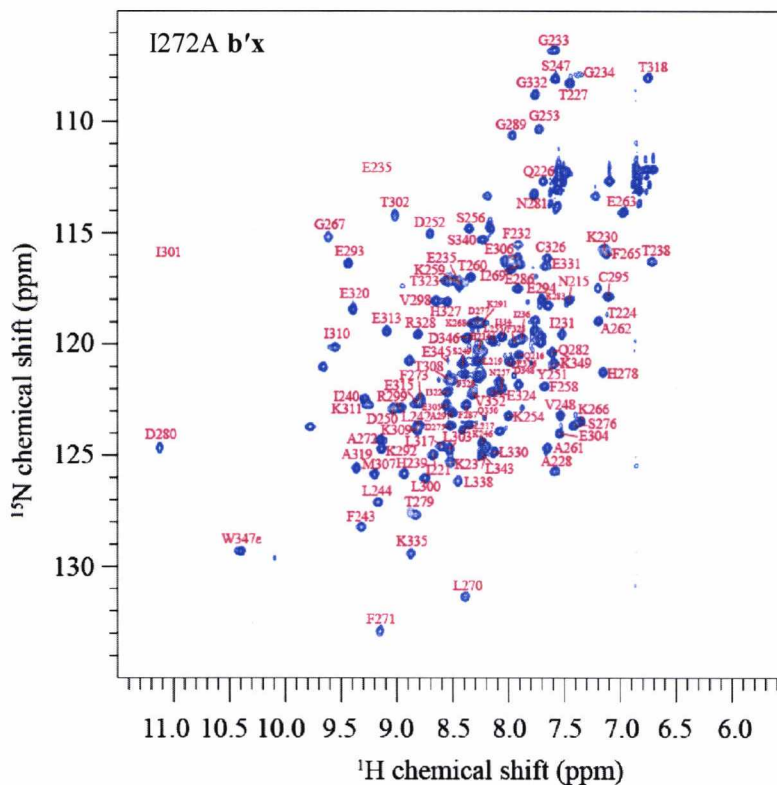


Figure 4.14 $^{15}\text{N}/^1\text{H}$ HSQC spectrum for I272A **b'x** with mapped assignments from WT **b'x**. All known amide resonance assignments are labelled in red on the spectrum. Experiments were run at 25°C with the protein sample at pH 6.5.

4.3.5 Secondary Structure Prediction using DANGLE

The assignment of resonances to $^1\text{H}_\text{N}$, ^{15}N , C_α and C_β for **xa'c** as well as existing $^1\text{H}_\text{N}$ and ^{15}N assignments for **b'x** enabled secondary structure prediction to be carried out using DANGLE. The program predicted Ψ and Φ angles and the most likely secondary structure type for each backbone residue generating one of three possible outcomes: C (coil), E (strand) and H (helix). Typical outputs for predictions for **xa'c** residues F365, F372, K392 and G381 are shown in figure 4.15 in the form of Ramachandran plots.

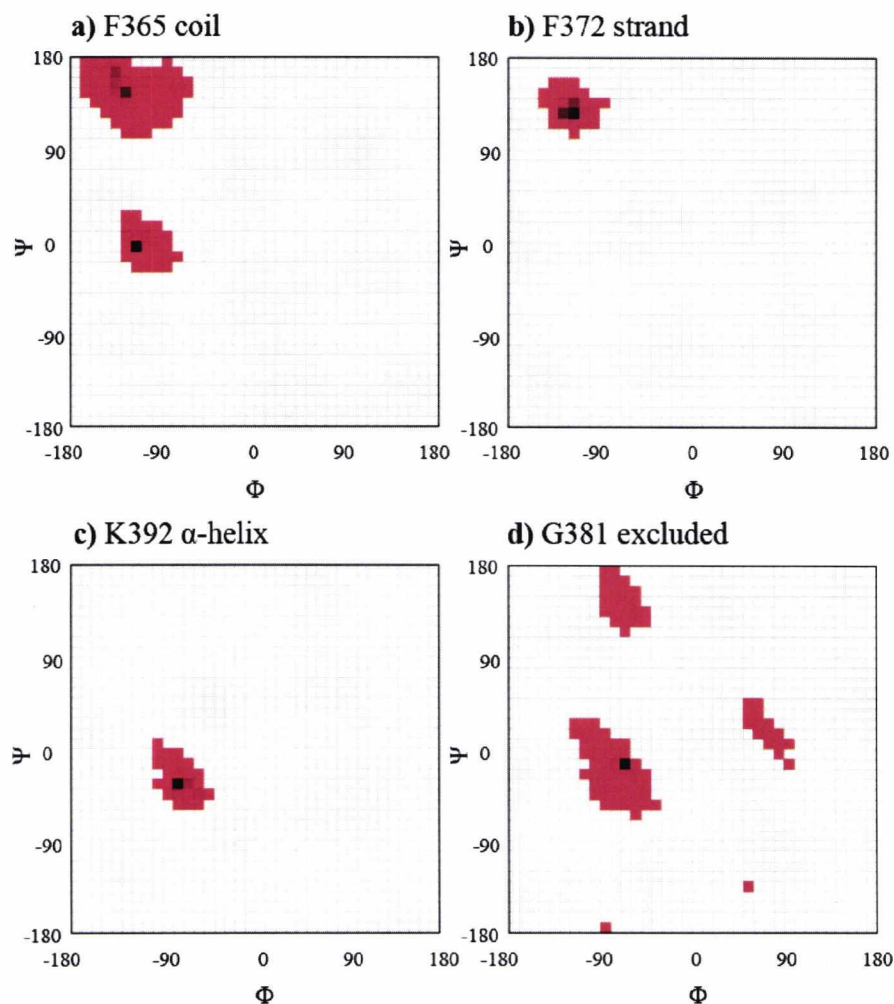


Figure 4.15 Examples of the types of structure given in DANGLE for secondary structure prediction. **a)** prediction of coil at this residue **b)** the single island predicts a β -strand and **c)** α -helix predicted **d)** where more than two islands at multiple locations were present, the residue was excluded from the prediction.

When angles were predicted in two islands such as those shown in figure 4.15a, coil was assigned to that residue. Angles predicted at a single location as in figure 4.15b and c suggested a β -strand and α -helix for those residues respectively. On occasions where more than two islands were present, as shown in figure 4.15d, the residue was rejected from the prediction.

Figure 4.16 summarises the predictions made by DANGLE for $\mathbf{xa'c}$, WT and I272A $\mathbf{b'x}$. The cylinders represent α -helix, arrows represent β -sheet and line coil. The residues in bold represented excluded data.

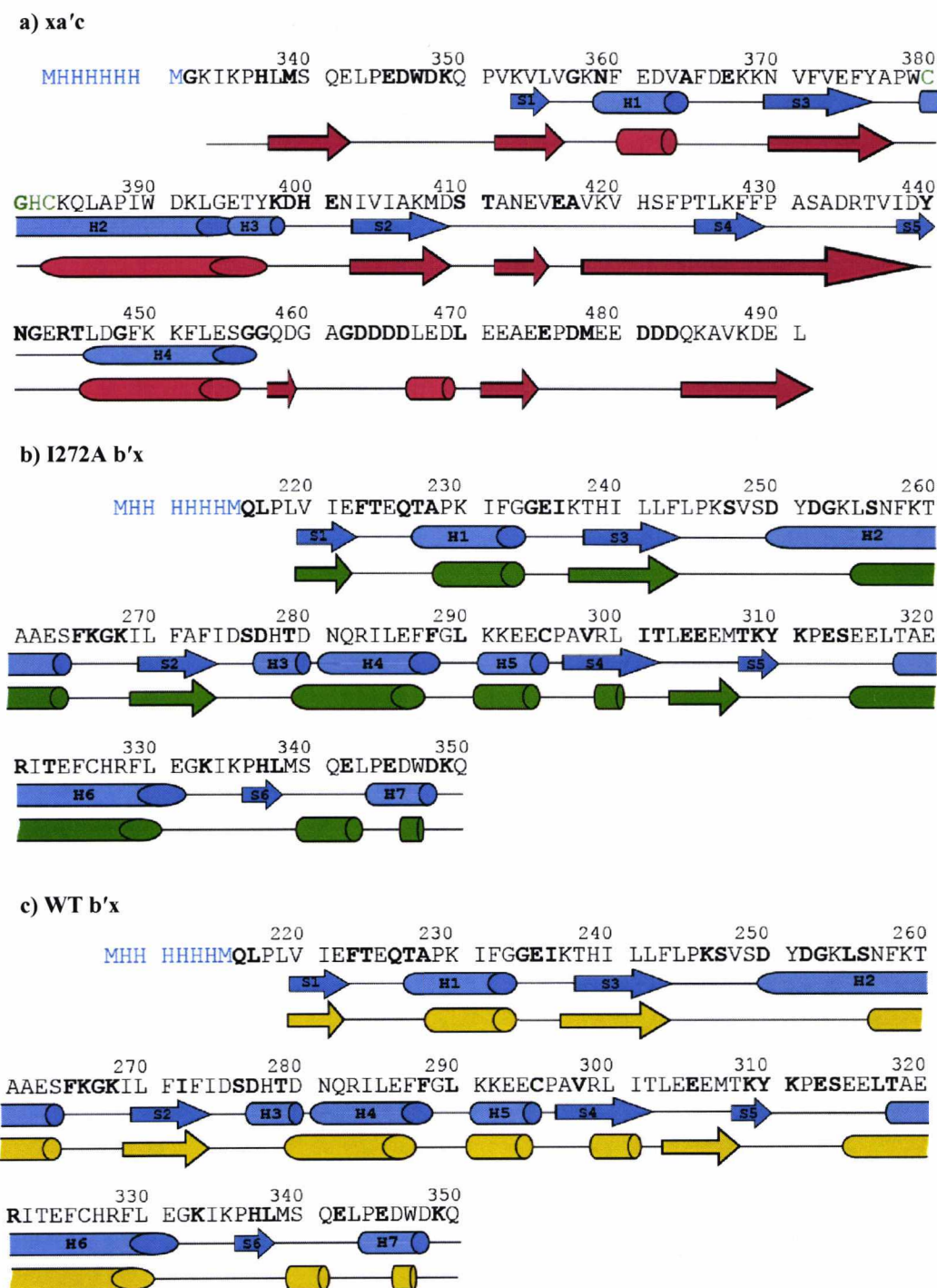


Figure 4.16 The predicted secondary structure of xa'c and WT and I272A b'x using DANGLE. The predicted structures are shown as a) red for xa'c b) green for I272A b'x and c) yellow for WT b'x, where α -helices are shown as cylinders, β -sheets as arrows and coils represented by lines. The His tag is shown in blue and the active site motif is shown in green. Residue numbers are those for mature PDI. The secondary structure motifs of a' (1X5C.pdb) and I272A b'x (3BJ5.pdb) are shown in blue for comparison under their respective amino acid sequences. The residues in bold were excluded from the structure prediction.

DANGLE was successful in predicting the secondary structures for **xa'c** and **b'x**, both WT and I272A. The predicted structure for **xa'c** and **b'x** were in agreement with the published structures for these constructs. Small differences were seen in the DANGLE prediction for **xa'c** where α -helices 2 and 3 were predicted as a single long helix, possibly due to their proximity to each other. The residues in the long loop between strands 2 and 4 were predicted to have β -sheet elements. The C-terminal extension has been predicted to consist of secondary structure elements such as a short α -helix and three β -sheets.

The predicted structures for WT and I272A **b'x** were very similar to each other and matched the I272A crystal structure with the exception of residues around strands 4 and 6, which DANGLE has predicted to consist of an α -helix, possibly because of a number of excluded residues in these regions of the sequence. Numbering of strands in the sequences of **a'** and **b'x** was taken from their respective PDB files 1X5C.pdb and 3BJ5.pdb.

4.3.6 Ligand binding

The ligand binding site on the **b'** domain of PDI has been mapped by NMR chemical shift perturbation in the presence of peptide ligands, such as Δ -somatostatin, which have been shown to compete with the **x** linker region for the binding site. As this is published data, ligand binding experiments with **b'x** in the presence of Δ -somatostatin were not repeated. Unlike **b'x**, **xa'c** did not appear to bind Δ -somatostatin as there were no visible chemical shift changes in the NMR $^{15}\text{N}/^1\text{H}$ HSQC spectra of the protein upon addition of the unlabelled peptide ligand (figure 4.17).

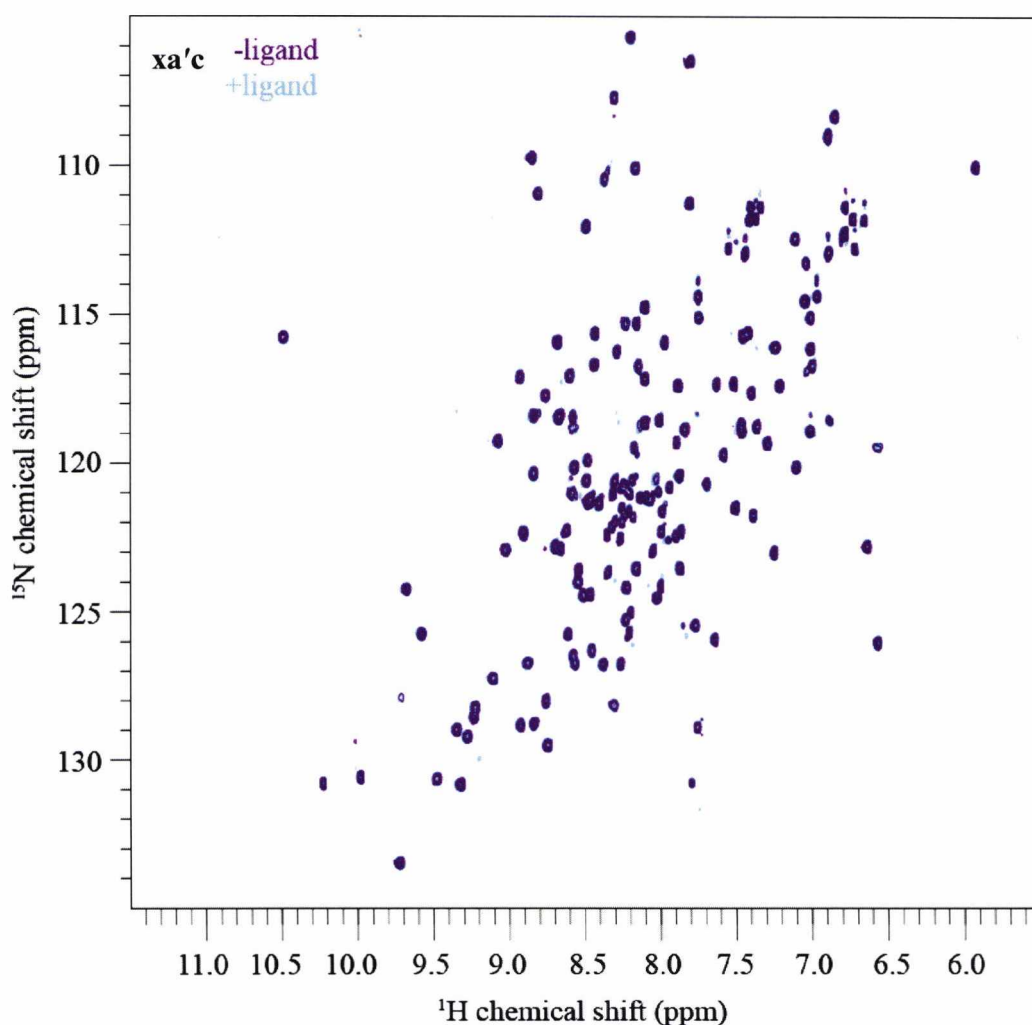


Figure 4.17 Overlaid $^{15}\text{N}/^1\text{H}$ HSQC spectra of **xa'c** in the presence and absence of Δ -somatostatin. Addition of the unlabelled peptide ligand to **xa'c** caused no chemical shift changes in the protein as the spectrum containing Δ -somatostatin (cyan) was exactly the same as the spectrum without Δ -somatostatin (purple).

4.3.7 ^{15}N NMR Relaxation Dynamics of **xa'c** and **b'x**

Tables of peak heights for ^{15}N T_1 , T_2 and hetNOE data as well as tables of results from model-free analysis for **xa'c** and **b'x**, both WT and I272A, are available in Appendix 4.5, 4.6 and 4.7 respectively. All graphs shown in this section were created from data in these appendices.

4.3.7.1 ^{15}N T_1 , T_2 and hetNOE data for **xa'c**

Figure 4.18 shows an example of the data fitting carried out in KaleidaGraph for the relaxation decay of **xa'c** residues D348 (located in the **x** linker region), W379 (located next to the active site motif just before the start of helix 2) and K427 (located in strand 4).

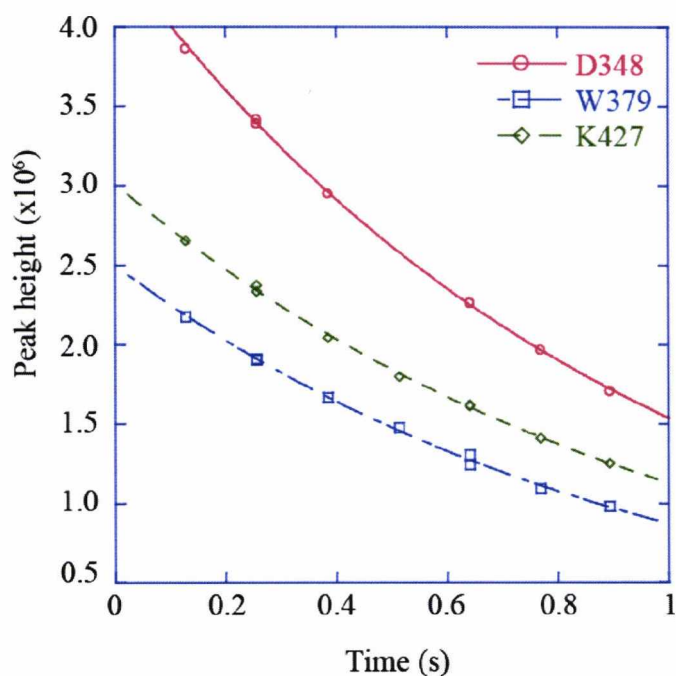


Figure 4.18 Example of the relaxation decay curves for the calculation of T_1 values with fitting by KaleidaGraph. Residues shown are D348 in red, W379 in blue and K427 in green.

Figure 4.19 displays plots of amide ^{15}N T_1 and T_2 on the same graph and hetNOE for **xa'c**. The secondary structure for **a'** (1X5C.pdb) is shown at the top of the graphs. Values for ^{15}N T_1 and T_2 times and their respective errors were calculated using equation 4.6 in CCPN Analysis. The mean ^{15}N T_1 and T_2 values obtained from rplot were 952 ± 52 ms and 68 ± 6 ms, respectively, giving a mean T_1/T_2 ratio of 14. Rplot also provided a correlation time, τ_m , of 11.8 ns and an average S^2 value of 0.87 ± 0.02 . This set of data

was then exported into ModelFree4 in order to obtain individual residue model-free parameter results. Deviations in data were observed in loops between stands 2 and 4 as well as strands 4 and 5. Greatest deviations were seen at the termini of the protein, with the longest T_2 times and lowest hetNOE values observed at the C-terminus.

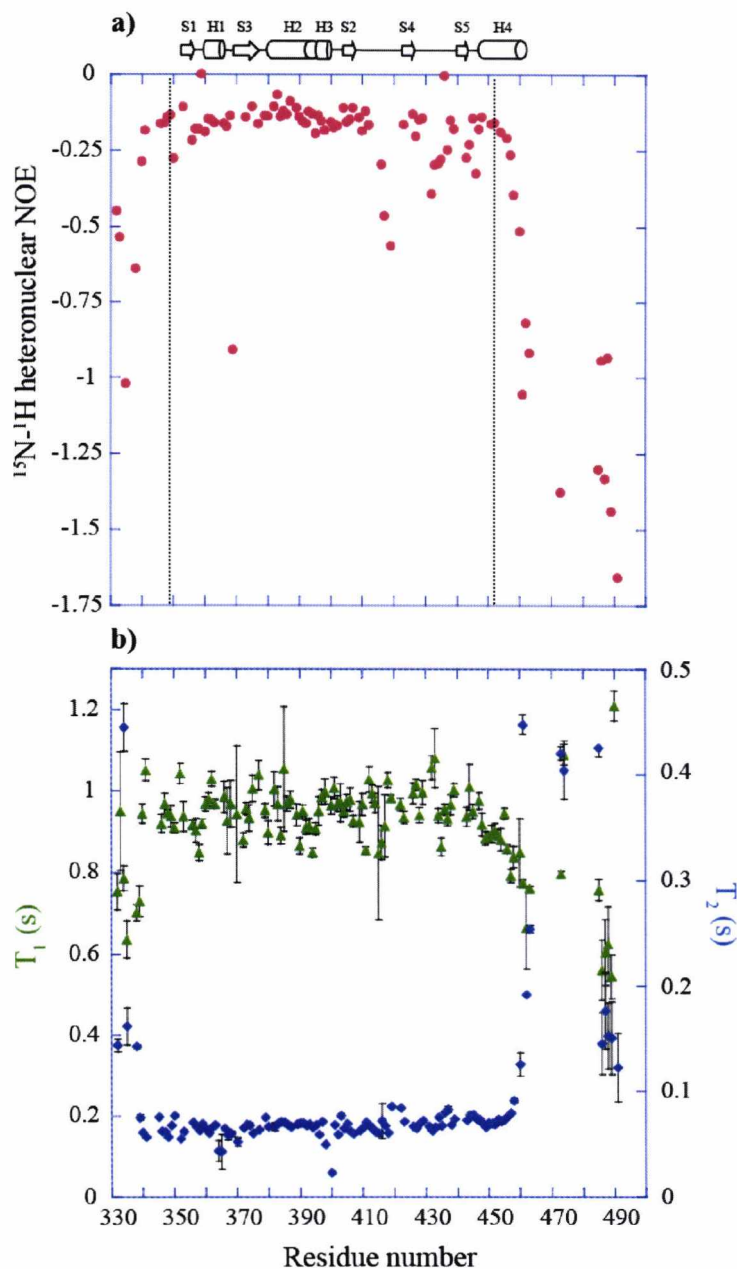


Figure 4.19 **xa'c** ^{15}N T_1 , T_2 and heteronuclear NOE data. a) hetNOE data is shown in red and b) T_1 and T_2 data shown in green and blue, respectively. Values for the error bars were taken directly from CCPN Analysis. The schematic secondary structure of **a'** is displayed at the top of the figure with cylinders as helices and arrows as strands.

4.3.7.2 ^{15}N T_1 , T_2 and *hetNOE* data for **b'x**

Figure 4.20 displays ^{15}N T_1 , T_2 and *hetNOE* data for WT **b'x** and its I272A mutant. The secondary structure of the I272A **b'x** mutant (3BJ5.pdb) is shown at the top for each protein.

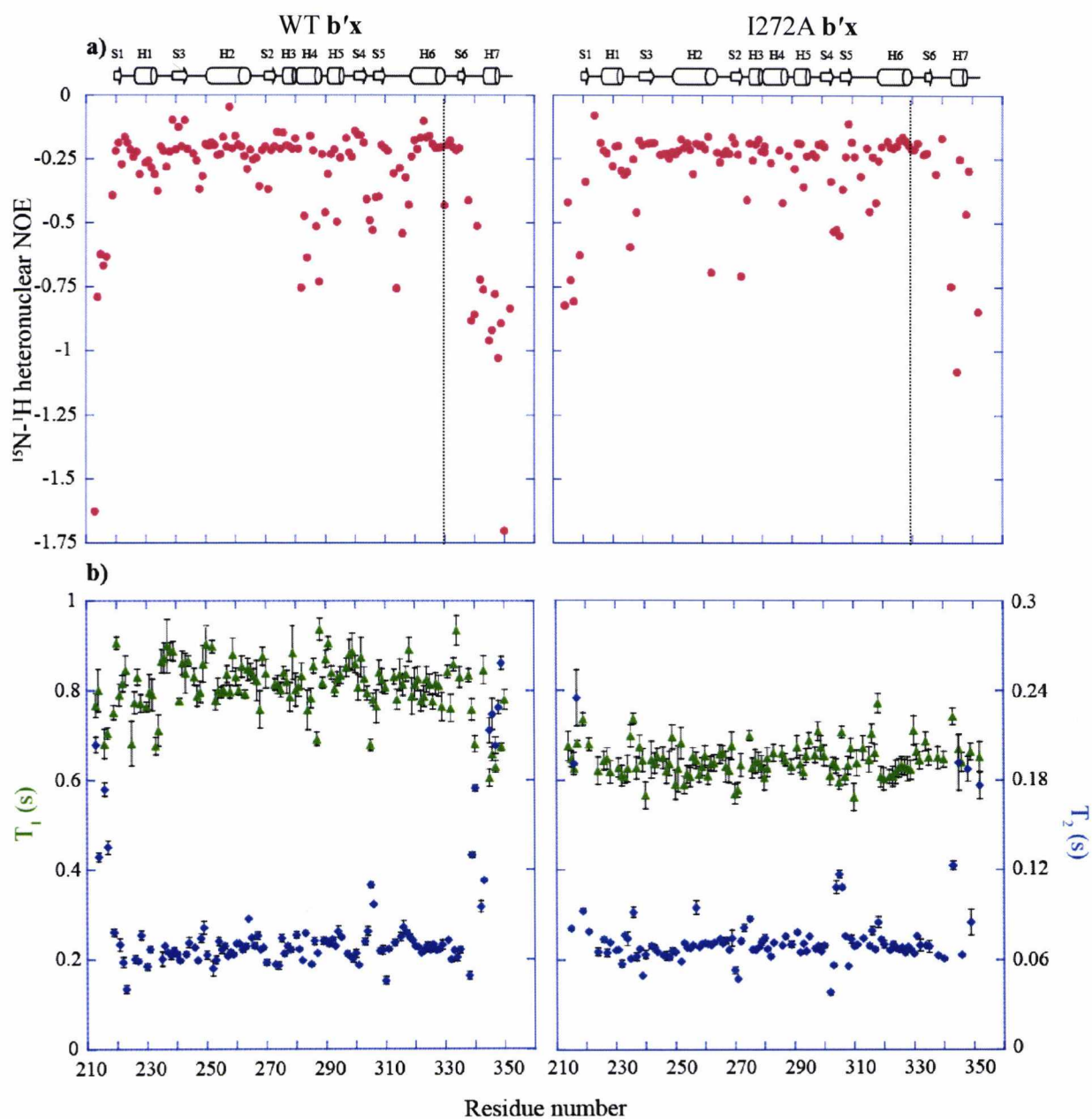


Figure 4.20 ^{15}N T_1 , T_2 and heteronuclear NOE data for WT **b'x** and its I272A mutant. a) *hetNOE* data for each protein is displayed in red with WT **b'x** on the graph on the left and I272A on the right. b) T_1 and T_2 data are shown in green and blue, respectively. The schematic secondary structure of I272A **b'x** is displayed at the top of the figure with cylinders as helices and arrows as strands.

Values for ^{15}N T_1 and T_2 times and their respective errors were calculated using equation 4.6 in CCPN Analysis as for **xa'c**. The mean ^{15}N T_1 values calculated from rplot were 833 ± 37 ms and 645 ± 22 ms for WT and I272A **b'x** respectively. The mean T_2 values for each of the proteins were 70 ± 6 ms for WT and 71 ± 2 ms for I272A **b'x**. The mean T_1/T_2 ratios were 11.9 for WT **b'x** and 9.1 for the I272A mutant. Surprisingly, the T_1 times for I272A appear to be low compared to the WT protein, whereas T_2 values are very similar to those of WT **b'x**. Rplot calculated average τ_m and S^2 values of 10.7 ns and 0.93 ± 0.02 respectively for WT **b'x**. However due to the abnormally low T_1 values for I272A it was not possible to calculate an accurate τ_m value for I272A but an average S^2 value of 0.91 ± 0.03 was taken from rplot. Deviations in hetNOE values between WT and I272A **b'x** highlighted the different dynamic properties of the **x** linker region, as in WT **b'x** hetNOE values drop after K333 whereas in I272A it is not until L343 that a decrease in hetNOE values is observed.

4.3.7.3 ^{15}N T_1 versus T_2 relaxation data for **xa'c** and **b'x**

A plot of T_1 against T_2 is a simple analysis of relaxation data that allows the first insight into the motional events within proteins. Figure 4.21 shows a plot of T_1 vs T_2 for ^{15}N amides of **xa'c**, WT **b'x** and I272A. Data points considered to be outliers are coloured in black and were excluded from rplot analysis. Both **xa'c** and WT **b'x** fall inside the theoretical lines for the order parameter S^2 . Almost all of the I272A **b'x** residues fall below the $S^2 = 1.0$ contour line suggesting that I272A has an average S^2 value that is greater than 1.0 which is not theoretically possible in model-free, but this anomaly is due to the low T_2 values observed in figure 4.20b. This observation now explains the differences in figure 4.20b between WT and I272A **b'x** as demonstrating a reduction in T_2 for I272A and not T_1 , making these times similar for those observed for WT. In fact, all those residues that fall under the $S^2 = 1.0$ line could be undergoing conformational exchange and suggest contribution from R_{ex} . If no R_{ex} contribution existed in the I272A **b'x** data, the expected T_2 values would be larger and all I272A points in figure 4.21 would fall within the $S^2=1.0$ line. Visual observation of both **xa'c** and **b'x** proteins in the T_1 vs T_2 plot showed them to be anisotropic due to the elongated shape the residues form in the plot.

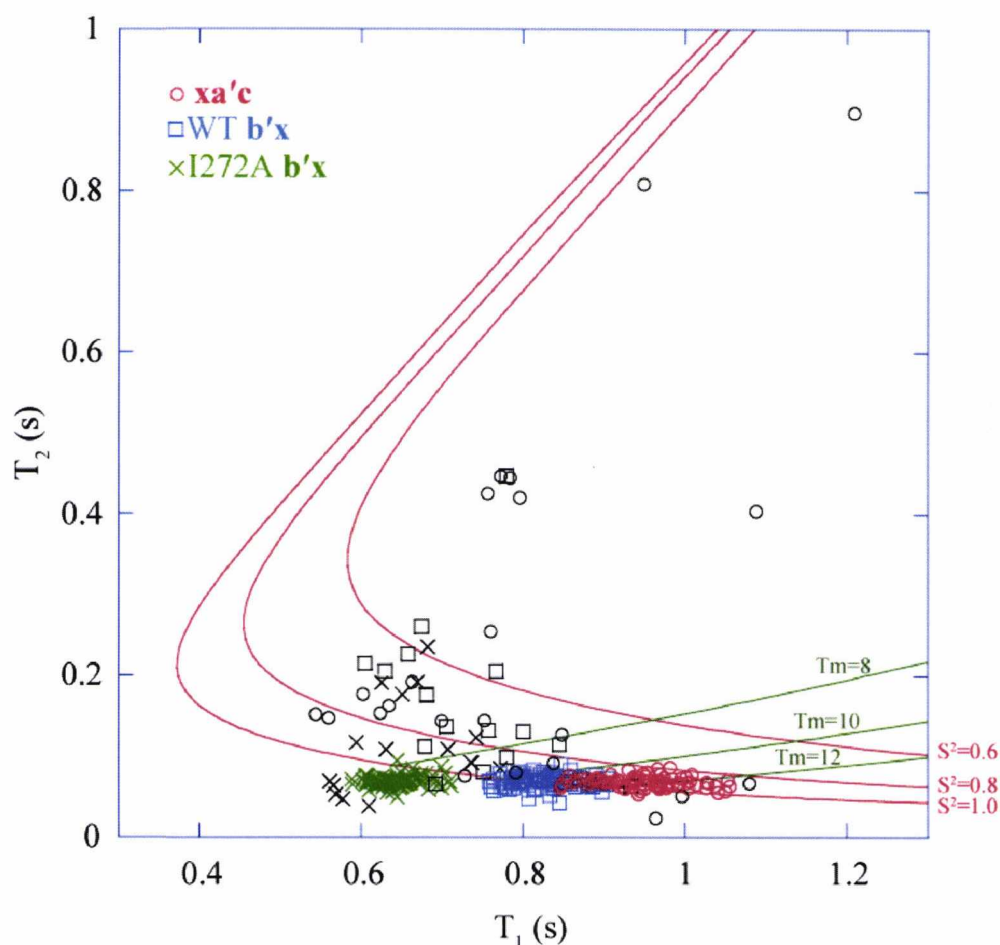


Figure 4.21 Plot of experimental T_1 against T_2 of amide protons for **xa'c**, **WT b'x** and **I272A**. Experimental data plotted with theoretical values of order parameter $S^2 = 0.6, 0.8$ and 1.0 and τ_m of $8, 10$ and 12 ns when τ_e is 50 ps at 600 MHz. **xa'c** data are shown in red circles, **WT b'x** in blue squares and **I272A b'x** in green crosses. Residues coloured in black that deviate from the main plot were excluded from rplot analysis.

The theoretical τ_m and S^2 contour lines in figure 4.21 were drawn using model values of ^{15}N T_1 and T_2 with a constant τ_e of 50 ps and do not take into account hetNOE values, but they are useful in estimating τ_m values required for model-free analysis. **I272A b'x** was estimated to have an average τ_m of 8.0 ns from the T_1 versus T_2 plot due to the low T_2 values. Table 4.3 shows a summary of the τ_m values exported from rplot, which uses a plot of T_1 vs T_2 to calculate the average τ_m . The expected τ_m values for **xa'c** and the **b'x** proteins, if isotropic, were estimated using equation 4.1 which takes into account the number of residues in the protein.

Protein	Molecular weight (kDa)	Theoretical τ_m (ns) based on MW	Measured τ_m (ns) from rplot
WT b'x	16.8	10.1	10.7
I272A b'x	16.8	10.1	8.0
xa'c	19.3	11.8	11.8

Table 4.3 A summary of the τ_m times estimated from experimentally derived ^{15}N T_1 and T_2 times. The theoretical τ_m values were calculated using the Daragan and Mayo equation 4.1 which takes into account the residue number and therefore the molecular weight for each protein.

The average τ_m value obtained for **xa'c** matched the expected theoretical value exactly for a protein of its size. The average τ_m observed for WT **b'x** was also very close to the expected τ_m value. Conversely, the I272A **b'x** mutant displayed a τ_m value that was ~ 2 ns shorter than the expected value, and as I272A tumbles faster in solution than WT **b'x** and **xa'c** it must be more structurally compact than WT **b'x**.

4.3.7.4 ^{15}N Model-free analysis of **xa'c** and **b'x**

All three relaxation measurements T_1 , T_2 and hetNOE were used to define the molecular motion of each protein using the Modelfree4 program. Figure 4.22 summarises the model-free data for **xa'c**, WT **b'x** and the I272A mutant. τ_e values obtained from Modelfree4 for **xa'c** showed that most of the residues have internal motion in the region of 100 ps or less, justifying the use of 50 ps for τ_e used in the contour lines in figure 4.20. Exceptions were seen for residues at the termini of the protein, in the **x** linker and the **c** extension, as well as residues K358, G394, T411, E416, A417, G456, G457, Q458. S^2 values for this protein displayed a similar trend with the residues in **x** and **c** having lower average S^2 values of 0.32 ± 0.02 and 0.16 ± 0.03 respectively, than the rest of the protein, suggesting fast internal motion and high flexibility in these regions of the protein. WT **b'x** presented similar trends in dynamics with the residues in the **x** linker region showing lower S^2 values, 0.35 ± 0.05 , than the rest of the protein. As with **xa'c**, loops in the secondary structure also displayed lower S^2 values than the structured regions. A larger number of WT **b'x** residues displayed τ_e values greater than 100 ps, suggesting a higher degree of internal motion for this protein than **xa'c**. Only two residues in **xa'c** displayed contribution

from R_{ex} , K398 and H400, whereas in WT **b'x**, a number of residues making up structured region of the protein, such as strand 2 and helix 4, displayed conformational exchange. These residues are V220, E222, F223, E225, Q226, T227, K230, F232, G233, S247, D252, G253, L270, F271, I272, F273, I274, Q282, L285, I301, T302, I322, G332 and L338. The majority of residues in I272A **b'x** showed contribution from R_{ex} , suggesting that the whole protein was undergoing conformational exchange. As for WT **b'x**, the **x** region in I272A showed lower S^2 values than the rest of the protein but to lesser extent than the WT protein.

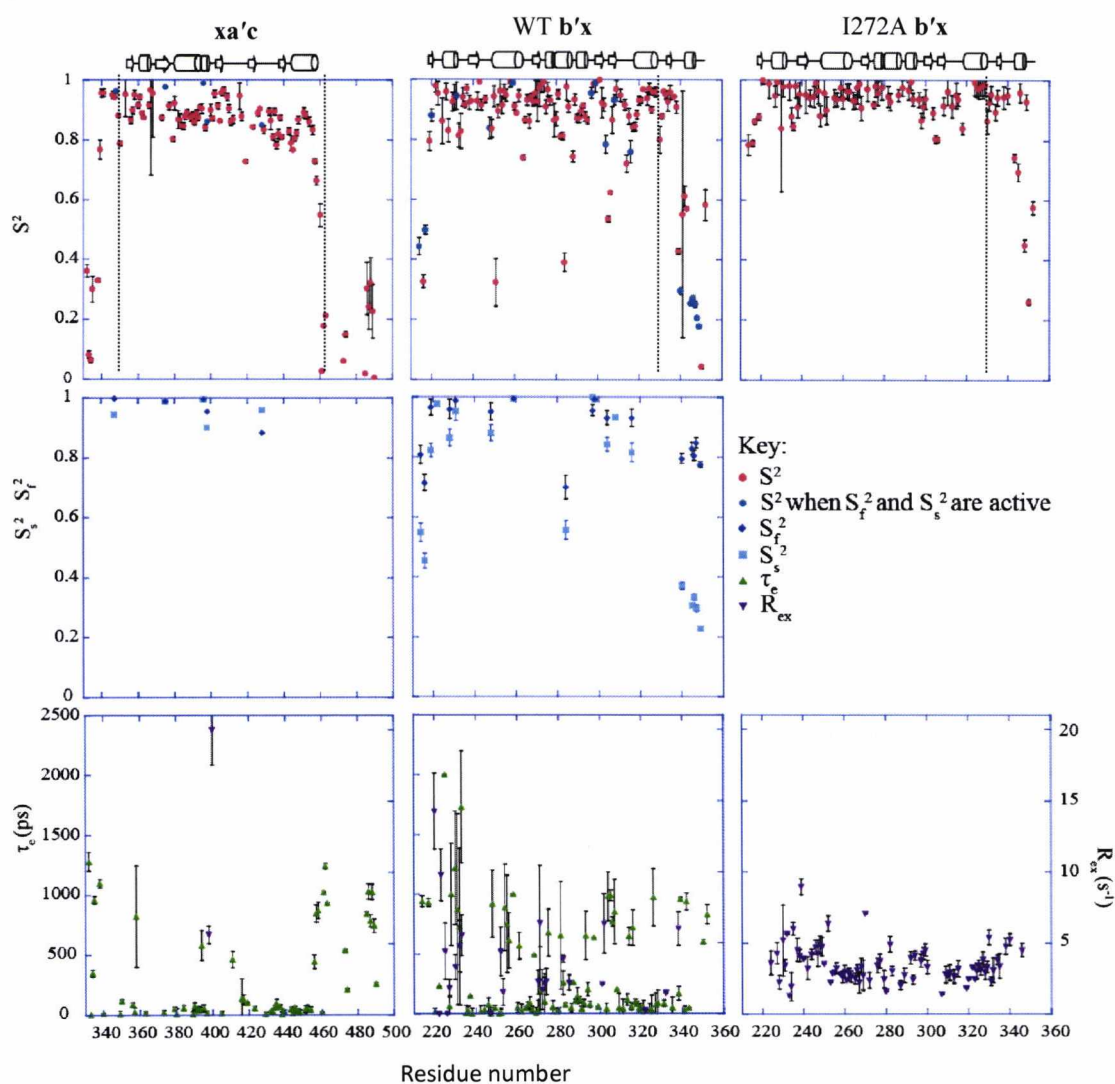


Figure 4.22 Graphs of model-free fit parameters for the characterisation of motion of **xa'c**, WT **b'x** and I272A **b'x**. ^{15}N T_1 , T_2 and hetNOE values for each protein were used with analysis program Modelfree4 to calculate residue specific order parameter values of S^2 (S_s^2 and S_f^2 where needed) and τ_c . R_{ex} contribution was used for those residues that were detected by Modelfree4 to display conformational exchange.

4.4 Discussion

4.4.1 The conformational stability of the **a'** domain

Studies on the conformational stability of PDI constructs, in chapter 3, showed that the **b'xa'c** fragment of PDI is a challenging construct to study due to the low conformational stability of the **a'** domain. Indeed, expression of the **a'** domain in isolation (domain boundaries D348 – G462) in *E. coli* proved difficult as protein yields were too low for NMR studies. However, addition of the acidic **c** tail improved expression levels suggesting that **c** may be involved in increasing the stability of **a'** and in turn regulating expression of PDI in addition to its role in retaining the protein in the ER. The resulting protein, **a'c**, was so conformationally active and exchange broadened beyond detection that it gave an initial appearance that the protein was unfolded, but close analysis of $^{15}\text{N}/^1\text{H}$ HSQC NMR spectra indicated insufficient number of peaks to account for all residues. Reduction of the **a'** active site proved that the **a'** domain was never unfolded, but in fact experiencing a high degree of conformational exchange and the oxidation state of this domain played a crucial role in its conformational stability, as seen previously in chapter 3. Recently it was shown that the structure of the reduced **bb'xa'** revealed that **a'** was tightly packed with both the **b'** domain and the **x** linker region to form a compact structural module. Oxidation of **a'** released the protein from this compact conformation in order to facilitate its high chaperone activity (Wang et al., 2012b). Therefore, it is not surprising that addition of DTT to **a'** and **a'c** had such a dramatic effect on the peak dispersion and resolution of the NMR spectra and supports the observations from the crystal structure seen by Wang and co-workers (Wang et al., 2012b). The **x** linker region also increased the conformational stability of the **a'** domain even further as the reduced **xa'c** construct presented with such good quality spectra that it could be used for further NMR studies, such as triple resonance assignments and relaxation dynamics. Minimal chemical shift mapping of **xa'c** against **a'c** revealed shifts throughout the **a'** domain, especially the core β -sheet, as can be seen in figure 4.9, suggesting that **x** could be packing into the **a'** domain and as a consequence lowering its degree of conformational exchange. So, the **x** linker region not only binds to the ligand binding site on the **b'** domain, an event known as “capping”, but it may also interact with the **a'** domain and consequently affecting its conformational stability. Interaction of the **x** linker region with the **a'** domain has been recently reported in the crystal structure of the **bb'xa'** construct in which hydrogen bonds are formed between **x** and **a'** residues M339–E414, Q341–N413, S334–G357, W347–V354

and D348–K353 (Wang et al., 2012b). The effect of temperature and pH on the conformational stability of **xa'c** were also investigated as they had a major impact on the **b'xa'c** protein, the results of which are discussed in the next chapter. It was found that running NMR experiments at 25°C with the protein in a pH 7.0 buffer gave suitable data and did not compromise the conformational stability of **xa'c**.

4.4.2 Sequential backbone assignment of **xa'c**

The percentage of NMR backbone assignments for **xa'c** was reported in section 4.3.3. 100% assignment was not possible due to the challenging nature of the **a'** domain and the presence of the acidic **c** tail of the protein which gave very sharp resonances in the random coil region of the spectrum between 8.0 and 8.5 ppm. The assignments carried out using the triple resonance experiments CBCA(CO)NH and CBCANH as well as the residues which could not be assigned are shown in appendix 4.4. ^1H , ^{13}C and ^{15}N resonance assignment of the **a'** domain has been previously carried out but the assigned protein was lacking the C-terminal extension, **c**, as well as the **x** linker region (Dijkstra et al., 1999) and as figure 4.9 showed, the presence of **x** and **c** causes significant shifts in the backbone ^1H and ^{15}N resonance of the **a'** domain, so these assignments were essential for further studies. Figure 4.23 shows the amino acid sequence of **xa'c** with the residues that could not be assigned from triple resonance experiments, due to missing C_α or C_β resonances or ambiguous data, highlighted in red. The secondary structure of the **a'** domain (1X5C.pdb) as well as the predicted secondary structure from DANGLE are also shown.

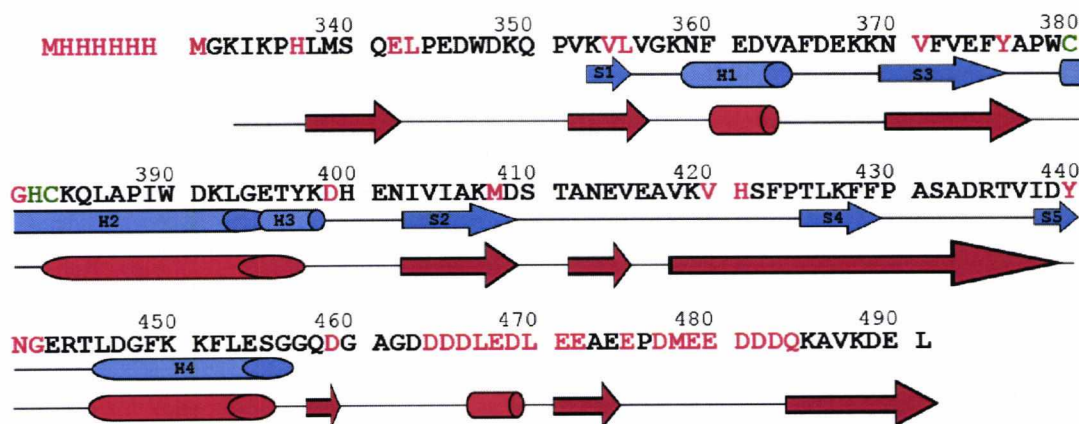


Figure 4.23 Amino acid sequence of **xa'c** showing the unassigned residues. The residues that could not be assigned from triple resonance experiments are shown in red and the active site residues in green. The secondary structure elements of the **a'** structure and those of the DANGLE predictions are displayed in blue and red respectively.

The majority of the unassigned residues were present in the C-terminal region of the protein due to sharp resonances from the **c** extension present in the random coil region of the spectrum, between 8.0 and 8.5 ppm. This suggests that the **c** tail of PDI is in fact unstructured despite the secondary structure elements predicted by DANGLE. This highlights one of the differences between yeast and human PDI, as the acidic C-terminal extension of yeast PDI forms a helix in the crystal structure (Tian et al., 2006).

Protein backbone NMR assignments for **xa'c** were invaluable for providing $^{15}\text{N}/^1\text{H}$ HSQC assigned spectra which could be used for backbone dynamics studies of this construct.

4.4.3 Secondary Structure Prediction using DANGLE

DANGLE was used to predict the secondary structure of **xa'c** as well as WT and I272A **b'x** using the protein NMR data analysis software package CCPN Analysis Version 2. The package used the sequence information and the ^1H and ^{15}N chemical shifts for the three proteins as well as the C_α and C_β shifts for **xa'c** to predict the likely secondary structure. Overall, DANGLE was successful in predicting the secondary structure of **xa'c**, WT and I272A **b'x**, as can be seen from figure 4.16, proving confidence in the chemical shift data being used in further NMR studies.

The discrepancies seen in the structure predictions, such as helices 2 and 3 in **xa'c** are predicted as a single helix and in **b'x** strand 6 in **x** is predicted as a helix, could be due to the fact that the software interrogates chemical shift values used in conjunction with published protein structures to generate a prediction rather than calculate an NMR structure, therefore some deviations are to be expected.

4.4.4 Ligand binding

The **b'** domain of PDI has been shown to provide the ligand binding site which has been mapped by NMR and shown to consist of residues from the core β -sheet and α -helices 1 and 3 highlighted in figure 1.10 in the introduction (Byrne et al., 2009). To date there is no evidence to show that the **a'** domain in isolation is capable of binding peptide ligands without cysteine residues, however it does contribute to binding of larger misfolded proteins. Therefore it was not surprising to find that the **xa'c** construct did not bind Δ -somatostatin as it does not contain a ligand binding site.

4.4.5 ^{15}N NMR Relaxation Dynamics of **xa'c** and **b'x**

We have already seen in chapter 3 that PDI constructs **xa'c**, WT **b'x** and I272A **b'x** have different conformational stabilities as studied by intrinsic fluorescence in the presence of denaturant. Therefore it is not surprising that the backbone dynamic properties of these proteins are also very diverse, not only between the different domains but also between the WT and mutant **b'x**. To date, there are no published NMR relaxation studies on human PDI or any of its domains separately, therefore the molecular motions of this protein in solution are still unknown but crystallographic data from hPDI and yPDI suggest that conformational motions do exist in the protein (Wang et al., 2012a, Tian et al., 2008).

4.4.5.1 Relaxation data

All three relaxation parameters, ^{15}N T_1 , T_2 and hetNOE, must be taken into account when describing the molecular motions of the proteins under investigation. HetNOE values for **xa'c**, WT and I272A **b'x**, which report on internal motion, successfully highlight loops and regions of structure in all three proteins. Upon visual observation of the hetNOE graph of **xa'c** it is instantly obvious that the C-terminal extension, **c**, is unstructured and free in solution due to the low hetNOE values and motions in the sub-nanosecond timescale compared to the rest of the protein, in agreement with the **xa'c** assignments described in section 4.4.2. Figure 4.19a also presents evidence that the **x** linker region in **xa'c** must be behaving as a structured part of the protein, possibly due to interaction with the **a'** domain. The secondary structure for the **a'** domain displayed at the top of the hetNOE graph in 4.19a shows that strand 1 of **a'** starts at residue 352, but residues 340-351 of **x** also appear structured, adding to the evidence that **x** must be interacting with the **a'** domain. The I272A **b'x** crystal structure shows **x** consists of a β -strand in residues P336 – L338 as well as an α -helix made up of residues P344 – D348. The strand in **x** is conserved in the recent **bb'xa'** structure (3EUM.pdb), where it is present in residues M339 – S340, whereas the α -helix does not appear as part of the structure in this construct. T_1 and T_2 values for **xa'c** are also in agreement with the hetNOE data with the longest T_2 and shortest T_1 values observed at the unstructured **c** tail.

HetNOE data for **b'x** highlighted flexible regions of the WT and I272A proteins. Figure 4.24 shows the crystal structure of the I272A mutant with the deviated hetNOE values for the WT (a) and I272A (b) proteins.

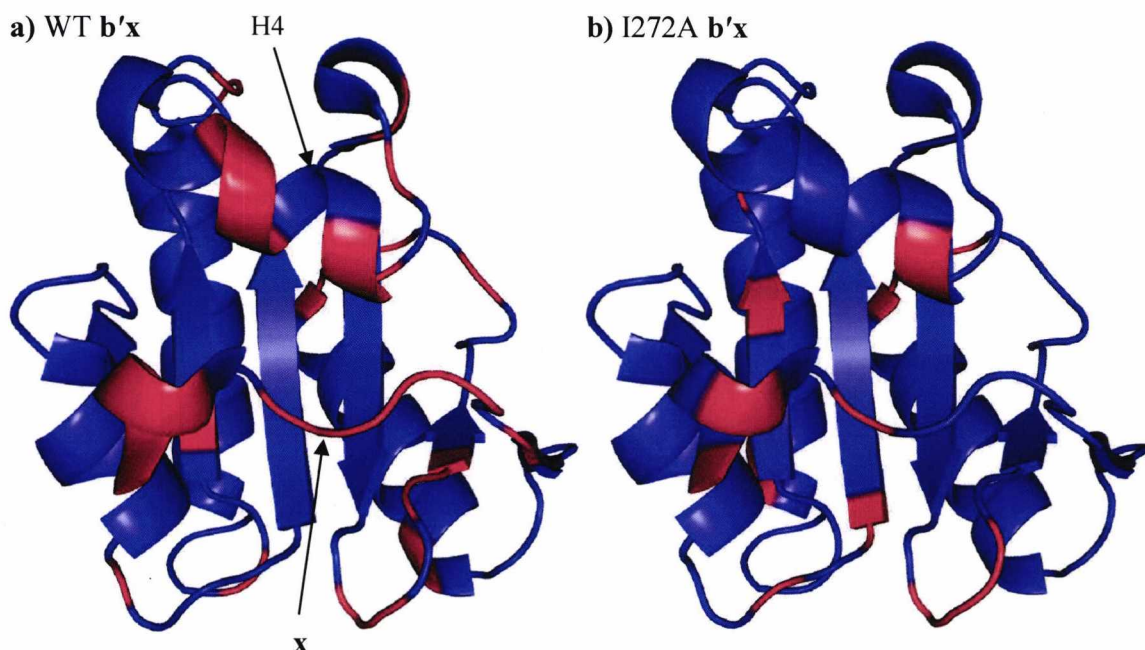


Figure 4.24 I272A **b'x** crystal structure (3BJ5.pdb) showing flexible regions highlighted by the hetNOE data. Residues in red show were deviations from the normal hetNOE values (>-0.5) for a) WT and b) I272A **b'x**.

WT **b'x** appeared more flexible than I272A, especially in the **x** region and helix 4, possibly due to the fact that WT **b'x** exists in equilibrium as a mixture of capped and uncapped protein, whereas the I272A mutant favours the uncapped conformation therefore displaying a more rigid backbone. Residues involved in ligand binding in the core β -sheet are highlighted in both proteins, confirming interactions with the **x** linker region.

T_1 values are sensitive to molecular motions in the nano to sub-nanosecond timescale, whereas T_2 times are also sensitive to motion in the micro to millisecond timescale as well as the overall size of the molecule and conformational exchange. Both **xa'c** and WT **b'x** present with T_1/T_2 and global correlation times of 11.8 ns and 10.7 ns, at 25°C, appropriate for their respective sizes. The estimated τ_m at 25°C for **b'x** is 10.1 ns assuming a globular compact shape. **xa'c** has a longer τ_m than **b'x** so tumbles slower in solution. This is not surprising as **xa'c** not only has a higher molecular weight than **b'x** but also a larger hydrodynamic volume, confirmed by gel filtration data in figure 4.5, due to the presence of the unstructured **c** tail. In contrast, the I272A **b'x** mutant displayed T_2 values that were too low for a protein of 143 amino acids and as a consequence the estimated global correlation time was 8.0 ns, corresponding to a protein of 111 amino acids

long, assuming a compact and spherical molecule at 25°C and using the relationship in equation 4.1. As mentioned above, protein conformational exchange is usually detected by a change in T_2 times and T_1 times are less affected due to the longer time window. Although initial observations in figure 4.20b showed that the T_1 values for the I272A $\mathbf{b'x}$ mutant were lower than expected, it is in fact the T_2 values that are attenuated when taking into account the global correlation time of the protein at 8 ns rather than the expected 10 ns like WT $\mathbf{b'x}$. Figure 4.25 shows a simple representation of the relationship between the global correlation time, τ_m , and the expected T_1 and T_2 values.

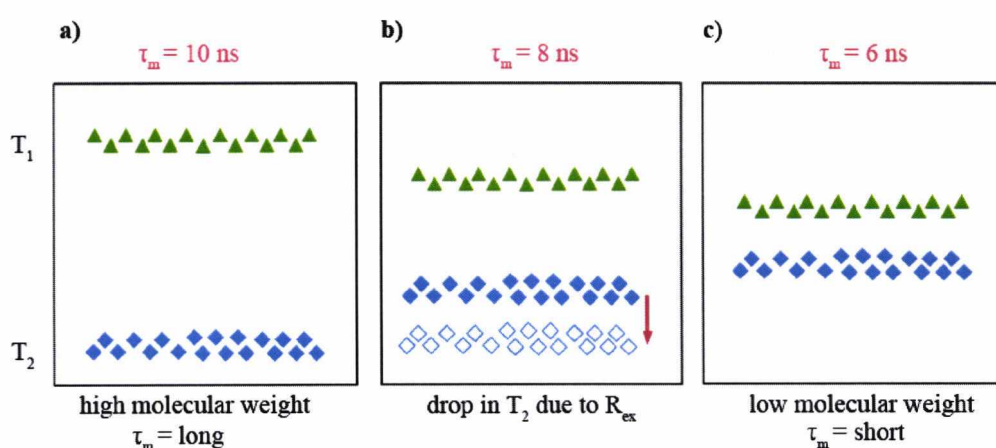


Figure 4.25 The relationship between τ_m and T_1 and T_2 values. Example T_1 (\blacktriangle) and T_2 (\blacklozenge) values for proteins exhibiting global correlation times of a) 10 ns b) 8 ns and c) 6 ns. Larger proteins present with higher T_1 and lower T_2 values than smaller proteins. T_2 values are sensitive to conformational exchange and become attenuated as shown in b) where T_1 values remain constant and T_2 values are lowered (\blacklozenge).

As already mentioned, the global correlation time of a protein is dependent on its molecular weight, as larger molecules tumble slower in solution than smaller molecules. As τ_m takes into account T_1 and T_2 values, these are also dependent on the molecular weight of the protein. Therefore a larger molecule with a τ_m of 10 ns, as shown in figure 4.25a, presents with T_1 values that are higher and T_2 values that are lower than those of a smaller molecule which tumbles at 6 ns in figure 4.25c. In the case of $\mathbf{b'x}$, the WT protein presents with the expected τ_m , T_1 and T_2 values for a protein of its size. Conversely, the I272A mutant tumbles faster in solution, with a τ_m of 8 ns suggesting a more compact conformation than the WT protein. Its T_1 values are as would be expected of a protein with that particular correlation time, but T_2 values are lower than expected, as shown in

figure 4.25b, suggesting conformational exchange between two or more states of I272A **b'x**.

4.4.5.2 Model-free analysis

Model-free analysis of **xa'c**, WT and I272A **b'x** was carried out to further characterise the motional events occurring within these proteins. All three proteins presented with similar average S^2 values of 0.87 ± 0.02 for **xa'c**, 0.93 ± 0.02 for WT **b'x** and 0.91 ± 0.03 for I272A **b'x** indicating the backbone rigidity of the thioredoxin fold. **xa'c** has a slightly lower S^2 than **b'x** possibly due to the lower conformational stability of **a'** compared to the **b'** domain. Surprisingly, **xa'c** had only two residues with contribution from R_{ex} , K398 and H400 but a number of residues exhibiting internal motion above 100 ps.

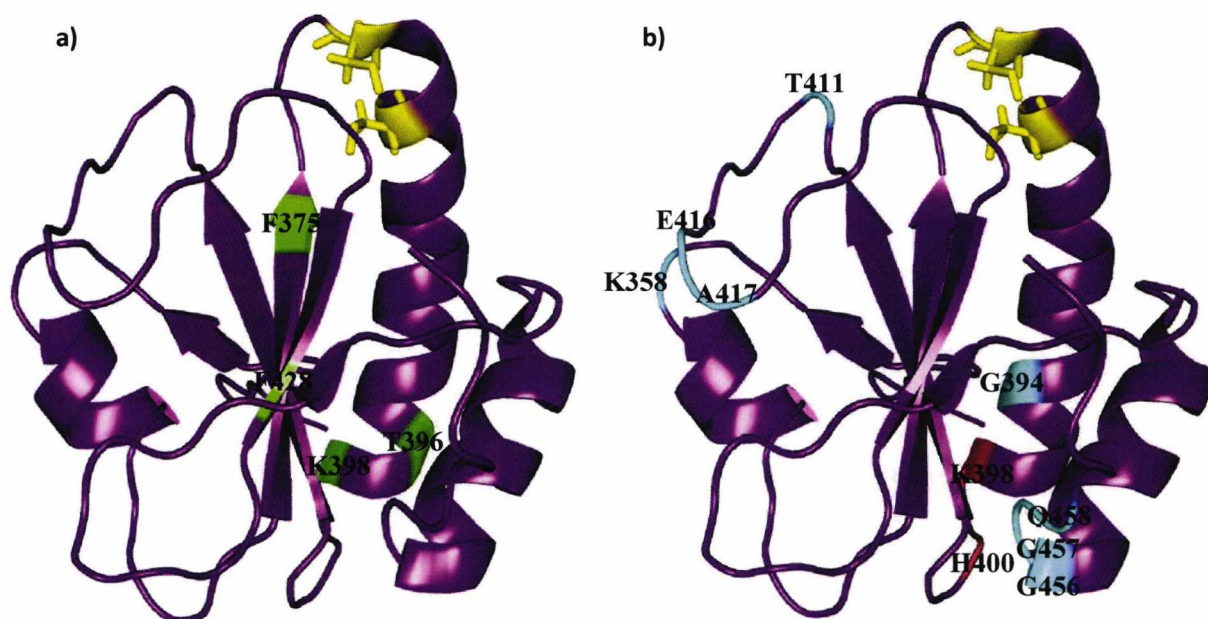


Figure 4.27 Model-free parameters displayed on the **a' structure.** a) residues with contribution from model-free parameters S_s^2 and S_f^2 are shown in green. b) residues experiencing R_{ex} and τ_e higher than 100 ps are shown in red and cyan, respectively. The cysteines of the active site are displayed as sticks in yellow.

As expected, residues with slow internal motion, above 100 ps, were also present in the **x** linker region, G332, I334, K335, L338 and Q350, as well as the **c** extension of **xa'c**, A461, G462, D463, A473, E474, K485, A486, V487, K488, D489 and E490, as the termini

b'x that is conformationally active and in exchange and raises the question: Is this event an “activation” in preparation for ligand binding?

WT **b'x**, regions experiencing conformational exchange also display internal motions above 100 ps, suggestive of slow conformational exchange. These residues are highlighted in cyan and are displayed on the I272A **b'x** structure in figure 4.29.

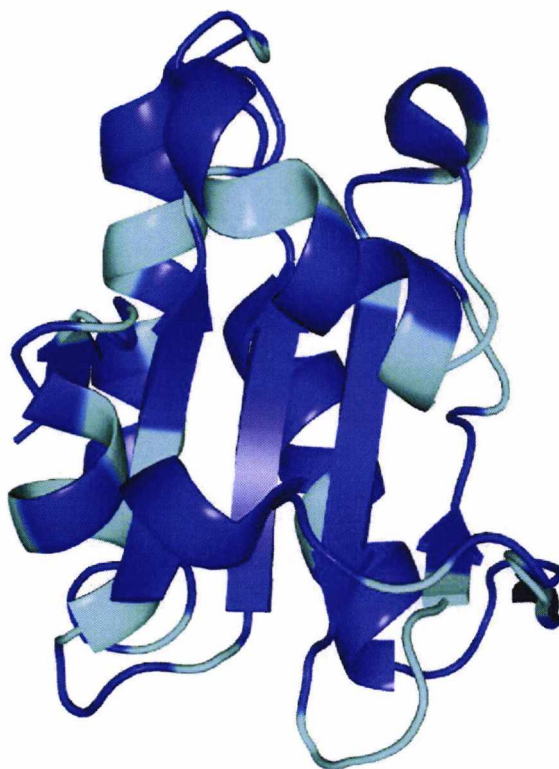


Figure 4.29 I272A **b'x** crystal structure (3BJ5.pdb) showing regions of slow internal motion. Residues in cyan displayed τ_e values higher than 100 ps for WT **b'x**.

WT **b'x** does not only display slow internal motions in loops between helix 4 and 5 as well as between strands 4 and 5, high τ_e values are also seen in helix 4, a region known to form part of the hydrophobic binding site, and helix 1 involved in ligand binding.

NMR relaxation data presented in this chapter assumes that all proteins under investigation are isotropic in solution. However, it is evident from the T_1 versus T_2 plots that these proteins fall on the theoretical contour line of $S^2 = 1$, with a number of residues falling below this line. In the case of I272A **b'x**, all of the residues fall below the $S^2 = 1$ line. Therefore the data indicated that **xa'c** and **b'x**, both WT and I272A, are anisotropic.

Modelling the data to an anisotropic case can be achieved using more advanced methods which were beyond the timescale available for this project when combined with protein production, generation of mutants and biophysical characterisation of the proteins discussed in this thesis (Palmer, 1997, Palmer, 2001). More in depth relaxation dynamics can also be achieved by ^{15}N and ^{13}C uniformly labelled proteins as ^{13}C can provide information that may have been missed by ^{15}N dynamics (Mittermaier and Kay, 2009).

However, as the NMR relaxation properties of full length human PDI or any of its domains have not been published in the literature, the purpose of this study was to characterise the **b'x** and **a'** domains separately in order to compare their behaviour in the **b'xa'c** protein.

CHAPTER 5

Characterisation of **b'xa'c** by NMR Spectroscopy

5.1 Introduction

As described in previous chapters, capping of the hydrophobic ligand binding site on the **b'** domain by the **x** linker region has been shown to act as a gate towards ligand binding in the **b'x** construct of hPDI (Nguyen et al., 2008). Therefore, one of the primary aims of this study is to determine if capping of the ligand binding site by the **x** linker occurs when the **a'** domain is also present in the **b'xa'c** C-terminal half of hPDI. If indeed capping of the ligand binding site does occur when **x** is tethered by both of its neighbouring domains, **b'** and **a'**, this could prove to be an important physiological mechanism controlling access to the hydrophobic binding site and could lead to wider structural arrangement of the protein as well as control other aspects of PDI structure/function.

Capping of the ligand binding site in the **b'x** fragment of PDI has been previously characterised by a combination of intrinsic fluorescence and NMR (Nguyen et al., 2008). As previously described in chapter 3, intrinsic fluorescence utilises the unique tryptophan residue, W347, in the **x** linker region to monitor capping by following the change in the environment of the Trp side chain upon binding to the hydrophobic ligand binding site of the **b'** domain. This change in the environment of W347 in **x**, consequently capping, can

also be monitored by NMR and has proven crucial in the determination of capping not only in the **b'x** construct, but also **bb'x** and now **b'xa'c** (Nguyen et al., 2008, Byrne et al., 2009). This chapter focuses on the characterisation of WT **b'xa'c** and mutants I272A, L343A and 2DA (D346A/D348A) as well as capping of the ligand binding site by the **x** linker region using NMR spectroscopy.

Characterisation of **b'xa'c** and its mutants has proved challenging, primarily due to the conformational stability and exchange of this protein. Denaturation studies revealed that the **a'** domain has very low conformational stability and presents with a high degree of conformational exchange in NMR data, both of which improved on addition of DTT. This chapter presents data on the effect of temperature, pH and the oxidation state of the **a'** domain on WT **b'xa'c** and the I272A, L343A and 2DA mutants by NMR. The effect of the peptide ligand Δ -somatostatin on WT **b'xa'c** and the I272A mutant was also investigated. However, in order to enable any kind of characterisation by NMR, assignment of the I272A **b'xa'c** mutant was carried out. In turn, this allowed mapping of these assignments onto $^{15}\text{N}/^1\text{H}$ HSQC spectra of WT **b'xa'c**, enabling backbone relaxation dynamic studies of WT **b'xa'c** and the I272A mutant to be completed.

5.1.1 NMR and conformational exchange

One advantage of NMR spectroscopy is its ability to probe timescales of motion through NMR relaxation but timescale effects can also have an effect on chemical shift. One protein molecular process that influences NMR observables is conformational exchange (chemical exchange). The effect of chemical exchange on a NMR spectrum depends on several factors, including the rate of exchange and the chemical shift difference between exchange environments and can be described below:

Consider a system that exchanges between two sites A and B with a rate k_{ex} . State A has a NMR chemical shift of δ_A and state B has a NMR chemical shift of δ_B and the chemical shift difference between the two states is $\Delta\delta_{AB}$. Depending on the magnitude of the rate, k_{AB} , the NMR spectrum will vary as shown in figure 5.1:

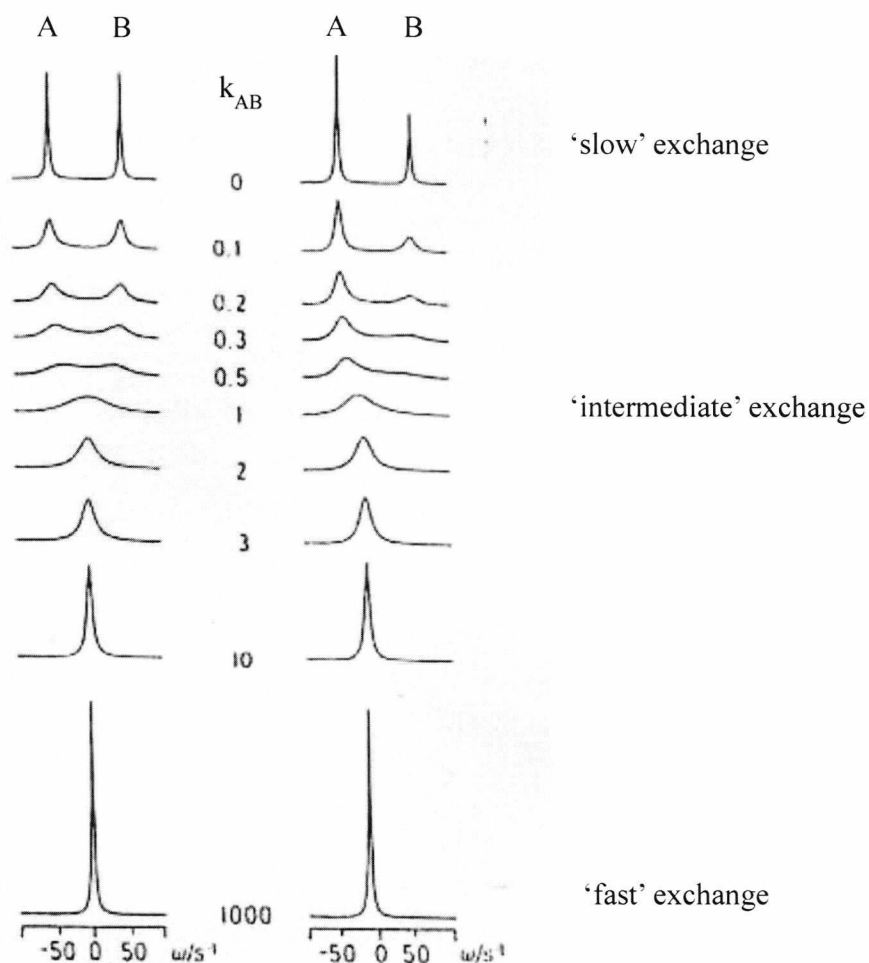


Figure 5.1. Two-site chemical exchange modelled using the equations of McConnell (McConnell, 1958). The left set are for when the proportions of A = B and right set for when the proportion of A > B.

When the rate is ‘slow’, we clearly see two states A and B. When the rate is ‘fast’, we see only an average state that is positioned according to the populations of A and B. When exchange events approach and pass through an intermediate condition; this is the point where we describe the system as experiencing conformational line broadening and this condition is seen throughout this thesis. To sharpen NMR lines you have to slow or speed up the exchange process and cooling or heating the sample can do this. The relationship when intermediate exchange exists is normally where $k_{AB} \sim \Delta\delta_{AB}$ and $\Delta\delta_{AB}$ is measured in Hz and so spectrometer dependent. This means exchange regimes can be moved by changing magnetic field and one method is to record data on 600, 700, 800 and

900 MHz spectrometers because 1 ppm of ^1H is 600, 700, 800 and 900 Hz respectively on these instruments.

5.1.2 Evidence of capping in PDI proteins

The first evidence of capping was confirmed by the WT **b'x** fragment of PDI, which presented as a mixture of capped monomer and uncapped dimer in solution (Nguyen et al., 2008). Generation of blue-shifted mutants like I272A and 2DA that promote capping of the ligand binding site by the **x** linker region, and the red-shifted L343A mutant that favours uncapping, allowed characterisation of this event by NMR spectroscopy. $^{15}\text{N}/^1\text{H}$ HSQC spectra for the capped monomeric species showed significant differences between the **b'x** mutants in the dispersion and resolution of resonances. I272A and 2DA **b'x** gave well resolved spectra consistent with a folded protein of the expected molecular size. Conversely, spectra of L343A appeared line broadened, indicative of conformational exchange on the NMR experimental timescale and suggested the possible presence of multimeric interactions. In the study conducted by Nguyen and co-workers (2008), as well as fluorescence which is described in chapter 3, NMR was used to track capping of the ligand binding site by the **x** linker by following the chemical shift of the W347 indole NH which shifts downfield upon capping. As **b'x** contains a single tryptophan residue, the $^{15}\text{N}/^1\text{H}$ HSQC spectrum of a homogenous protein sample should contain only one side chain peak at approximately 10.1 ppm (^1H) and 130.0 ppm (^{15}N), representing the indole NH of W347. However, two resonances were seen in this region indicating that at least the W347 side chain was present in two different environments, representing the two conformations of **x** in slow exchange, capped and uncapped seen in figure 5.2 (Nguyen et al., 2008).

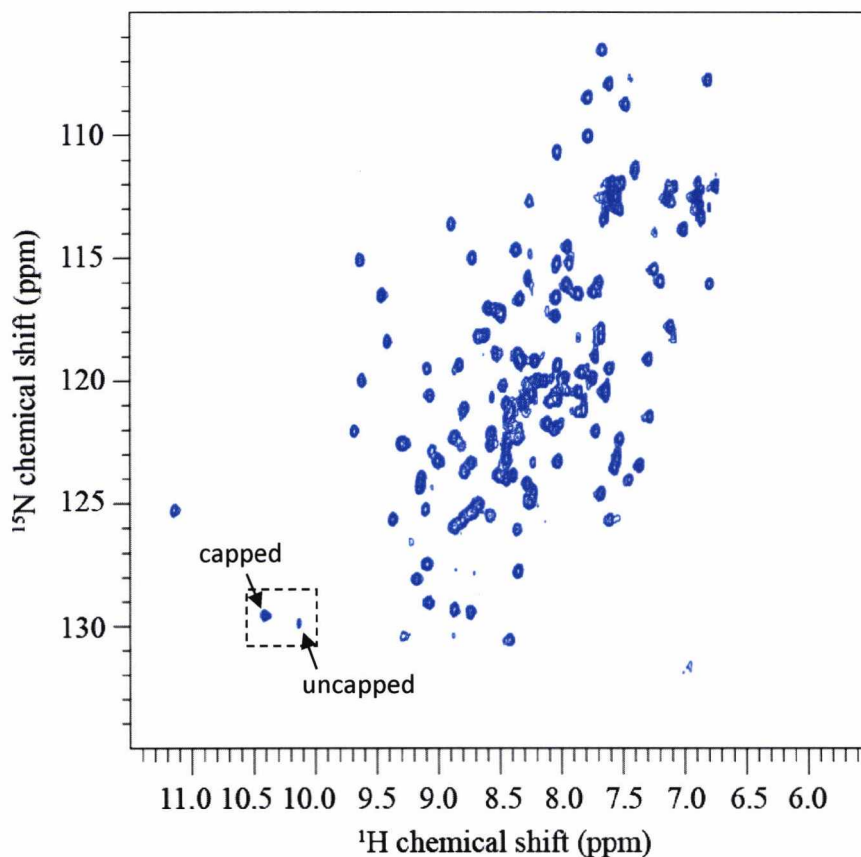


Figure 5.2 $^{15}\text{N}/^1\text{H}$ HSQC of WT **b'x**. WT monomeric **b'x** presents as a mixture of capped and uncapped protein as shown by the presence of two peaks for W347 in the boxed indole region.

Figure 5.3 shows a $^{15}\text{N}/^1\text{H}$ HSQC spectrum of full length hPDI focused on the indole region, overlaid with spectra of WT **b'x** and mutants I272A, L343A and 2DA. The I272A and 2DA mutants of the monomer form of **b'x** favour the capped conformation due to the shift of the W347 resonance at lower field in the spectrum, whereas L343A favours the uncapping of the ligand binding site by the **x** linker region. Monitoring the chemical shift of the W347 indole in combination with intrinsic fluorescence forms the basis of tracking the different conformations of the **x** linker region.

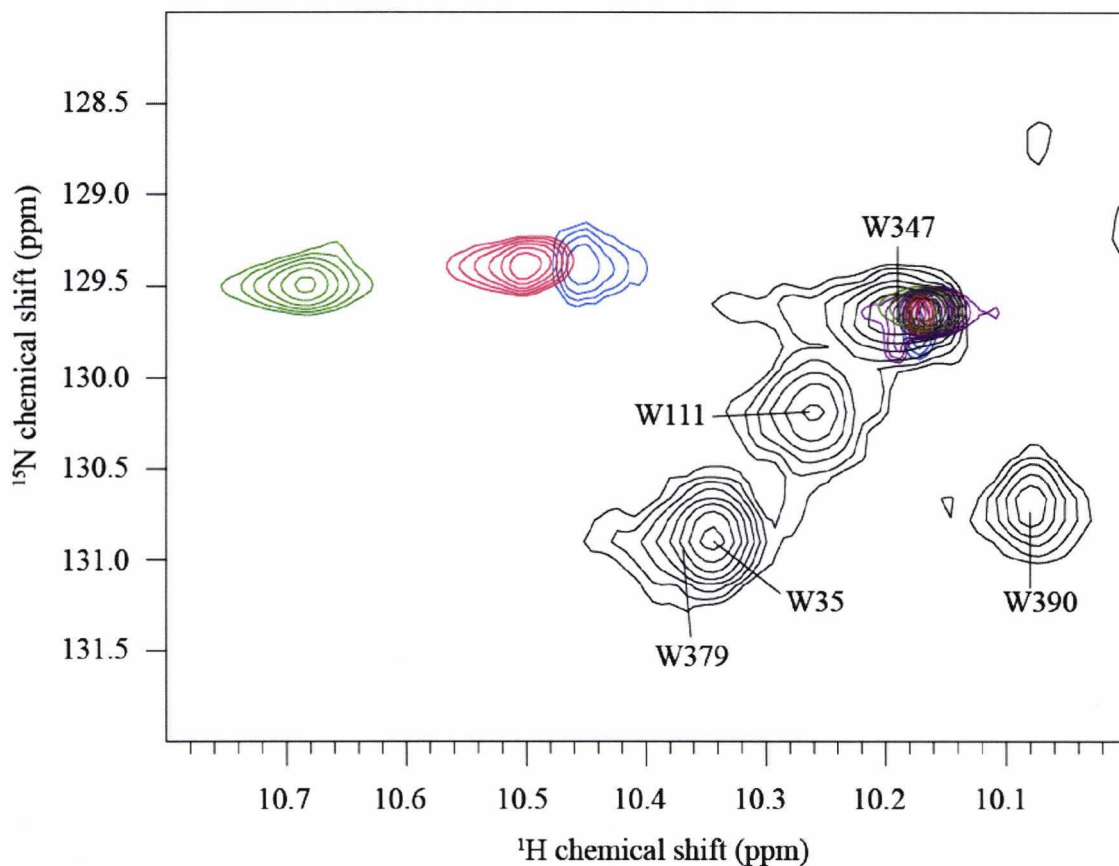


Figure 5.3 Conformational exchange in the environment of W347 of b'x. Overlay of the indole region of $^{15}\text{N}/^1\text{H}$ HSQC spectra of full length hPDI (black), WT **b'x** (blue) **b'x** mutants I272A (red), L343A (purple) and 2DA (green). Assignments of full length hPDI resonances are indicated.

5.1.3 Sequential backbone assignment of reduced **b'xa'c**

Backbone assignments were essential for the characterisation of **b'xa'c** by NMR. However, due to the large size of the WT protein and a high degree of line broadening, triple resonance assignments were challenging. In contrast to WT **b'xa'c**, $^{15}\text{N}/^1\text{H}$ HSQC spectra for the I272A **b'xa'c** mutant presented with resonances that were less broadened and better resolved, making this mutant a more ideal target for NMR triple resonance assignments.

Triple resonance experiments for the I272A mutant of **b'xa'c** required $^{15}\text{N}/^{13}\text{C}$ labelled protein to enable the assignment of $^1\text{H}_\text{N}$, ^{15}N and $^{13}\text{C}_\alpha$. The triple resonance pair of experiments used for the backbone assignment of **xa'c** in chapter 4, CBCA(CO)NH and CBCANH, could not be used for the assignment of **b'xa'c** as they were of poor quality due

to the larger size of the protein and the presence of the unstructured C-terminal extension, **c**, therefore C_{β} nuclei could not be assigned. Instead, the HN(CO)CA and HNCA pair of triple resonance experiments was used, in conjunction with previous assignments from WT **b'x** and **xa'c**. HNCA correlates the ^1H and ^{15}N with the $^{13}\text{C}_{\alpha}$ nuclei of the same (designated “i”) and preceding amino acid (designated “i-1”), whereas the HN(CO)CA correlates the ^1H and ^{15}N with only the $^{13}\text{C}_{\alpha}$ of the preceding amino acid. This approach utilises the distinctive C_{α} chemical shifts characteristic of the amino acid type as described in 4.1.2 for the CBCA(CO)NH and CBCANH. However, as there is a lack of information for the C_{β} nuclei from this pair of experiments, the already assigned spectra of WT **b'x** and **xa'c**, shown in figures 4.14 and 4.13 (chapter 4) were used to confirm the assignments that could be mapped onto I272A **b'xa'c**.

The data from the triple resonance experiments was loaded into Analysis Version 2 program which, in combination with assigned **b'x** and **xa'c** spectra, allowed the sequential assignment of $^1\text{H}_{\text{N}}$, ^{15}N and $^{13}\text{C}_{\alpha}$ nuclei of I272A **b'xa'c**.

5.1.4 The redox-dependent conformational changes of **bb'xa'** and **abb'xa'**

Recent full length PDI studies have shown the N-terminal **ab** region as less susceptible to proteolytic cleavage than the C-terminal region and is considered proof of a difference in flexibility between both halves of the protein, and most importantly, **b'xa'c** undergoes significant structural changes upon reduction of the **a'** active site (Wang et al., 2010, Wang et al., 2012b). Recent structures of **bb'xa'** and **abb'xa'**, the latter both in the reduced and the oxidised state, have shown that the conformational changes in PDI are triggered by the redox state switch of the active site in the **a'** domain (Wang et al., 2012b, Wang et al., 2012a). In addition to a catalytic domain, the **b'** domain and the **x** linker region are also required and are essential for the conformational change of hPDI. As a result, the **b'xa'c** fragment determines the redox-dependent conformational changes of hPDI.

As mentioned in the introduction of this thesis, the recently solved reduced and oxidised **abb'xa'** structures confirm the conformational flexibility of hPDI as the protein assumes a more compact conformation upon reduction of the **a'** active site when compared to the oxidised structure, where the **a'** domain moves out by $\sim 45^{\circ}$ releasing the compact nature of the molecule. ANS fluorescence studies showed that the reduced compact

structure of hPDI revealed a smaller hydrophobic area than the oxidised structure (Wang et al., 2012a). In addition, reduced PDI shows a blue-shift in fluorescence when compared to the oxidised form, consistent with the compact conformation. This blue shift in fluorescence has also been observed for **b'xa'c** in this project as described in chapter 3, confirming that redox-regulated conformational changes in PDI are mainly located in the C-terminal half of the protein. This redox dependent conformational change has allowed a glimpse into the mechanism of action of PDI as it has been shown to regulate not only the chaperone activity of PDI but also its ability to recognise different substrates depending on its redox state.

To date, there are no publications on the redox-dependent conformational changes of full-length hPDI or the **b'xa'c** fragment by NMR spectroscopy. However, Serve and co-workers (2010) showed chemical shift changes in $^{15}\text{N}/^1\text{H}$ HSQC spectra of the C-terminal half of PDI, from the thermophilic fungus *Humicola insolens*, in the reduced and oxidised states. Mapping these chemical shift changes onto their structural model of **b'xa'c** revealed that the redox state of the **a'** domain influenced the dynamic properties of the **b'** domain (Serve et al., 2010).

Therefore, this chapter focuses on the characterisation of the **b'xa'c** fragment of hPDI, its dynamic properties, ligand binding ability and the effect of the reducing agent DTT, temperature and pH on this protein.

5.2 Materials and Methods

5.2.1 NMR sample preparation

Samples for NMR analysis were prepared as described in chapter 4. I272A **b'xa'c** protein required for triple resonance experiments, was ^{13}C and ^{15}N labelled. Proteins for all other NMR experiments were only ^{15}N labelled. The monomer species were concentrated using Vivaspin 4 concentrators (MWCO 10 kDa) to 0.4 mM or 1.0 mM final sample concentration as determined by absorbance at $A_{280\text{nm}}$. Also, proteins were buffer exchanged into the required buffers containing 10% (v/v) D_2O (Goss Scientific Ltd.).

5.2.2 NMR Data acquisition and processing

NMR data was acquired with the assistance of Dr. Mark Howard and Dr. Michelle Rowe. pH, temperature, ligand binding and triple resonance experiments were completed on a 600 MHz Varian UnityINOVA spectrometer operating at 14.1 Tesla (^1H resonance frequency of 600 MHz) equipped with a 5 mm HCN z-pulse field gradient probe. Relaxation dynamics experiments for WT and I272A **b'xa'c** were run on a 600 MHz AV3 Bruker spectrometer with a QCI-F cryoprobe. Data acquisition and processing was completed as described in chapter 4, section 4.2.2.

5.2.2.1 Amino acid sequential backbone resonance assignment

Backbone resonance assignments were obtained from a 1.0 mM reduced $^{13}\text{C}/^{15}\text{N}$ labelled I272A **b'xa'c** sample in a pH 7.0 buffer consisting of 20 mM MOPS, 50 mM NaCl and 10 mM DTT. HNCA and HN(CO)CA triple resonance data sets were acquired at 40°C with 1024 points (9000 Hz) in the direct F3 dimension (^1H), 20 points (2100 Hz) in F2 (^{13}C) and 58 points (4564 Hz) in F1 (^{15}N) indirect dimensions. Carrier frequencies for the triple resonance experiments were set to 4.588 ppm, 55.704 ppm and 119.284 ppm for the ^1H , ^{13}C and ^{15}N respectively.

5.2.2.2 ^{15}N NMR Relaxation

^{15}N relaxation experiments were carried out for WT and I272A **b'xa'c**. T_1 and T_2 experiments were run with 2048 (9000 Hz) in the direct F2 dimension and 256 points

(2100 Hz) in the indirect F1 dimension. Experiments were collected at 40°C on the AV3 Bruker spectrometer at 600 MHz with the carrier frequencies set at 4.766 ppm, 59.143 ppm and 119.08 ppm for ^1H , ^{13}C and ^{15}N respectively. T_1 relaxation delays were set to 256, 384, 512, 640 (x2), 768, 894 and 1024 ms. T_2 relaxation delays were set to 33.92, 50.88, 67.84, 84.80 (x2), 101.76, 118.72 and 135.68 ms.

As with **b'x** and **xa'c** relaxation, ^{15}N HetNOE data were obtained by observing the intensity of the NH peaks with and without saturation of amide protons. HetNOE experiments were run with the same spectral widths as the T_1 and T_2 experiments.

5.2.3 NMR Data analysis

5.2.3.1 Triple Resonance experiments HNCA and HN(CO)CA

The completed backbone assignments for **xa'c** and WT **b'x** were used alongside triple resonance experiments to assign I272A **b'xa'c** by assignment mapping and sequential assignments of the backbone C_α . Data analysis was carried out in CcpNmr Analysis Version 2 (Vranken et al., 2005) as described in chapter 4.

5.2.3.2 NMR temperature titrations

Temperature titrations were carried out on 1.0mM reduced protein samples of WT **b'xa'c** and I272A, L343A and 2DA mutants. All proteins were in a pH 6.5 buffer containing 20 mM NaH_2PO_4 , 50 mM NaCl, 10 mM DTT and 10% (v/v) D_2O . $^{15}\text{N}/^1\text{H}$ HSQC experiments were run as described in 4.2.2.1 at 25°C, 15°C, 35°C and 40°C.

5.2.3.3 NMR pH titrations

pH titrations were carried out on 1.0 mM reduced protein samples of WT **b'xa'c** and I272A, L343A and 2DA mutants. $^{15}\text{N}/^1\text{H}$ HSQC experiments were run at 40°C as described in 4.2.2.1. Proteins were buffer exchanged into four different buffers at pH 6.0, 6.5, 7.0 and 8.0 as described in 4.2.4.

5.2.3.4 *Ligand binding using Δ -somatostatin*

Ligand binding experiments were performed with 0.4 mM protein, WT and I272Ab'**xa'c**, in the presence of varying concentrations of Δ -somatostatin. $^{15}\text{N}/^1\text{H}$ HSQC spectra were recorded at 40°C with the protein and ligand in a pH 7.0 buffer containing 20 mM MOPS, 50 mM NaCl, 4mM DTT and 10% (v/v) D_2O . Four samples were made up for each protein containing 0.4 mM **b'xa'c** and the peptide ligand was added to the protein at concentrations of 0.08 mM, 0.2 mM and 0.4 mM. A control sample was also used containing protein in the absence of peptide ligand. $^{15}\text{N}/^1\text{H}$ HSQC spectra were recorded for each sample and the differences in chemical shifts were calculated using equation 4.5 as described in section 4.2.3.2. The chemical shifts from the control spectrum containing no peptide were used as a reference for calculating the chemical shift that occurred due to the addition of peptide ligand.

5.2.3.5 *Secondary structure prediction using DANGLE*

Secondary structure prediction for **b'xa'c**, WT and I272A, was completed using the DANGLE (Dihedral Angles from Global Likelihood Estimates) program built into CcpNmr Analysis Version 2. ^{15}N , ^1H , $^{13}\text{C}_\alpha$ chemical shifts obtained from the triple resonance assignment of I272A **b'xa'c** were used for the structure prediction of this mutant. The ^1H and ^{15}N chemical shifts from the I272A triple resonance assignment were mapped on WT **b'xa'c** spectra and used to generate a structure prediction in DANGLE. Analysis was carried out as described in section 4.2.3.3.

5.2.3.6 *NMR Relaxation Analysis*

Analysis of ^{15}N T_1 , T_2 and hetNOE backbone relaxation and model-free data was carried out as described in section 4.2.3.4.

5.2.3.7 *Minimal Shift Mapping*

Minimal maps were constructed using assigned spectra of I272A **b'xa'c** against unassigned spectra of WT, L343A and 2DA **b'xa'c**. Equation 4.5 was used to estimate the chemical shift distances from the signals in the assigned I272A spectrum to the nearest peak of the unassigned WT **b'xa'c** spectrum. This was repeated with L343A and 2DA.

5.3 Results

5.3.1 Initial NMR investigations on WT **b'xa'c** and mutants

Initial NMR studies of WT **b'xa'c** revealed $^{15}\text{N}/^1\text{H}$ HSQC spectra with poorly resolved and line broadened peaks, indicative of conformational exchange. Similarly to the **a'c** construct discussed in chapter 4, reduction of the **a'** active site upon addition of DTT, improved spectral quality and reduced line broadening. Figure 5.4 shows an overlay of WT **b'xa'c** in the presence and absence of DTT. Both spectra were collected at 25°C with the protein in a pH 6.5 buffer.

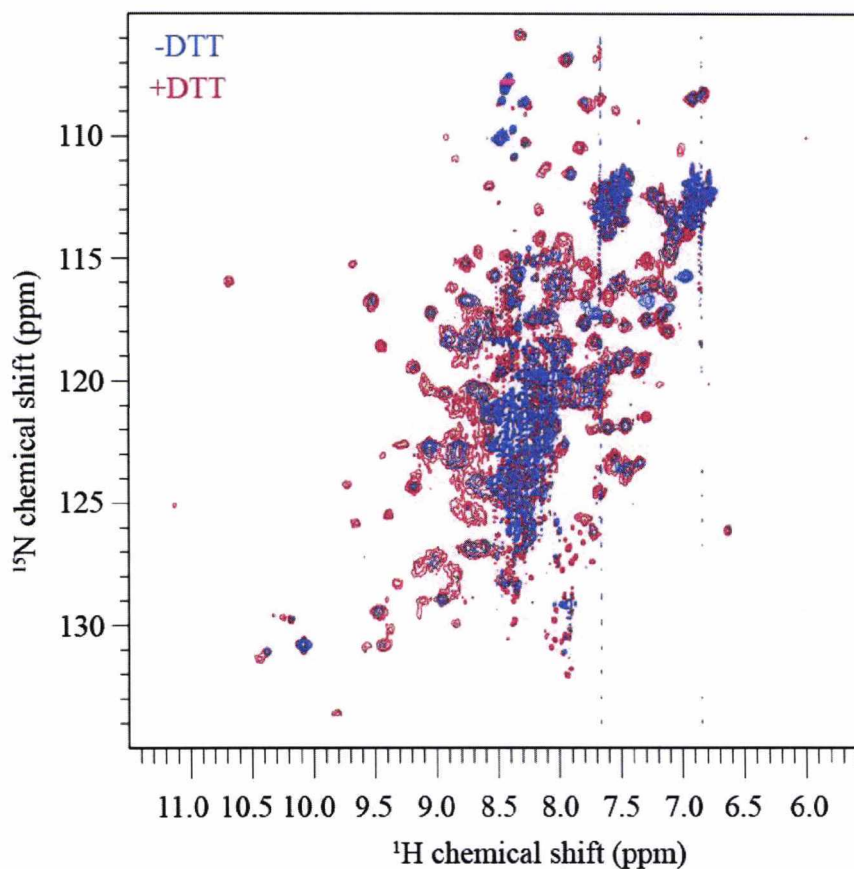


Figure 5.4 $^{15}\text{N}/^1\text{H}$ HSQC spectra of WT **b'xa'c** at 25°C, pH 6.5. Peak dispersion and resolution is improved on addition of DTT (red) compared to the spectrum in the absence of DTT (blue).

As addition of DTT resulted in reduced line broadening and better resolved peaks, all subsequent NMR experiments were carried out in the presence of DTT. This supports the chemical denaturation data which showed an increase in the conformational stability of

the **a'** domain in the reduced state. Although reduction of the **a'** active site resulted in improved peak dispersion, resonances were still broad and the quality of the spectra was insufficient for further NMR studies. Sharp resonances, due to the unstructured C-terminal extension, **c**, were present between 8.0 and 8.5 ppm and added to the poor quality of the spectrum. Spectra of I272A, L343A and 2DA mutants also presented with poorly resolved peaks at 25°C and pH 6.5 as shown in figures 5.5, 5.6 and 5.7 respectively.

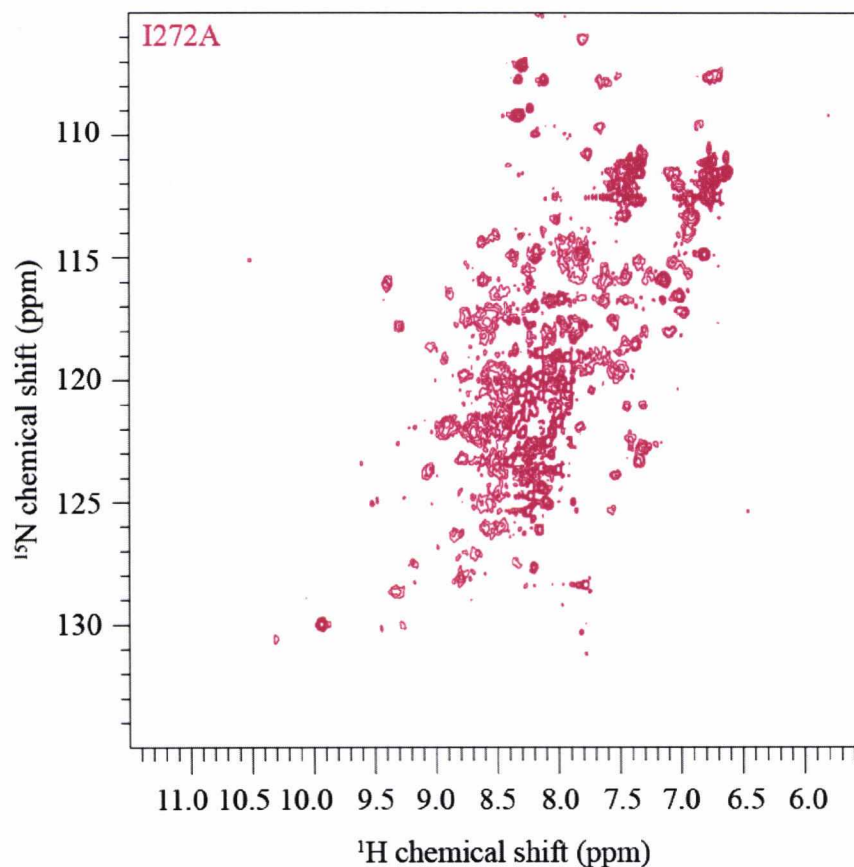


Figure 5.5 $^{15}\text{N}/^1\text{H}$ HSQC spectrum of the I272A **b'xa'c** mutant. Sample contained 1.0 mM protein and 10 mM DTT in a pH 6.5 buffer containing 20 mM NaH_2PO_4 , 50 mM NaCl and 10% (v/v) D_2O . Spectrum was collected at 25°C on the 600 MHz Varian UnityINOVA spectrometer.

The I272A **b'xa'c** mutant spectrum was very similar to that of the WT protein in figure 5.4 showing a high degree of line broadening. Although this mutant in **b'x** favours capping of the ligand binding site that results in better resolved NMR spectra than the WT protein, the I272A mutation causes only a marginal improvement in **b'xa'c** at 25°C and with the proteins at pH 6.5.

$^{15}\text{N}/^1\text{H}$ HSQC spectrum of the L343A **b'xa'c** mutant presented with peaks that were even more line broadened and less well resolved than WT and I272A **b'xa'c**, as can be seen in figure 5.6.

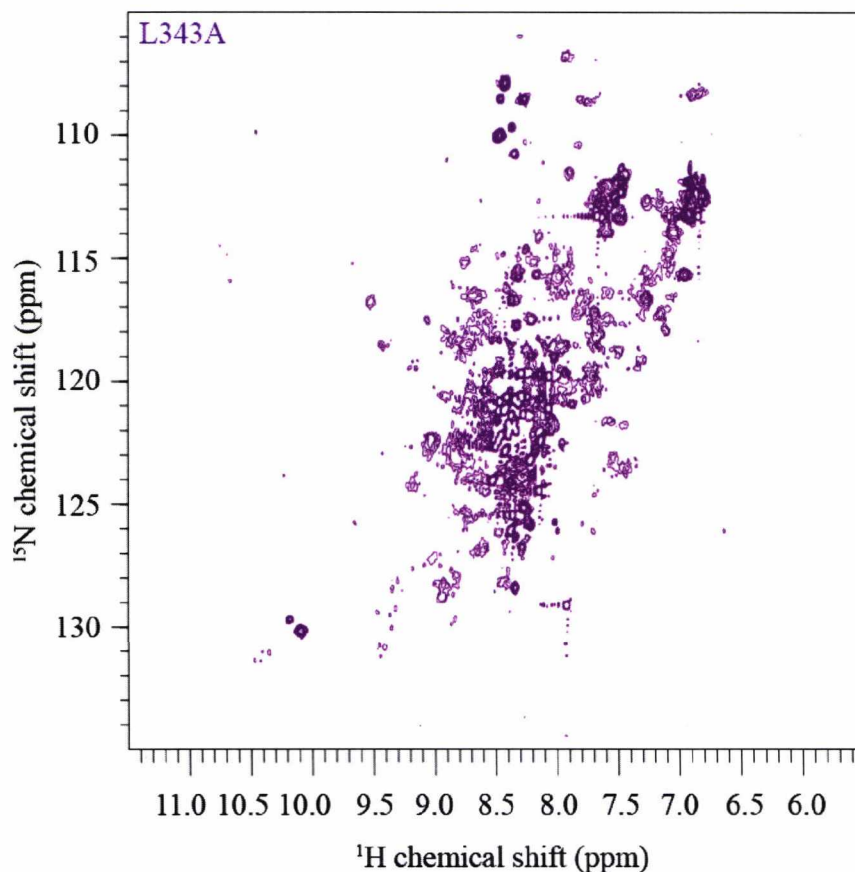


Figure 5.6 $^{15}\text{N}/^1\text{H}$ HSQC spectrum of the L343A **b'xa'c** mutant. The L343A protein sample was made up in the same buffer conditions as for WT and I272A **b'xa'c**. Spectrum was acquired at 25°C on the 600 MHz Varian UnityINOVA spectrometer.

In **b'x**, the L343A mutant favours uncapping of the ligand binding site by the **x** linker region resulting in poor quality NMR spectra due to increased line broadening. The L343A mutant of **b'xa'c**, shown in figure 5.6, presented with peaks that were even more line broadened and less well resolved than WT and I272A **b'xa'c** and therefore supports the previously published **b'x** data (Nguyen et al., 2008).

Like I272A, the 2DA mutant of **b'x** also favours capping of the ligand binding site by the **x** linker region and generated good quality NMR spectra with well resolved peaks (Nguyen et al., 2008). However, the 2DA mutant of **b'xa'c** does not appear to have the

same effect as in **b'x**. On the contrary, 2DA **b'xa'c** $^{15}\text{N}/^1\text{H}$ HSQC spectrum resembles those of **a'c** shown in chapter 4, where a number of peaks are line broadened beyond detection due to the high degree of conformational exchange. This is supported by the chemical denaturation data in chapter 3, where this mutant had the lowest conformational stability out of all the **b'xa'c** proteins.

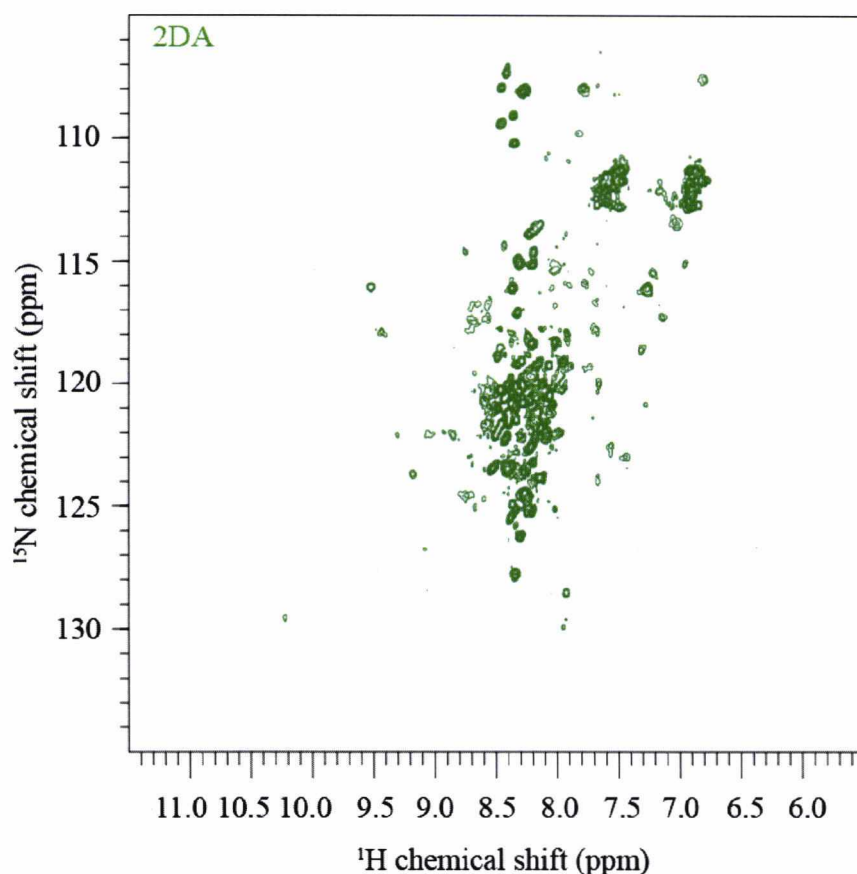


Figure 5.7 $^{15}\text{N}/^1\text{H}$ HSQC spectrum of the 2DA **b'xa'c** mutant. Protein sample and experimental conditions were as described for WT, I272A and L343A **b'xa'c**.

Visual inspection of the spectra revealed that the 2DA mutant presented with a higher degree of line broadening than WT **b'xa'c** and I272A and L343A mutants, demonstrated by the absence of peaks around the periphery of the spectrum in figure 5.7. L343A **b'xa'c** also showed diminished peak dispersion and increased line broadening compared to the WT protein. The I272A mutant was the only mutant that appeared to show a marginal improvement in peak resolution, but the quality of the NMR spectra was still insufficient for further NMR characterisation.

As WT **b'xa'c** and its mutants presented with $^{15}\text{N}/^1\text{H}$ HSQC spectra that were not of adequate quality for NMR investigations, it was necessary to find suitable experimental conditions to improve the quality of NMR spectra in order to allow backbone assignments and further NMR studies. Therefore, the effects of temperature and pH on WT and mutant **b'xa'c** proteins were investigated.

5.3.2 The effect of temperature and pH on WT **b'xa'c** and mutants

$^{15}\text{N}/^1\text{H}$ HSQC experiments were run at a range of temperatures with **b'xa'c** proteins in buffers of varying pH to find suitable conditions which lowered the degree of line broadening and as a consequence improved the quality of the NMR spectra to allow NMR characterisation. All $^{15}\text{N}/^1\text{H}$ HSQC spectra for WT and mutant **b'xa'c** proteins under different conditions are shown in Appendix 5.1.

Initially, NMR spectra of the proteins were run over a variety of temperatures in the order 25°C, 15°C, 35°C and 40°C with the protein samples at pH 6.5. It was evident by visual inspection of the spectra that an increase in temperature resulted in an increase in peak resolution and dispersion. An increase in temperature causes a shift from intermediate to fast exchange (figure 5.1) resulting in reduced line broadening as an average peak is seen between the states in exchange. Figure 5.8 shows $^{15}\text{N}/^1\text{H}$ HSQC spectra of WT **b'xa'c** and mutants I272A, L343A and 2DA acquired at 15°C and 40°C with the proteins at pH 6.5 in order to demonstrate the positive effect of increasing temperature on lowering the degree of line broadening in **b'xa'c**.

A similar effect was observed with an increase in pH, as spectra collected from protein samples at pH 6.0 at 40°C displayed line broadening and lack of peak dispersion. Conversely, a change into pH 8.0 lowered the degree of line broadening, and in turn increased peak resolution and dispersion, as can be seen in figure 5.9.

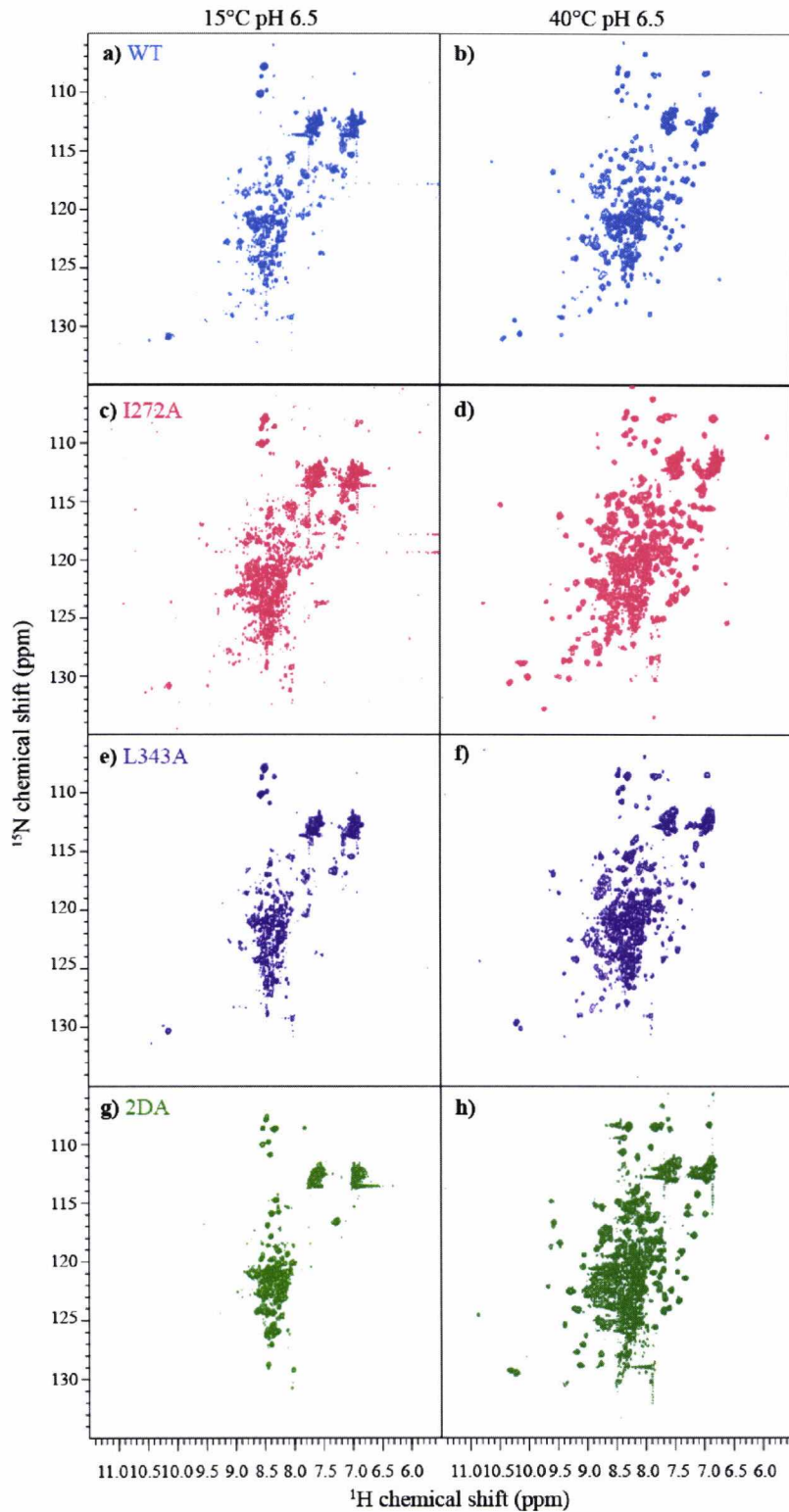


Figure 5.8 $^{15}\text{N}/^1\text{H}$ HSQC spectra of WT b'xa'c and mutants at pH 6.5. a) and b) WT at 15°C and 40°C c) and d) I272A at 15°C and 40°C e) and f) L343A at 15°C and 40°C g) and h) 2DA at 15°C and 40°C, respectively.

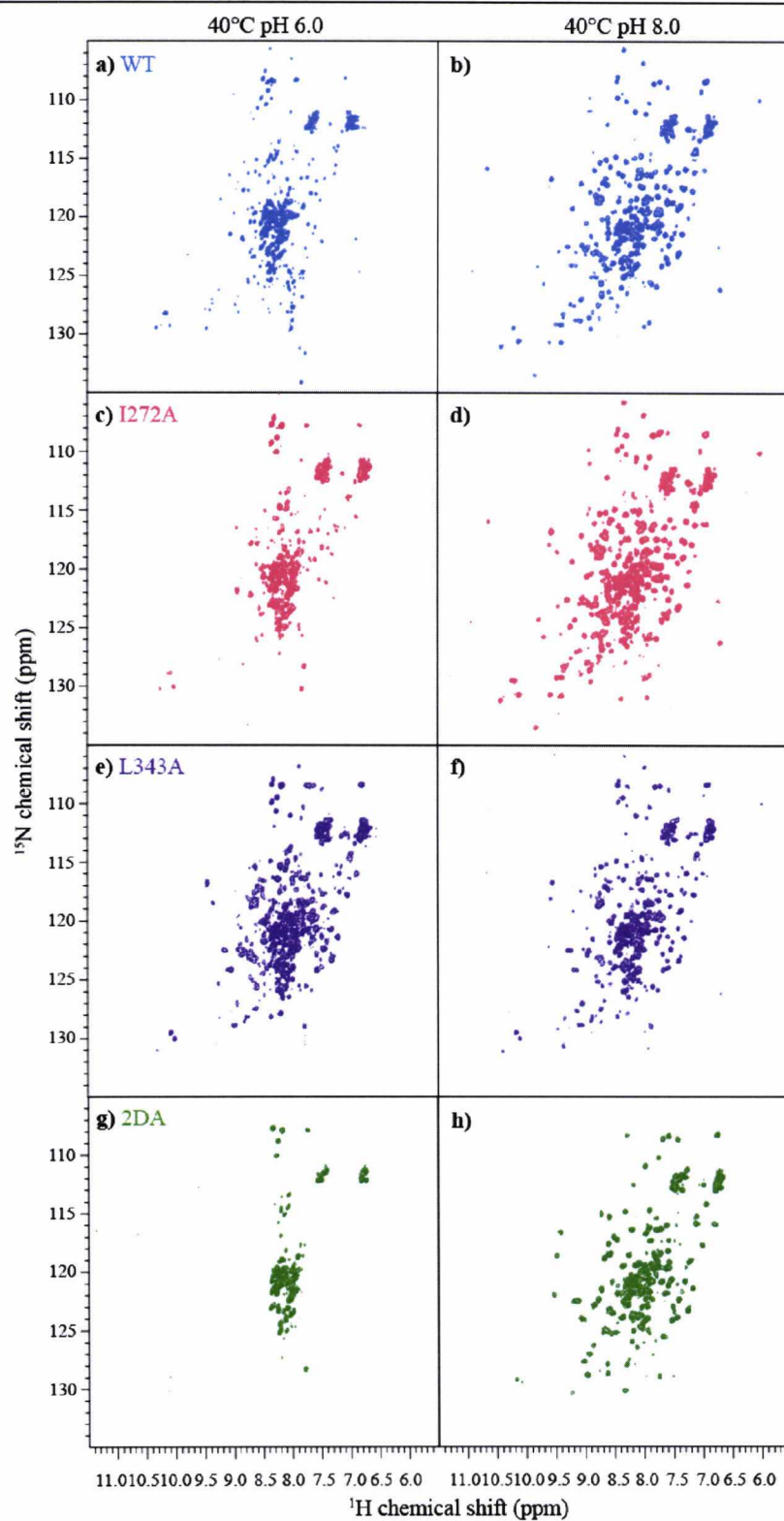


Figure 5.9 $^{15}\text{N}/^1\text{H}$ HSQC spectra of WT b'xa'c and mutants at 40°C. a) and b) WT at pH 6.0 and pH 8.0 c) and d) I272A at pH 6.0 and pH 8.0 e) and f) L343A at pH 6.0 and pH 8.0 g) and h) 2DA at pH 6.0 and pH 8.0, respectively.

As can be seen from figures 5.8 and 5.9, an increase in temperature and pH dramatically increased the quality of the NMR spectra with the best spectra acquired at 40°C and pH 8.0. The I272A mutant produced the best quality spectra out of all the **b'xa'c** proteins, followed by WT, L343A and finally 2DA. Both L343A and 2DA showed an improvement as temperature and pH increased, but the quality of the spectra was still insufficient for NMR studies. As the I272A spectra showed better resolved and dispersed peaks, its peak heights at the different temperature and pH conditions are shown in figure 5.10. The data used for the temperature and pH graphs is shown in appendices 5.2 and 5.3 respectively. The graphs were constructed using the triple resonance assignments obtained for I272A **b'xa'c**, which will be discussed in section 5.3.4.

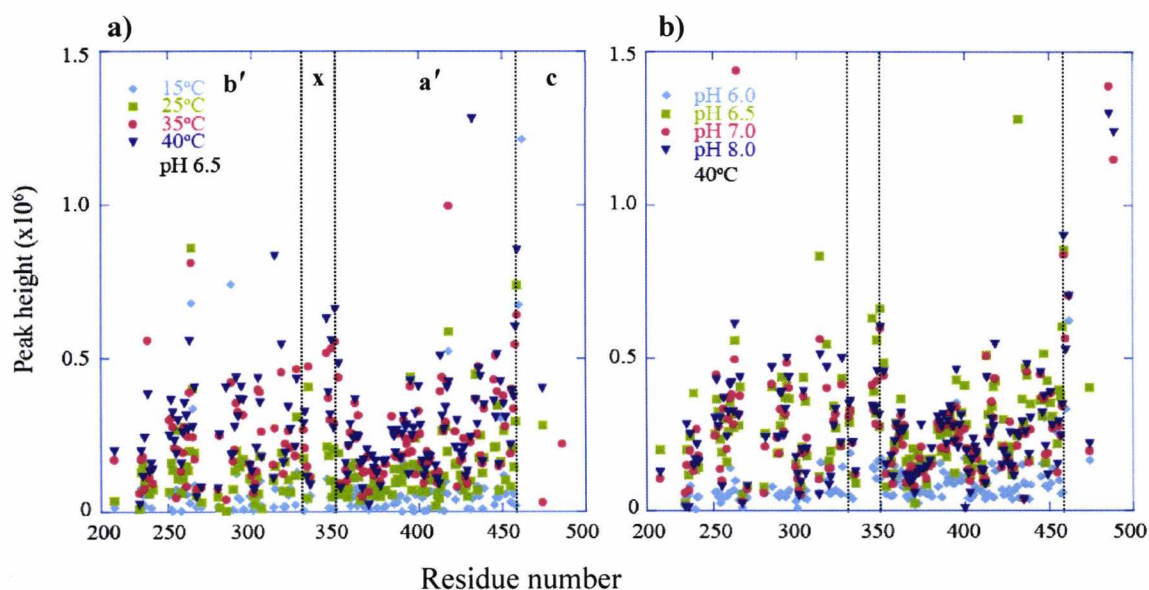


Figure 5.10 The effect of temperature and pH on the peak height of reduced I272A **b'xa'c** $^{15}\text{N}/^1\text{H}$ HSQC spectra. a) peak heights increase with temperature and b) increasing pH. The horizontal dotted lines show the domain boundaries.

The highest peak heights were observed at 40°C, therefore subsequent experiments were run at this temperature. The lowest peak heights were observed at pH 6.0 followed by an increase at pH 6.5. However, there was no significant increase in peak height from pH 7.0 to pH 8.0 and as some degradation was seen at the higher pH, therefore, it was decided for all future experiments to be run with the protein in a pH 7.0 buffer.

Surprisingly, I272A **b'xa'c** assignments obtained from triple resonance experiments, discussed in section 5.3.4, could be later mapped onto the WT **b'xa'c** NMR

spectra run at 40°C and pH 7.0 and pH 8.0 but not at the lower temperatures or pH conditions due to missing peaks caused by excessive line broadening. WT **b'xa'c** peak heights from spectra acquired at 40°C and pH 7.0 and pH 8.0 were very similar to those obtained for the I272A mutant and therefore are not shown here.

I272A **b'xa'c** assignments could not be mapped on L343A or 2DA $^{15}\text{N}/^1\text{H}$ HSQC NMR spectra due to large scale spectrum differences. Therefore, the analysis of temperature and pH spectra for these mutants was limited. However, it is evident from figures 5.8 and 5.9 that the L343A and 2DA mutations have had a negative effect on the conformational rigidity of the **b'xa'c** protein over the NMR timescale of detection. This supports the denaturation data in chapter 3, which shows that these mutants have a lower conformational stability than WT and I272A **b'xa'c**, especially in the **a'** domain.

5.3.3 Assignment of I272A **b'xa'c** and mapping assignments on WT **b'xa'c**

Sequential backbone assignments for $^1\text{H}_\text{N}$, $^{15}\text{N}_\text{H}$ and $^{13}\text{C}_\alpha$ nuclei for the I272A **b'xa'c** mutant, with the **a'** active site in the reduced state, were made from the pair of triple resonance experiments HNCA and HN(CO)CA. Previously assigned $^{15}\text{N}/^1\text{H}$ HSQC spectra of **xa'c** and WT **b'x** were used in combination with the triple resonance experiments to aid in the assignment of I272A **b'xa'c**. Ignoring the His tag, backbone assignments for I272A **b'xa'c** were obtained for 66% of amide $^1\text{H}_\text{N}$ and ^{15}N and 70% $^{13}\text{C}_\alpha$. The indole H_ϵ and N_ϵ for W347, W379 and W390 were also assigned using an overlay of full length PDI for which the indole region had been previously assigned. Residues were linked sequentially as the example shown in figure 5.11 for the amino acid stretch Ile398 to Leu393.

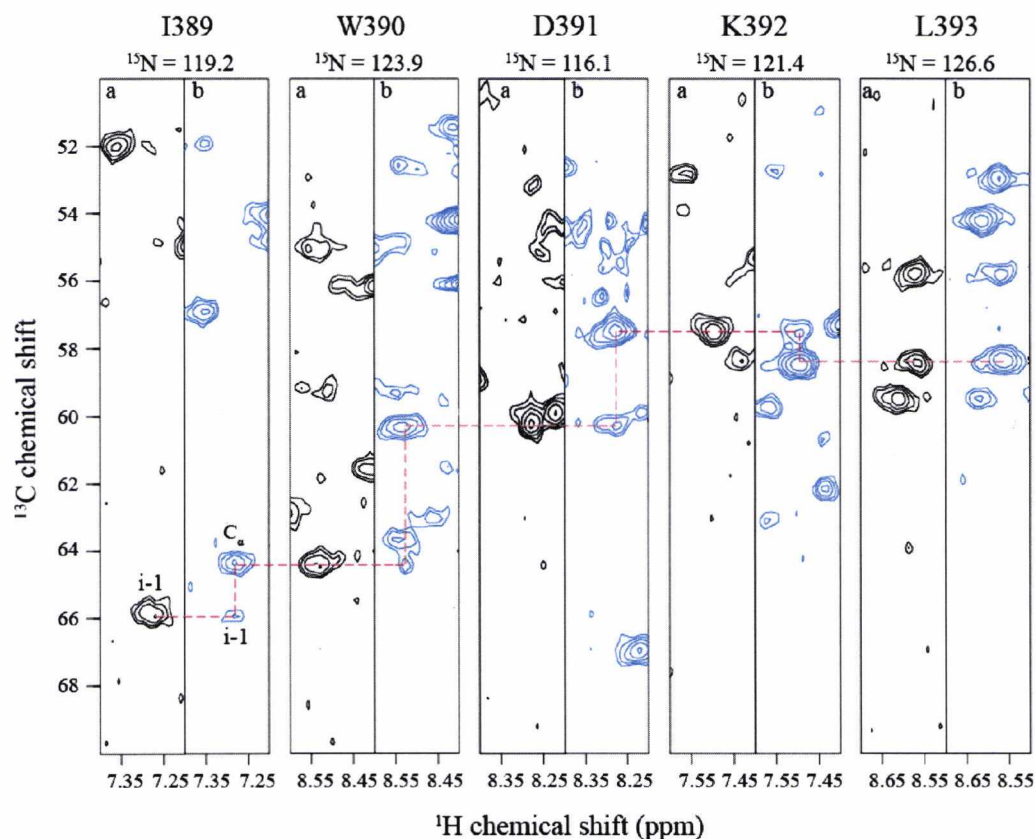


Figure 5.11 Backbone triple resonance sequential assignment example from I272A **b'xa'c experiments.** HN(CO)CA and HNCA strips, **a** and **b** respectively, show the assignment stretch for residues I389 to L393. The red dotted line shows the walkthrough along the sequential matches. The relevant ^{15}N chemical shift position of each plane is shown above the strip.

Residue-labelled $^{15}\text{N}/^1\text{H}$ HSQC spectrum of reduced I272A **b'xa'c** is shown in figure 5.12. Assignments were obtained as part of the triple resonance sequential assignments shown above and the $^1\text{H}_\text{N}$, ^{15}N and $^{13}\text{C}_\alpha$ chemical shifts are summarised in Appendix 5.4.

mutant onto $^1\text{H}/^{15}\text{N}$ HSQC spectrum of WT **b'xa'c** run at 40°C , with the sample in a pH 7.0 buffer.

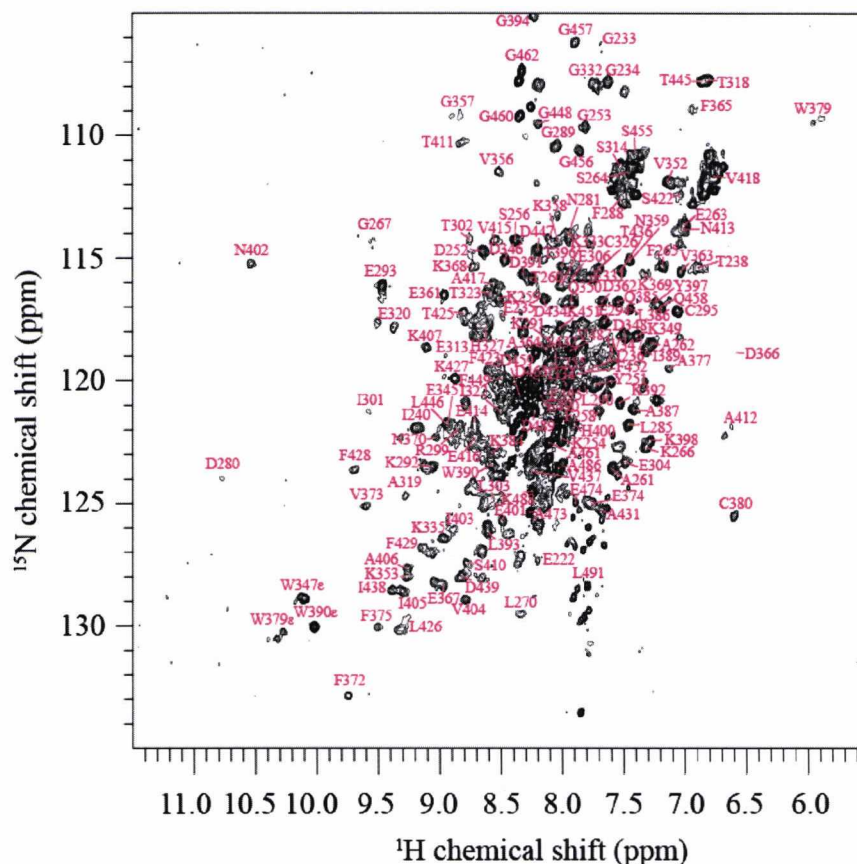


Figure 5.13 $^{15}\text{N}/^1\text{H}$ HSQC spectrum for reduced WT **b'xa'c**. Assignments shown in red are those mapped from an overlay of the I272A **b'xa'c** spectrum in figure 5.12. Experiments were run at 40°C with the protein sample in a pH 7.0 buffer.

5.3.4 The effect of the mutations on **b'xa'c**

In order to observe the effect of mutation on the **b'xa'c** protein, the I272A assignments were used to follow the chemical shift in WT **b'xa'c** and the L343A and 2DA mutants. Minimal maps can be constructed by measuring distances from signals in an assigned spectrum, in this case I272A, to the nearest peak in the unassigned spectra of WT, L343A and 2DA **b'xa'c**. Figure 5.14 shows the shifts that occurred due to the I272A mutation of **b'xa'c**.

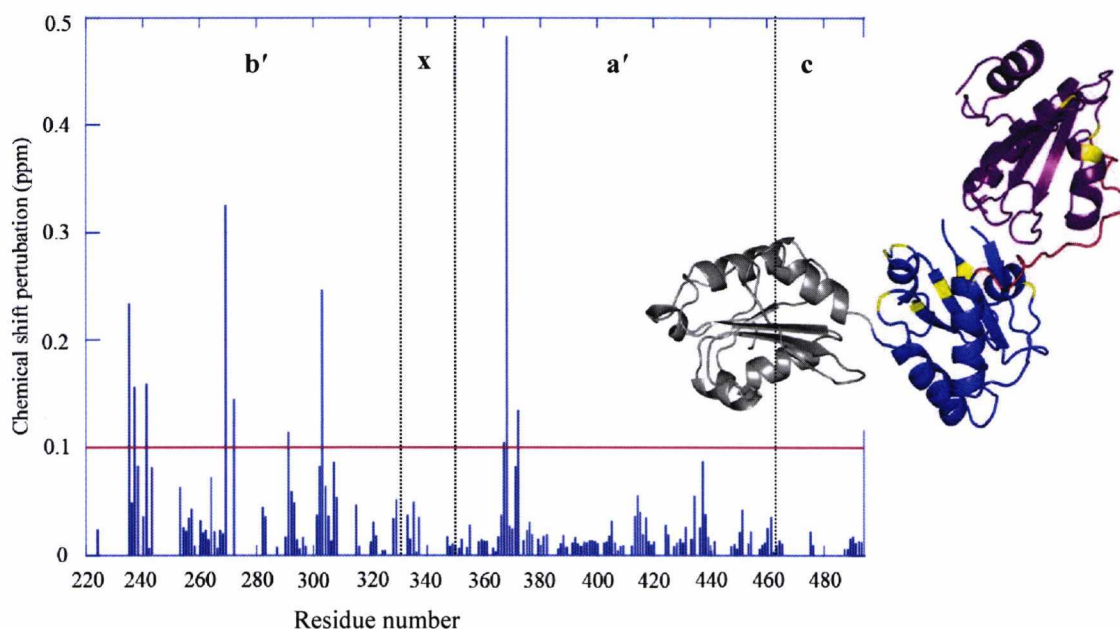


Figure 5.14 Minimal chemical shift map to the nearest peak for backbone amide ^1H and ^{15}N resonances of I272A with WT **b'xa'c** and mapped on the crystal structure of **bb'xa'** (3UEM.pdb). Chemical shift differences greater than 0.1 ppm are highlighted in yellow on the structure. The **b** domain is shown in grey, **b'** in blue, **x** in red and **a'** in purple. The dotted lines represent the domain boundaries.

Chemical shifts observed in figure 5.14, due the I272A mutation in **b'xa'c**, were mainly located in the core β sheet of the **b'** domain, consisting of the site of the mutation and the ligand binding site. Only three residues in the **a'** domain appeared to be affected by the mutation in the **b'x** domain. They were F365, D366 located at the end of helix 7 and N370 located at the start of strand 8. These residues are usually weak and line broadened in the HSQC spectra. The average chemical shift due to the I272A mutation was 0.03 ppm with a standard deviation of 0.06 ppm.

Figure 5.15 shows that the L343A mutation had a greater effect on the **a'** domain than I272A, as more **a'** residues appeared to have shifted.

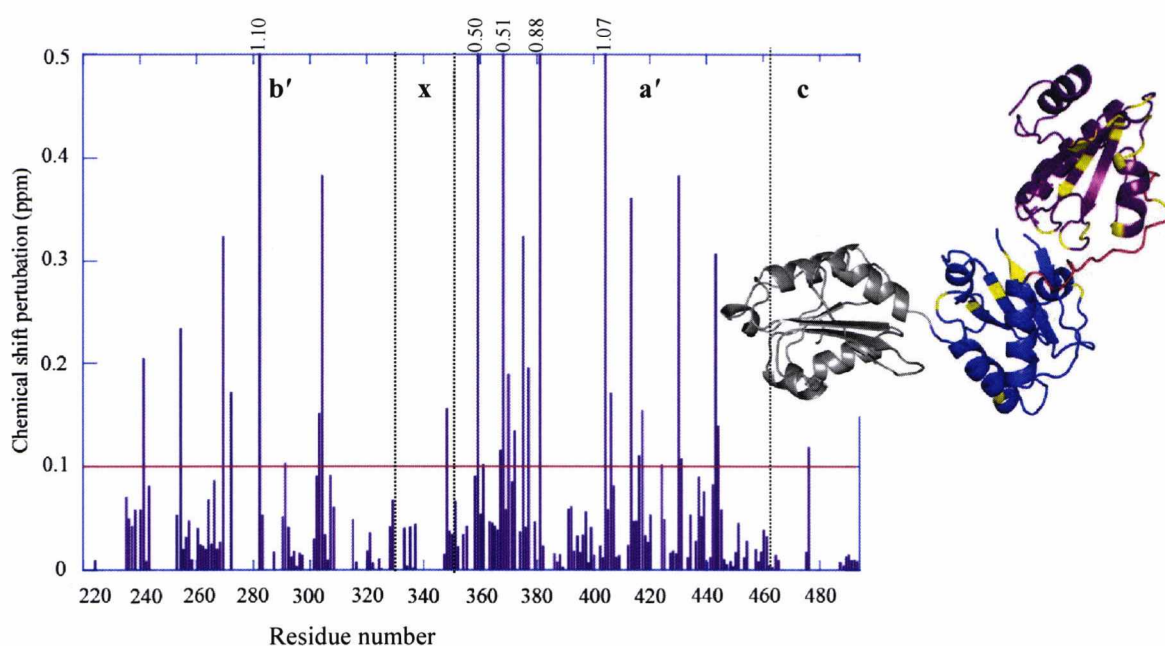


Figure 5.15 Minimal chemical shift map for backbone amide ^1H and ^{15}N resonances of I272A with L343A **b'xa'c** and mapped on the crystal structure of **bb'xa'** (3UEM.pdb). Chemical shift differences greater than 0.1 ppm are highlighted in yellow on the structure. The **b** domain is shown in grey, **b'** in blue, **x** in red and **a'** in purple. The dotted lines represent the domain boundaries.

The majority of the chemical shifts in the **b'** domain are the same as those in figure 5.14 caused by the I272A mutation, with the addition of L241 in strand 3 and L300 in strand 4 which forms part of the ligand binding site. The L343A mutation in the **x** linker region had a greater effect on the conformational stability of the **a'** domain more than that of **b'**, as most of the chemical shifts were mapped on the **a'** domain. The average chemical shift due to the L343A mutation was 0.08 ppm with a standard deviation of 0.16 ppm.

The double mutation at the C-terminus of the **x** linker region, 2DA, had a more pronounced effect on the chemical shift perturbations of the **a'** domain than the other mutants as apart from the same **b'** shifts observed in figures 5.14 and 5.15, the biggest chemical shifts were located in the **a'** domain.

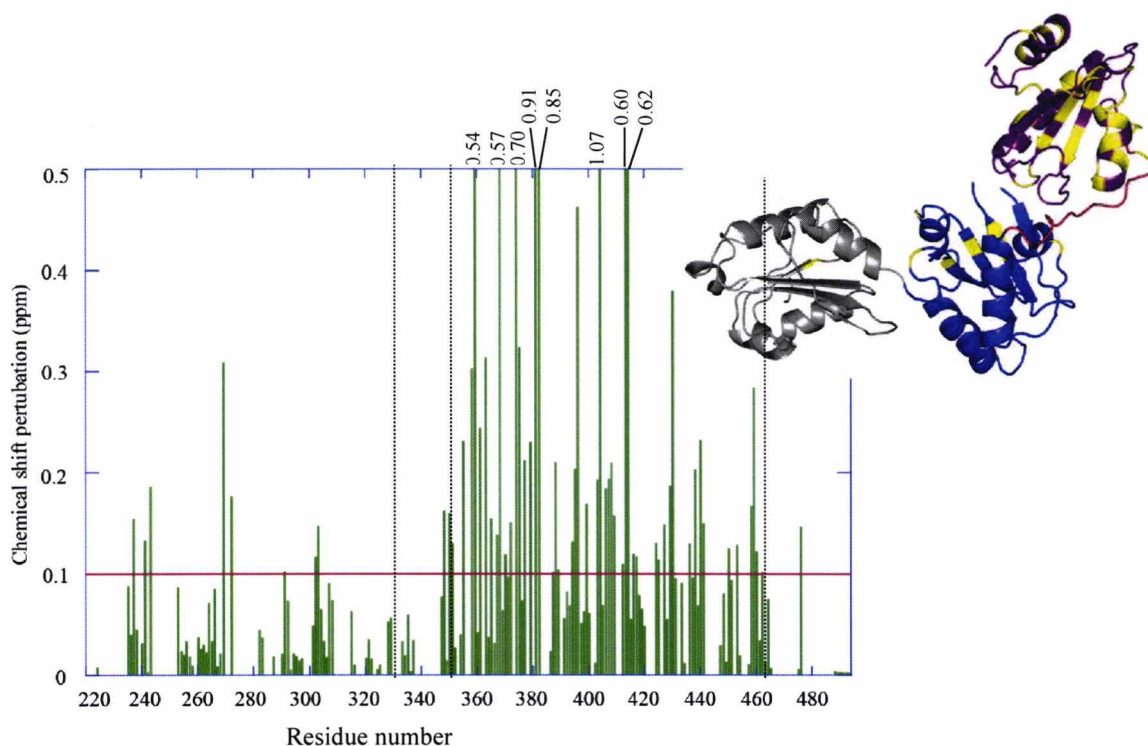


Figure 5.16 Minimal chemical shift map for backbone amide ^1H and ^{15}N resonances of I272A with 2DA **b'xa'c and mapped on the crystal structure of **bb'xa'** (3UEM.pdb).** Chemical shift differences greater than 0.1 ppm are highlighted in yellow on the structure. The **b** domain is shown in grey, **b'** in blue, **x** in red and **a'** in purple. The dotted lines represent the domain boundaries.

The large degree of line broadening in 2DA manifests as large minimal shift differences when compared to I272A; clearly many **a'** peaks are broadened beyond detection, giving rise to large shift perturbations. The average chemical shift perturbation due to the 2DA mutation was 0.12 ppm with a standard deviation of 0.17 ppm.

However, it is important to remember the differences between the NMR spectra of the L343A and 2DA mutants when comparing the minimal maps. This is because L343A and 2DA displayed a high degree of line broadening; therefore some of the shifts in figures 5.15 and 5.16 may not necessarily be due to actual peak shifts but due to peaks that are line broadened beyond detection, consequently missing in the L343A and 2DA spectra. However, the minimal maps did successfully show the effect of the mutations on **b'xa'c**, whether it was due to peak shifts or line broadening.

5.3.5 Determination of capping in **b'xa'c** by NMR spectroscopy

Capping of the ligand binding site by the **x** linker region has been previously determined by tracking of the indole NH of W347 in the **b'x** fragment of PDI. As described in 5.1.2, capping is characterised by a downfield shift of the W347 indole NH. In order to determine if capping of the ligand binding site on the **b'** domain by the **x** linker region did occur in the presence of the **a'** domain, the indole region of full length PDI was overlaid with spectra of WT **b'xa'c** and the I272A, L343A and 2DA mutants, shown in figure 5.17.

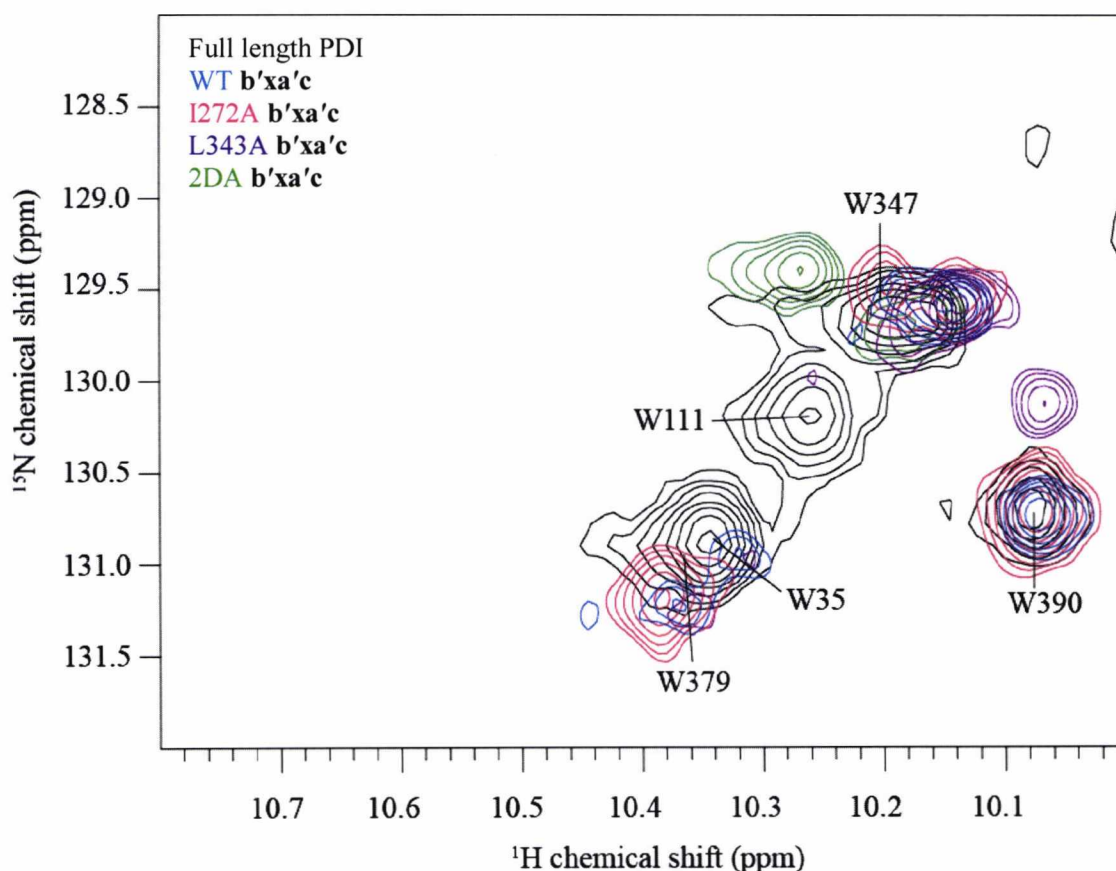


Figure 5.17 Conformational exchange in the environment of W347 of **b'xa'c.** Overlay of the indole region of $^{15}\text{N}/^1\text{H}$ HSQC spectra of full length hPDI (black), WT **b'xa'c** (blue) **b'xa'c** mutants I272A (red), L343A (purple) and 2DA (green). Assignments of full length PDI resonances are indicated.

As can be seen from figure 5.17, there was no significant downfield shift in the indole NH of WT **b'xa'c** and its mutants. However, the indole NH of W347 of I272A, shown in red, was split into two peaks and indicative of at least two different environments

of the tryptophan indole that could support two conformations of the **x** linker region. The indole NH of the 2DA mutant also appeared in two states, confirming the conformational exchange observed in figure 5.16. However, as the quality of the 2DA NMR spectra was very poor, splitting of indole W347 may not always be due to capping. In addition, the indole NH of W379 and W390 in 2DA did not appear on the spectra due to line broadening. The W347 indole NH of WT **b'xa'c** did not show a downfield shift and therefore evidence of binding site capping was inconclusive, but the protein was in at least two conformational states demonstrated by the splitting of the W379 peak. The L343A mutation in **b'xa'c** also showed no evidence of capping, but this was to be expected as this mutant does not favour capping in the **b'x** construct of PDI.

5.3.6 Secondary Structure Prediction using DANGLE

The assignment of resonances to $^1\text{H}_\text{N}$, ^{15}N and C_α for the I272A mutant of **b'xa'c** and mapping of the ^1H and ^{15}N assignments on WT **b'xa'c** $^{15}\text{N}/^1\text{H}$ HSQC spectra enabled a secondary structure prediction to be carried out using DANGLE. Similarly to **xa'c** and **b'x** in chapter 4, the program predicted Ψ and Φ angles and the most likely secondary structure for each backbone residue generating one of three possible outcomes: C (coil), E (strand) and H (helix). Outputs are also shown in the form of Ramachandran plots, as shown in figure 4.15 for **xa'c**, where the number of islands is limited to two and on occasions where more than two islands were present, the residue was excluded from the prediction. Figure 5.18 summarises the predictions made by DANGLE for WT **b'xa'c**. An α -helix is represented by a cylinder, an arrow predicts a β -sheet and line predicts coil. Residues in bold represent data that were excluded from the structure prediction due to missing assignments or ambiguous predictions from DANGLE.

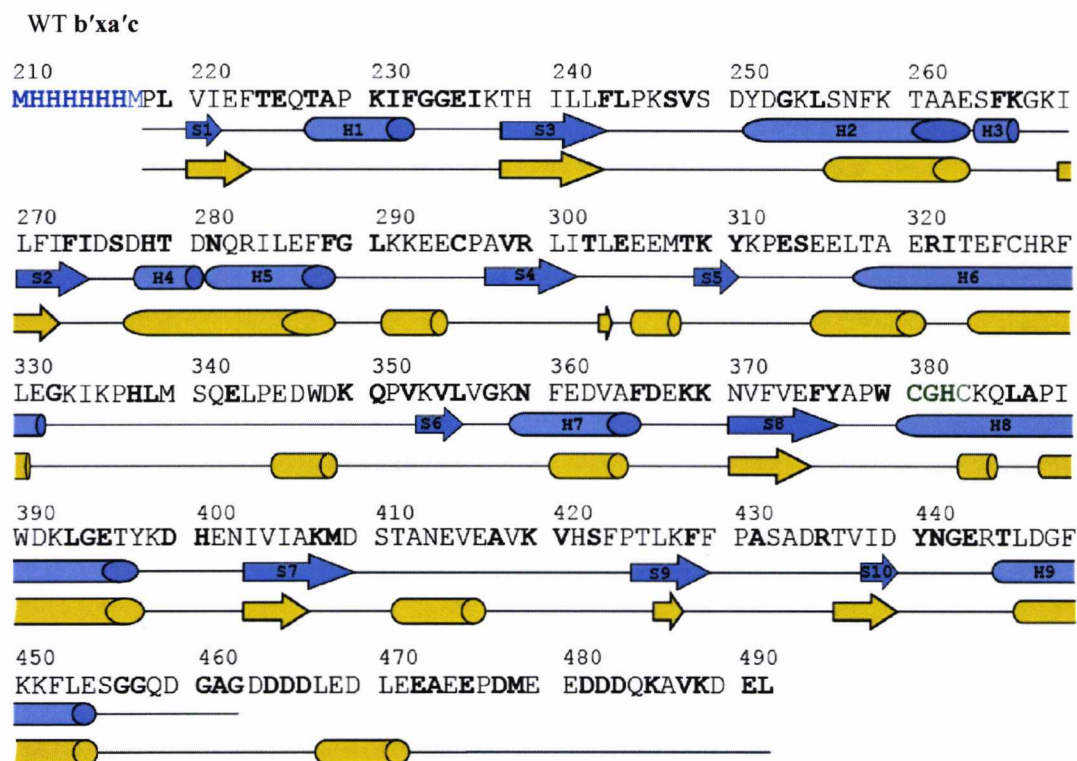


Figure 5.18 The predicted secondary structure of WT **b'xa'c** using DANGLE. The predicted structure is shown in yellow with α -helices shown as cylinders, β -sheets as arrows and coils as lines. The secondary structure motifs of **b'xa'c** are those from the **bb'xa'** crystal structure (3UEM.pdb) and are shown in blue under the amino acid sequence for comparison with the DANGLE predicted structure. The His tag is shown in blue and the active site motif is shown in green. Residues in bold were excluded from the structure prediction. Residue numbering is for mature PDI.

Overall, DANGLE was successful in its prediction of the secondary structure of WT **b'xa'c** with few small differences from the published **bb'xa'** crystal structure (3UEM.pdb). Most of the predicted secondary structure elements were in agreement with the published structure except for helices 2 and 3 which were predicted as a single helix possibly because the program was unable to distinguish the two helices due to their close proximity. Helices 4 and 5 were also predicted as a single helix. Helix 1 is missing from the prediction due to missing assignments for that part of the amino acid sequence. Other small differences included an extra small helix between helix 5 and strand 4 as well as strands 4 and 5. Helix 6 was predicted as two helices instead of one. Interestingly, part of the **x** linker region has been predicted to consist of a helix between residues E345-D348, which is not seen in the crystal structure of **bb'xa'** but is present in the crystal

structure of the I272A mutant of **b'x** (3BJ5.pdb). An extra helix has also been predicted in the loop between strands 7 and 9 as well as part of the **c** extension.

The same structure prediction was run for the I272A mutant of **b'xa'c** and the results generated from DANGLE are shown in figure 5.19.

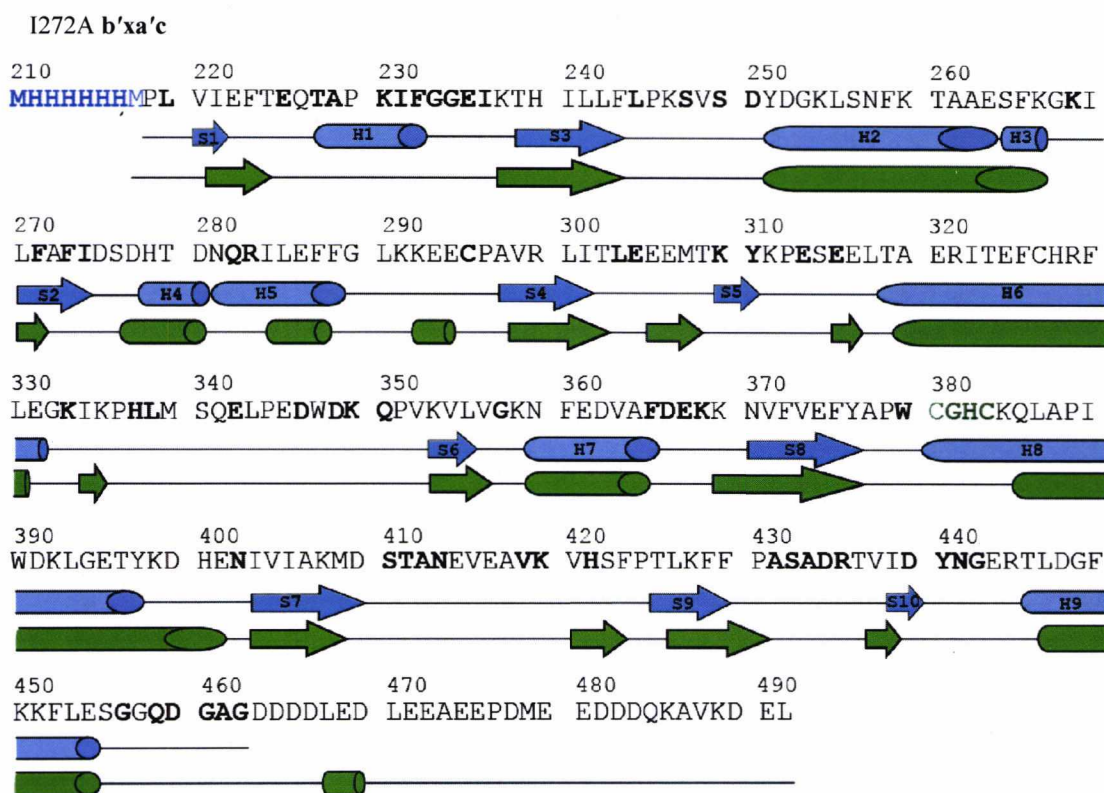


Figure 5.19 The predicted secondary structure of I272A **b'xa'c** using DANGLE. The predicted structure of I272A **b'xa'c** is shown in green below the secondary structure of **bb'xa'** (3UEM.pdb) in blue. Symbols for the secondary structure elements are as stated in figure 5.18.

DANGLE was also successful in predicting the secondary structure elements for I272A **b'xa'c** as overall, the predicted structure was in agreement with the published **bb'xa'** structure of PDI. As with WT **b'xa'c**, helix 1 was missing from the prediction due to lack of assignments and helices 2 and 3 were predicted as a single helix. An extra helix was predicted between helix 5 and strand 4 and small strands were present just before helix 6 and strand 9. A small helix similar to that predicted in WT **b'xa'c** was predicted in the **c** tail.

Generally, the predicted structures for WT and I272A **b'xa'c** were similar not only to the published **bb'xa'** structure (3UEM.pdb) but also to each other, giving confidence in mapping of the assignments from I272A to the WT protein.

5.3.7 Ligand binding

The effect of the peptide ligand Δ -somatostatin on WT and I272A **b'xa'c** were investigated by adding 0.08 mM, 0.2 mM and 0.4 mM of Δ -somatostatin to 0.4 mM protein so that the concentration ratio of protein to ligand was 5:1, 2:1 and 1:1 respectively. Figure 5.20 shows an overlay of $^{15}\text{N}/^1\text{H}$ HSQC spectra of WT **b'xa'c** in the presence of the different concentrations of Δ -somatostatin.

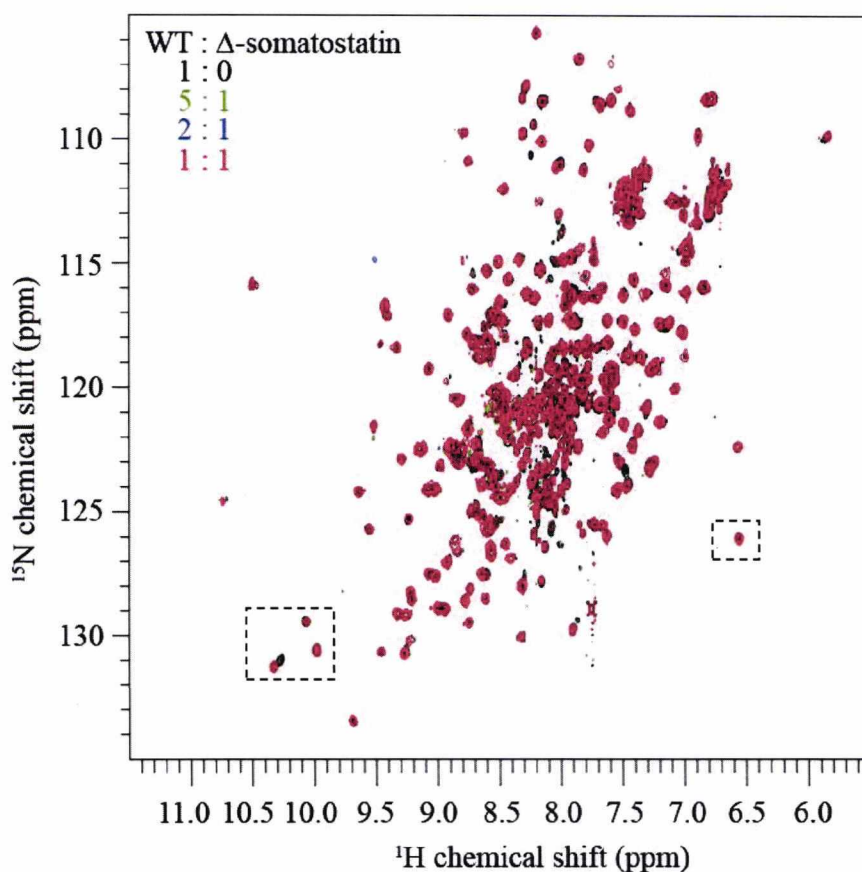


Figure 5.20 Overlay of $^{15}\text{N}/^1\text{H}$ HSQC spectra of WT **b'xa'c** with and without Δ -somatostatin. A control spectrum with the protein in the absence of Δ -somatostatin was run as a reference (black). Δ -somatostatin was added at a protein to ligand ratio of 5:1 (green), 2:1 (blue) and 1:1 (red). Dashed boxes show two of the regions that show chemical shift changes due to addition of ligand.

The chemical shift changes that occurred due to the addition of Δ -somatostatin to WT **b'xa'c** were so small that they could not be easily seen from figure 5.20, so as an example, figure 5.21 shows a zoomed in view of two of the regions that undergo chemical shift changes, highlighted by the dashed boxes in figure 5.20.

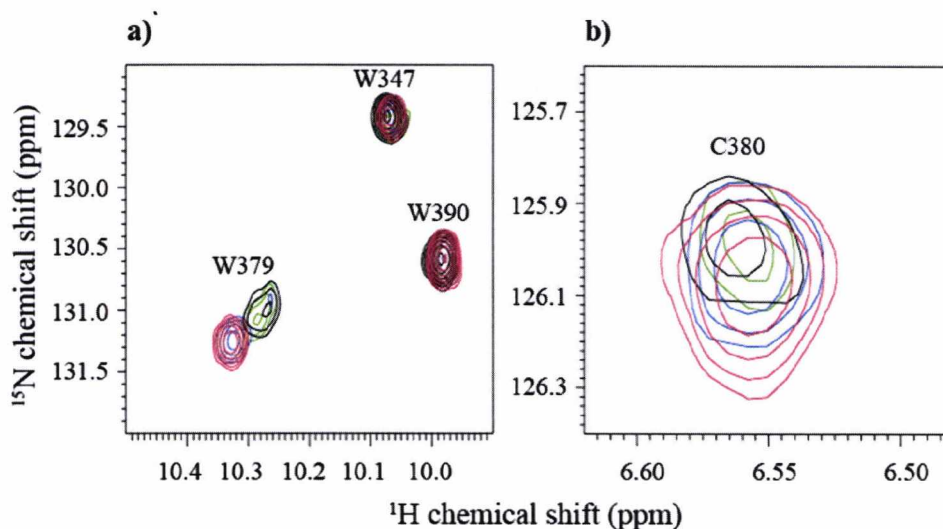


Figure 5.21 Chemical shift changes in WT **b'xa'c due to the addition of Δ -somatostatin.** A zoom in of **a)** the indole region and **b)** C380 of the active site. The black peaks represent the protein in the absence of ligand. Green, blue and red peaks represent the sample with a protein to ligand ratio of 5:1, 2:1 and 1:1 respectively.

The indole NH of W347 did not appear to be affected by the addition of ligand, but there were small shifts in the indole NH of W379, the Trp residue adjacent to the active site, and C380, the N-terminal cysteine residue of the **a'** active site.

Although changes in chemical shifts upon addition of ligand are not easily seen in the spectra overlay in figure 5.20, chemical shifts were monitored by comparison to a control spectrum of the protein in the absence of ligand. The chemical shift perturbation due to binding of the peptide ligand was calculated using equation 4.5 and the minimal map is shown in figure 5.22. The changes in chemical shifts were mapped onto the crystal structure of **bb'xa'** (3UEM.pdb).

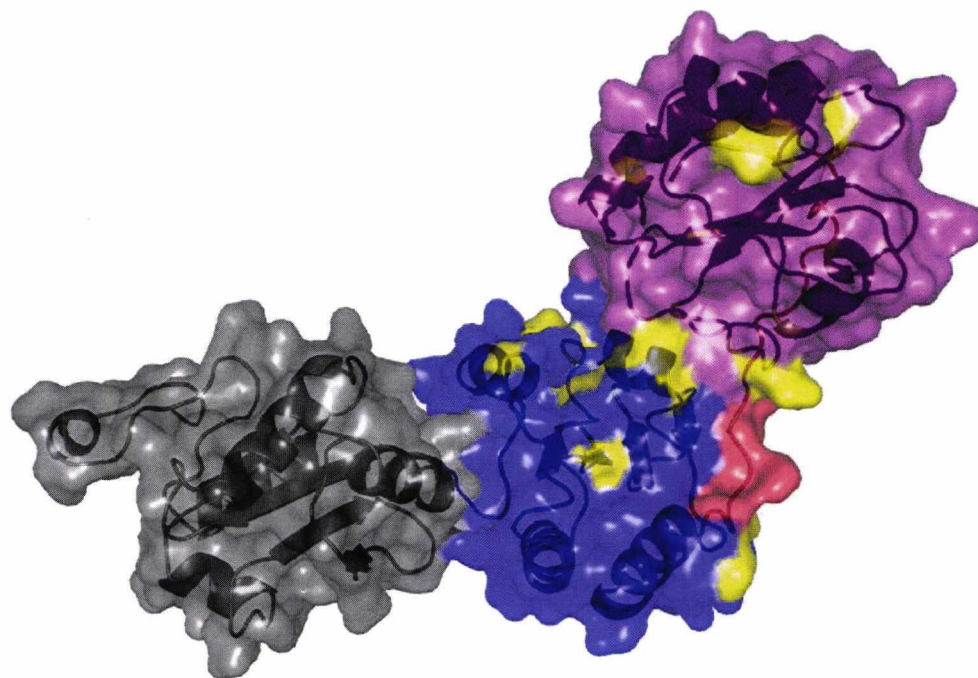
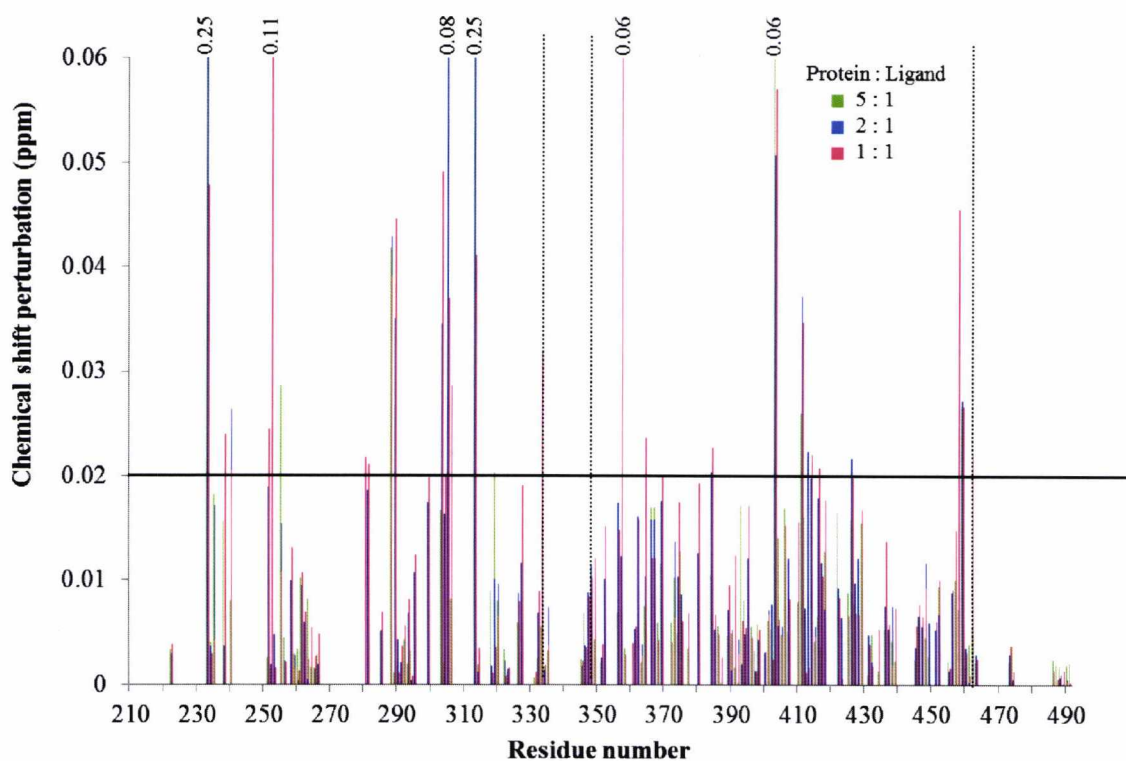


Figure 5.22 Minimal map showing the chemical shift changes of WT **b'xa'c** due to the addition of Δ -somatostatin. The concentration ratio of protein to ligand is 5:1 (green), 2:1 (blue) and 1:1 (red). Chemical shifts above 0.02 ppm were mapped in yellow on the surface representation of the crystal structure of **bb'xa'** (3UEM.pdb) where the **b** domain is shown in grey, **b'** in blue, **x** in red and **a'** in purple.

Chemical shift changes due to addition of Δ -somatostatin to WT **b'xa'c** were seen throughout the structure and most importantly in the core β -sheet of the **b'** domain consisting of the ligand binding site. Shifts were also seen throughout the **a'** domain, such as helices 7 and 9 of the catalytic domain. The loop connecting strands 4 and 5 in the **b'** domain was also highlighted by large chemical shifts, possibly due to its proximity to the active site on the **a'** domain. Chemical shift perturbations observed throughout the spectra appeared to be dependent on the concentration of ligand, with the greatest shifts seen when the protein to ligand concentration ratio was 1:1 with an average and standard deviation of 0.012 and 0.015 ppm respectively. Shifts changes seen at 5:1 protein to ligand in the minimal map presented with an average of 0.011 and a standard deviation of 0.03. A protein to ligand ratio of 2:1 caused chemical shifts with an average of 0.013 ppm and a standard deviation of 0.023 ppm. However the chemical shifts observed in WT **b'xa'c** were very small, in fact on the same scale as figures 5.14 – 5.16, they were almost insignificant. However, shifts in the **b'** domain do map with those observed by Byrne et al. (2009), but they are a whole order of magnitude smaller. Most of the shifts appeared to form an extended surface at the interface between the **b'** and **a'** domains. It is also worth mentioning here that assignments for WT **b'xa'c** spectra were those mapped from I272A and not confirmed by triple resonance.

The I272A **b'xa'c** mutant also appeared to bind to the peptide ligand Δ -somatostatin as chemical shift changes were observed in the $^{15}\text{N}/^1\text{H}$ HSQC spectra at different concentrations of ligand. The largest shifts were observed at the protein to ligand concentration ratio of 1:1 and the lowest chemical shift perturbations were seen at the protein to ligand concentration of 5:1. Figure 5.23 shows an overlay of the $^{15}\text{N}/^1\text{H}$ HSQC spectra of I272A **b'xa'c** in the presence of the different concentrations of Δ -somatostatin.

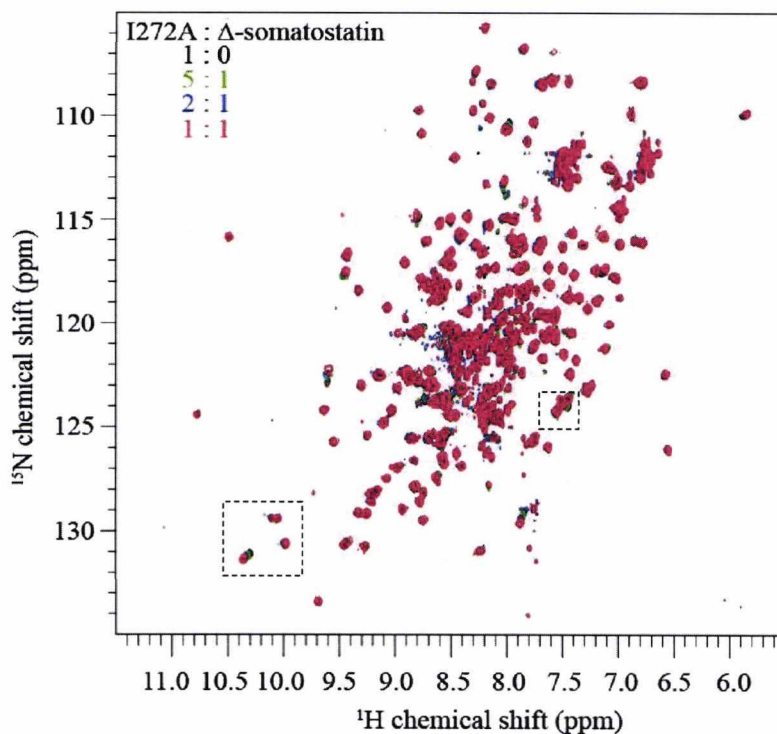


Figure 5.23 Overlay of $^{15}\text{N}/^1\text{H}$ HSQC spectra of I272A b'xa'c with and without Δ -somatostatin. A control spectrum with the protein in the absence of Δ -somatostatin was run as a reference (black). Δ -somatostatin was added at a protein to ligand ratio of 5:1 (green), 2:1 (blue) and 1:1 (red). Dashed boxes show two of the regions that show chemical shift changes due to addition of ligand.

A closer view of the areas in the dashed boxes is displayed in figure 5.24.

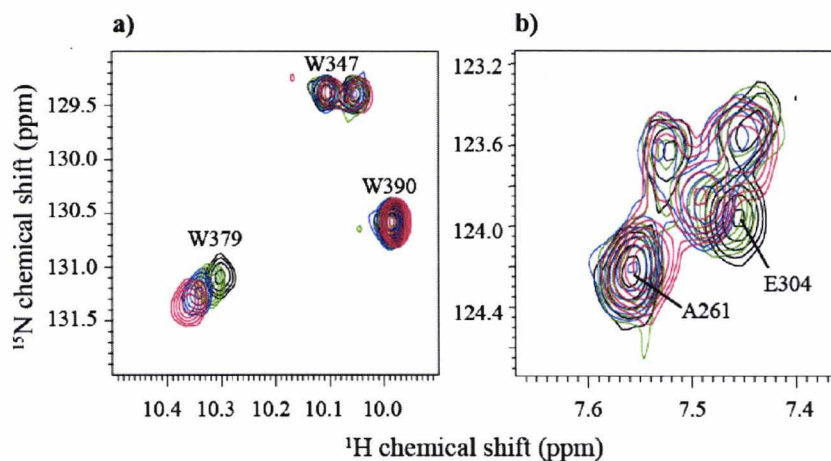


Figure 5.24 Chemical shift changes in I272A b'xa'c due to the addition of Δ -somatostatin. A zoom in of a) the indole region and b) A261 in helix 2 and E304 in the loop linking strands 4 and 5. The black peaks represent the protein in the absence of ligand. Green, blue and red peaks represent the sample with a protein to ligand ratio of 5:1, 2:1 and 1:1 respectively.

Similarly to WT **b'xa'c**, chemical shift changes in the I272A mutant also appeared to be dependent on ligand concentrations as a linear relationship is observed between the peak shift and the concentration of ligand. Also, the main shift in the indole region, in figure 5.24a, was due to W379 and W347 in **x** did not appear affected by Δ -somatostatin. Figure 5.25 shows the chemical shift perturbations that occurred due to the addition of Δ -somatostatin at different concentrations to the I272A **b'xa'c** mutant.

The chemical changes in the I272A **b'xa'c** spectra caused by Δ -somatostatin were mainly located in the core β sheet of the **b'** domain consisting of the ligand binding site as previously mapped by NMR (Byrne et al., 2009). Chemical shift perturbations were also seen in helix 2 and loops between helix 5 and strand 5 as well as strand 4 and strand 5, the latter of which is in close proximity to the **a'** active site in the **bb'xa'** structure. The interface between **b'** and **a'**, especially residues 300 – 305, which have been described by Byrne et al. (2009) to be implicated in ligand binding were also highlighted. Shift differences were observed in residues of helix 7 in **a'** similarly to the WT protein, in addition to shifts in helix 9.

Overall, the average chemical shifts caused to the I272A mutant by the addition of Δ -somatostatin were also very small, similar to those observed in the WT protein, but as assignments were carried out by triple resonance experiments, there was more confidence in the data. At a protein to ligand ratio of 5:1 the average chemical shift for the I272A mutant was 0.006 ppm with the same value for the standard deviation. At 2:1 ratio of protein to ligand the average and standard deviation were 0.009 ppm and at 1:1 chemical shift perturbations had an average and standard deviation of 0.013 ppm.

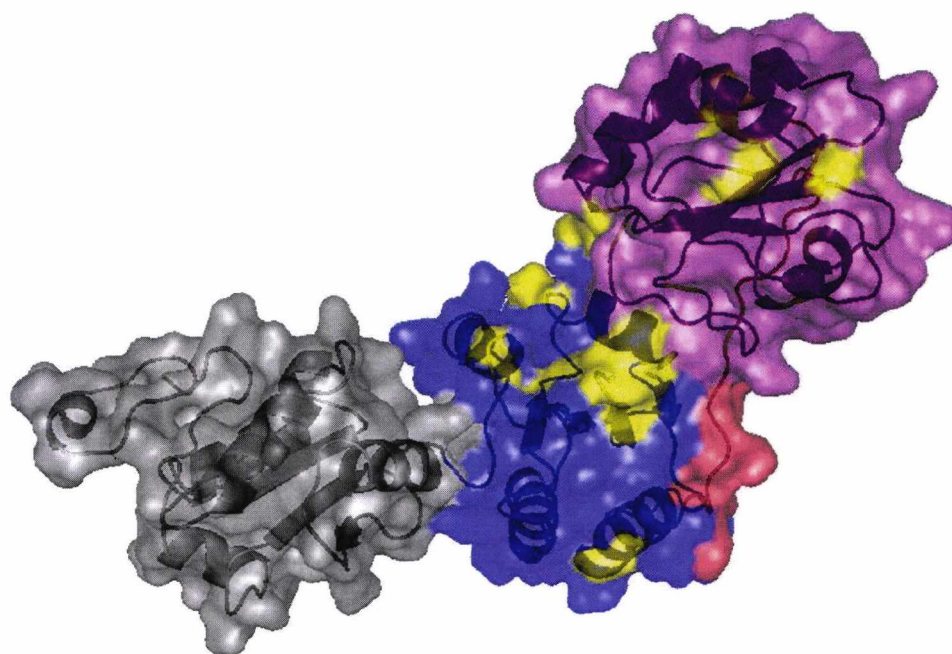
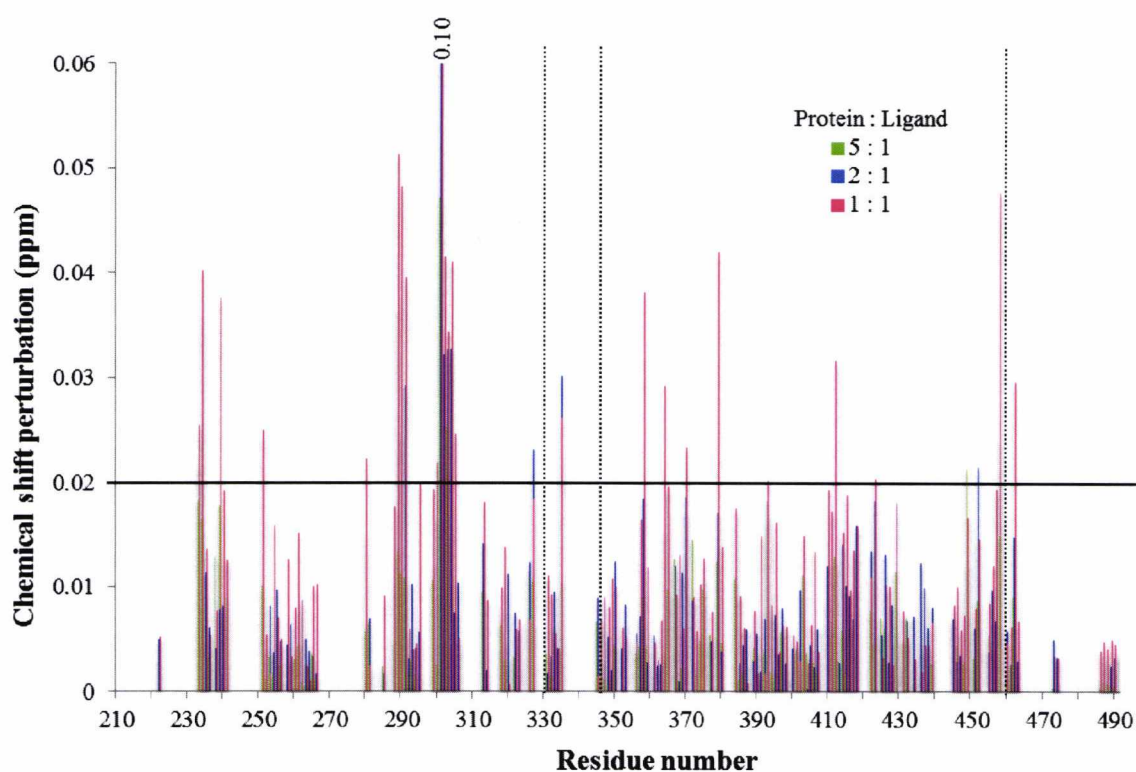


Figure 5.25 Minimal map showing the chemical shift perturbations of I272A **b'xa'c** due to the addition of Δ -somatostatin. The concentration ratio of protein to ligand is 5:1 (green), 2:1 (blue) and 1:1 (red). Chemical shifts above 0.02 ppm were mapped in yellow and green on the surface representation of the crystal structure of **bb'xa'** (3UEM.pdb) where the **b** domain is shown in grey, **b'** in blue, **x** in red and **a'** in purple.

5.3.8 Relaxation Dynamics

Tables of peak heights for ^{15}N T_1 , T_2 and hetNOE data as well as tables of results from model-free analysis for WT and I272A **b'xa'c** are available in Appendix 5.5 and 5.6 respectively. All graphs shown in this section were created from data in these appendices.

5.3.8.1 ^{15}N T_1 , T_2 and hetNOE data for WT and I272A **b'xa'c**

^{15}N T_1 , T_2 and hetNOE relaxation data for WT and I272A **b'xa'c** are shown in figure 5.26. The secondary structure elements of **b'xa'c** taken from the crystal structure of **bb'xa'** (3UEM.pdb) are shown at the top of the graphs. The mean ^{15}N T_1 and T_2 times obtained from rplot for WT **b'xa'c** were 1078 ± 187.9 ms and 88.20 ± 28.09 ms, respectively, giving a mean T_1/T_2 ratio of 12.22. The mean ^{15}N T_1 and T_2 values for the I272A **b'xa'c** mutant were similar to those of the WT protein at 1056 ± 136.6 for T_1 and 90.74 ± 38.50 for T_2 , giving a mean T_1/T_2 ratio of 11.64. Rplot provided average S^2 and τ_m values of 0.68 ± 0.07 and 11.30 ns for WT **b'xa'c** and 0.71 ± 0.04 and 11.00 ns for the I272A mutant, respectively. The ^{15}N T_1 , T_2 and hetNOE relaxation data for both proteins showed similar trends with deviations observed in loops between helix 5 and strand 4 of the **b'** domain as well as helix 6 and strand 6 in **a'**. Deviations were also seen at residues comprising the **x** linker region as well as the termini of the proteins, which similarly to **xa'c** showed the longest T_2 times and lowest hetNOE values at the C-terminal extension, **c**.

T_1 errors for WT **b'xa'c** were notably larger than those of the I272A mutant, possibly because WT **b'xa'c** presents with NMR data that is of poorer quality than the I272A mutant. T_1 values were obtained using a curve fit and poor data with low signal to noise can reduce the precision of data fitting.

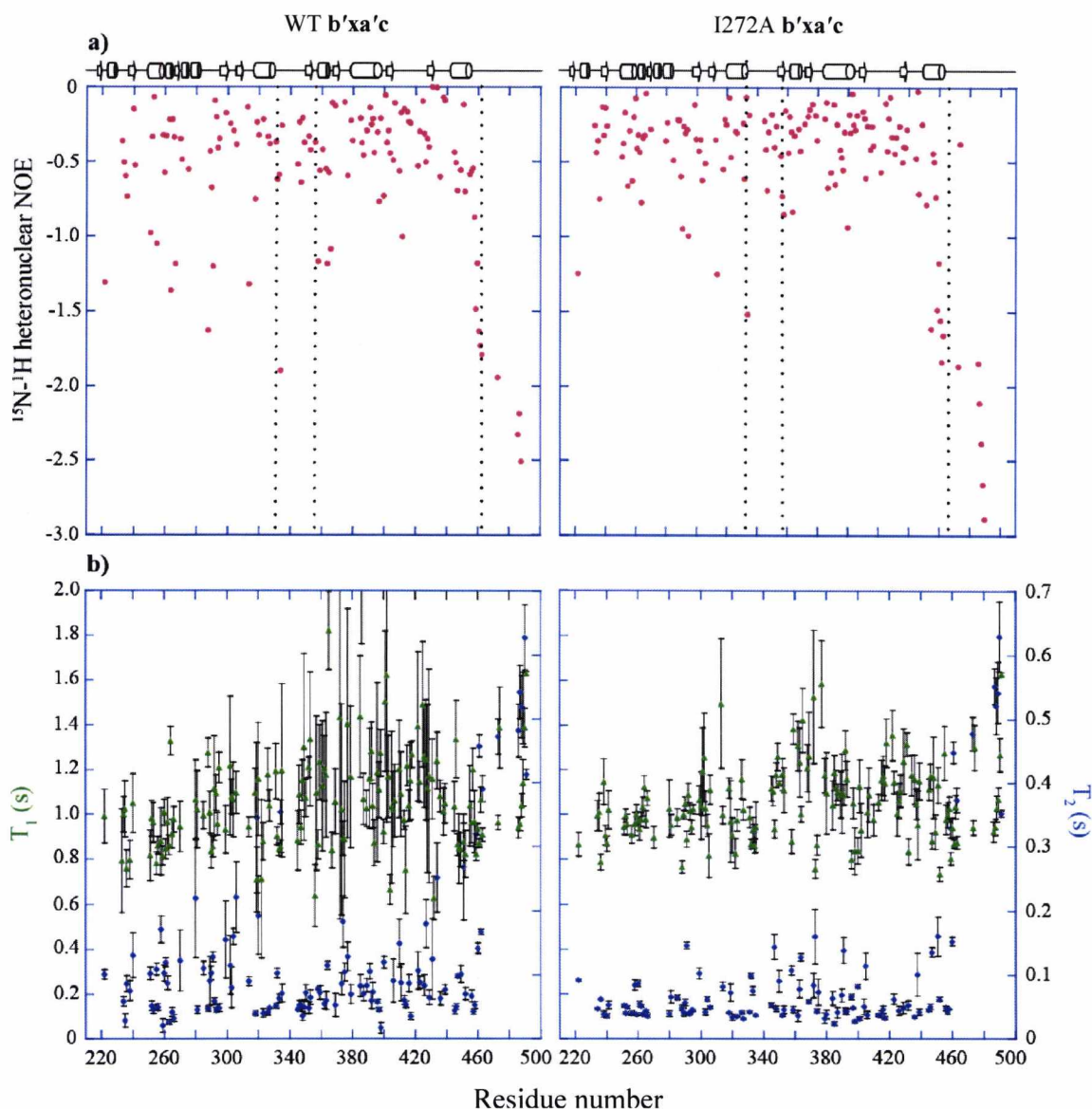


Figure 5.26 ^{15}N T_1 , T_2 and heteronuclear NOE data for WT **b'xa'c** and its I272A mutant. **a)** hetNOE data for each protein is displayed in red with WT **b'xa'c** on the graph on the left and I272A on the right. **b)** T_1 and T_2 data are shown in green and blue, respectively. The secondary structure elements of the **bb'xa'** structure are displayed at the top of the figure with helices represented by cylinders and strands shown as arrows. The dotted lines represent the domain boundaries.

The ^{15}N T_1 , T_2 and hetNOE relaxation data was exported into ModelFree4 to allow the calculation of model-free parameters for individual residues in both WT and I272A **b'xa'c**.

5.3.8.2 ^{15}N T_1 versus T_2 relaxation data for WT and I272A **b'xa'c**

As described in chapter 4, plots of T_1 versus T_2 relaxation data provide information about the global correlation time, τ_m , and the order parameter, S^2 , of the protein under investigation. Figure 5.27 shows a plot of T_1 versus T_2 for ^{15}N amides of WT **b'xa'c**. Data points considered to be outliers are displayed as empty circles and were excluded from rplot analysis in order to get accurate τ_m and S^2 values.

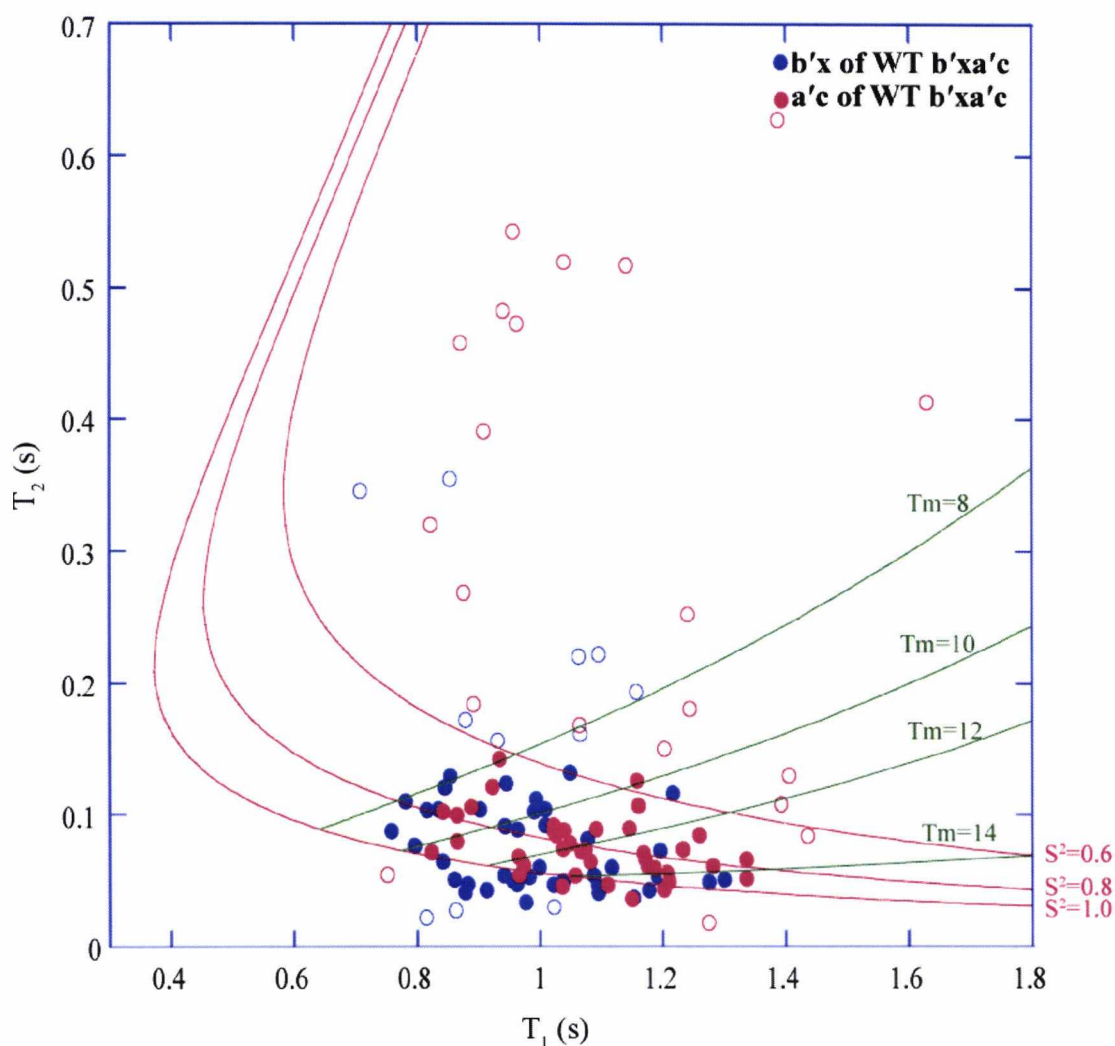


Figure 5.27 Plot of experimental T_1 against T_2 of amide protons for WT **b'xa'c**. Experimental data plotted with theoretical values of order parameter $S^2=0.6, 0.8$ and 1.0 and τ_m of 8, 10, 12 and 14 ns when τ_e is 50 ps at 600 MHz. **b'x** residues are displayed as blue circles and **a'c** residues as red circles. Empty circles represent residues that deviate from the plot and were excluded from rplot analysis.

The majority of the residues for WT **b'xa'c** fell inside the theoretical lines for the order parameter S^2 . A number of residues that fell under the $S^2 = 1.0$ contour line suggested an S^2 value greater than 1.0 for these residues. However, as this is not theoretically possible in model-free, it is possible that these residues were undergoing conformational exchange and suggest contribution from R_{ex} . The wide T_1 vs T_2 plot for WT **b'xa'c** signifies an anisotropic system because there is a large variation in T_1 .

The theoretical τ_m and S^2 lines were constructed using model values of ^{15}N T_1 and T_2 with a constant τ_e of 50 ps, without taking into account hetNOE values, and are useful in estimating τ_m values required for model-free analysis. Taking into account the number of residues and temperature that was used during the acquisition of the relaxation data, the theoretical value for **b'xa'c** is 12.29 ns, using equation 4.1. The measured τ_m from rplot is 11.30 ns, approximately 1.0 ns lower than expected. This is probably due to interdomain motion which lowers the apparent τ_m . Interestingly, **b'x** and **a'c** residues of WT **b'xa'c** fell in the same area of the plot, therefore displaying similar tumbling times.

The T_1 vs T_2 plot generated from the I272A **b'xa'c** experimental data displayed similar results to the WT protein with an average τ_m of 11.00 ns. Unlike WT **b'xa'c**, I272A residues clustered in a smaller area of the plot suggesting a more compact structure. A greater number of residues fell outside of the theoretical $S^2 = 1.0$ line due to conformational exchange similar to I272A **b'x** seen in chapter 4. The T_1 vs T_2 plot for I272A **b'xa'c** with theoretical τ_m and S^2 lines is shown in figure 5.28.

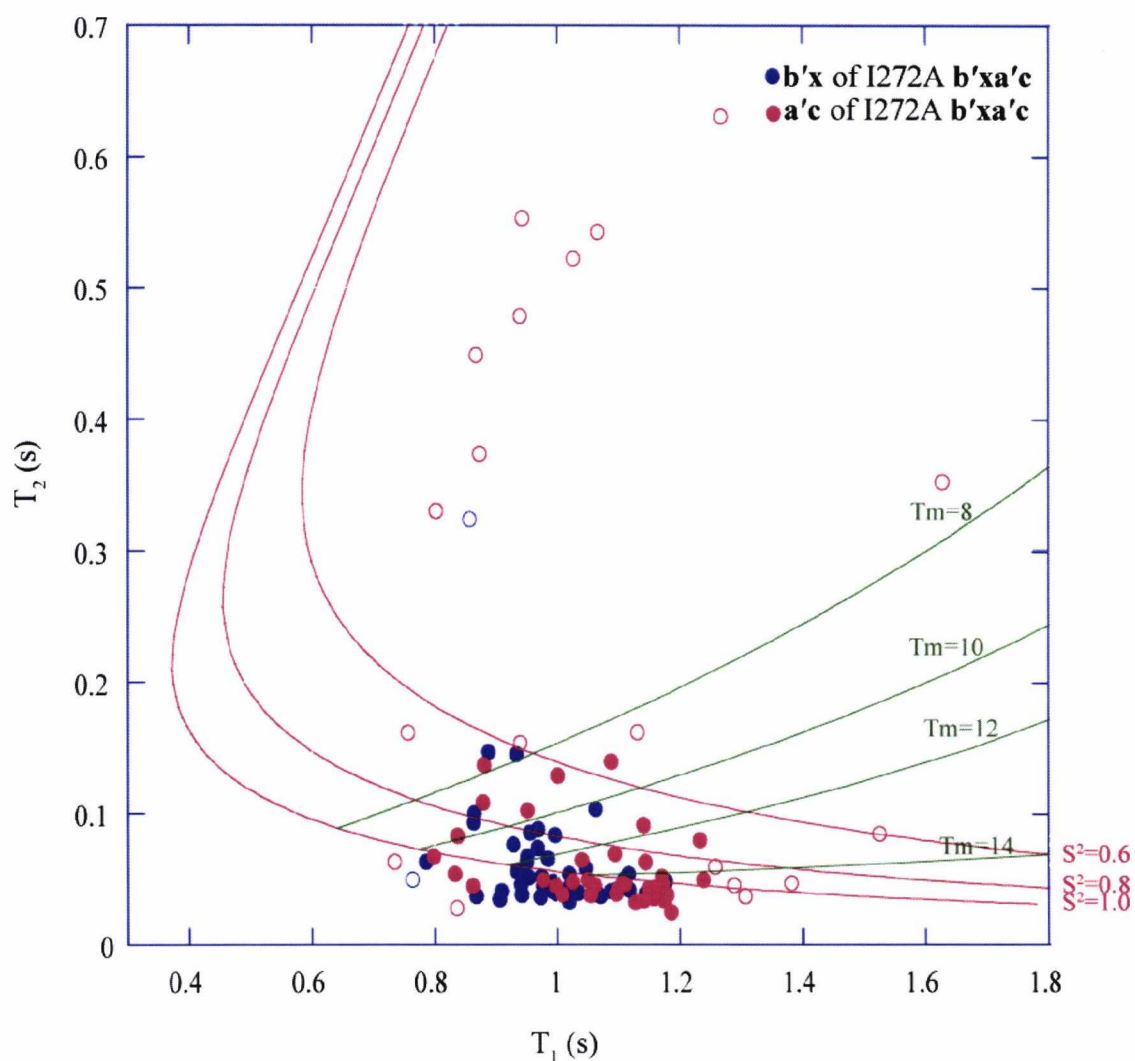


Figure 5.28 Plot of experimental T_1 against T_2 of amide protons for I272A **b'xa'c**. Experimental data plotted with theoretical values of order parameter $S^2 = 0.6, 0.8$ and 1.0 and τ_m of 8, 10, 12 and 14 ns when τ_c is 50 ps at 600 MHz. **b'x** residues are displayed as blue circles and **a'c** residues as red circles. Empty circles represent residues that deviate from the plot and were excluded from rplot analysis.

5.3.8.3 Model-free analysis of WT and I272A **b'xa'c**

Model-free analysis of **b'xa'c**, both WT and I272A, used all three relaxation parameters ^{15}N T_1 , T_2 and hetNOE in order to define the molecular motions of each protein, using the Modelfree4 program. Model-free data for WT and I272A **b'xa'c** is summarised in figure 5.29. As both proteins displayed values outside the S^2 curves in figures 5.27 and 5.28, the model 3, which calculated S^2 and R_{ex} values for each residue, was used.

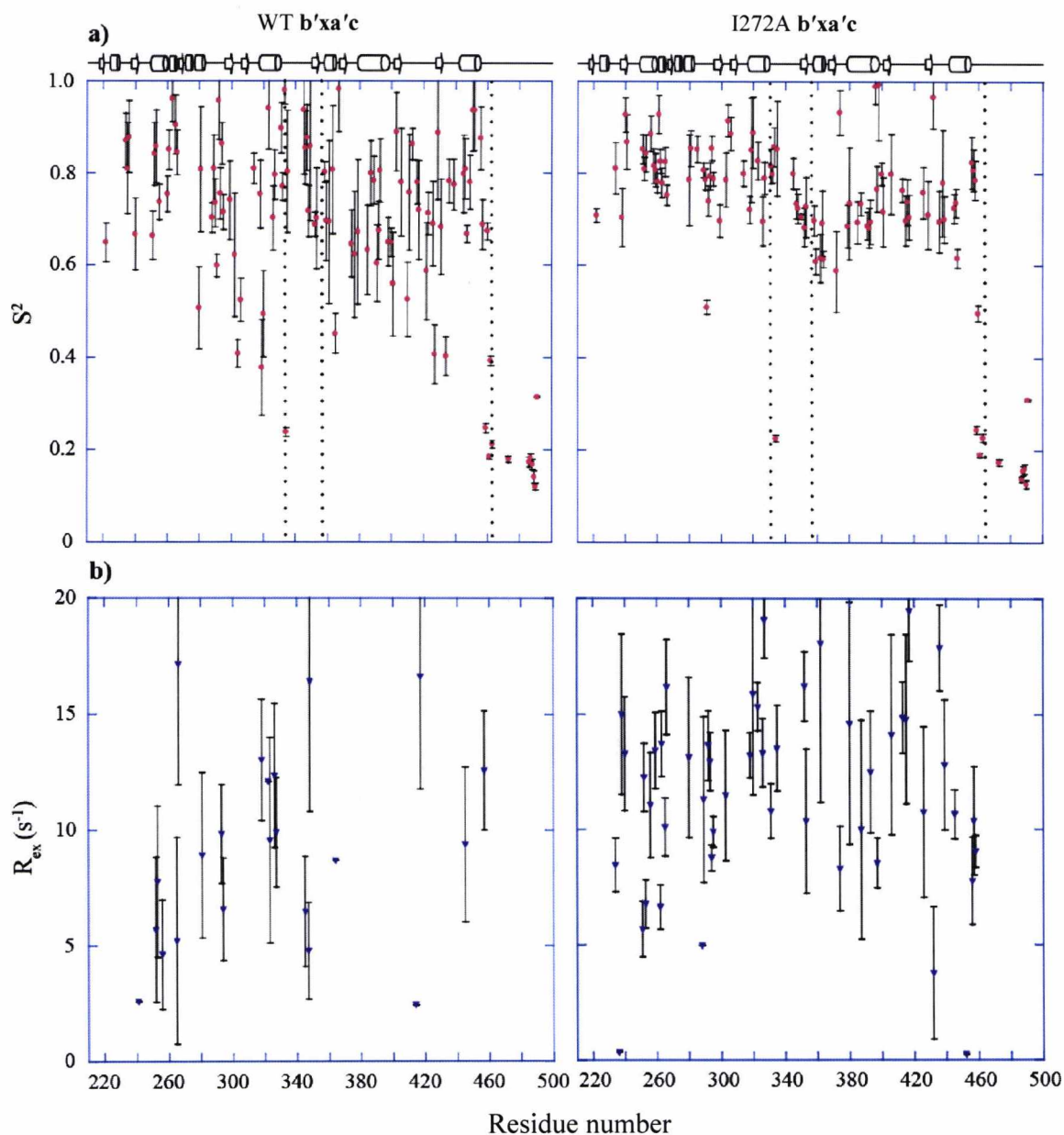


Figure 5.29 Graphs of model-free fit parameters for the characterisation of motion of WT and I272A **b'xa'c**. a) ^{15}N T_1 , T_2 and hetNOE values for each protein were used with analysis program Modelfree4 to calculate residue specific order parameter values of S^2 b) R_{ex} contribution was used for those residues that were detected by Modelfree4 to display conformational exchange. The secondary structure elements of the **bb'xa'** structure (3UEM.pdb) are displayed at the top of the graphs. The dotted lines represent the domain boundaries.

WT **b'xa'c** had an average S^2 of 0.68 ± 0.07 , whereas I272A presented with an S^2 value which was slightly higher at 0.71 ± 0.04 . Both proteins showed the C-terminal extension, **c**, to have a lower S^2 value than the rest of the residues in the protein, with **c** in

WT **b'xa'c** having an S^2 of 0.25 ± 0.01 and in I272A presenting with an S^2 of 0.22 ± 0.01 . Table 5.1 summarises the average S^2 values in specific regions of the **b'xa'c** WT and I272A proteins derived from the model-free analysis.

Fragment of b'xa'c	S^2 of WT protein	S^2 of I272A mutant
b'xa'c	0.68 ± 0.07	0.71 ± 0.04
b'x	0.75 ± 0.07	0.79 ± 0.04
b'	0.74 ± 0.07	0.80 ± 0.04
x	0.79 ± 0.08	0.71 ± 0.04
a'c	0.62 ± 0.06	0.64 ± 0.04
a'	0.72 ± 0.08	0.70 ± 0.05
c	0.25 ± 0.01	0.22 ± 0.01

Table 5.1 A summary of the average S^2 values of specific regions of WT and I272A **b'xa'c** calculated from model-free analysis.

Overall the **b'x** fragment of WT **b'xa'c** had a lower S^2 value than the I272A mutant. The **b'x** fragment of **b'xa'c** also had a higher S^2 than the **a'c** fragment in both the WT protein and the I272A mutant. The **b'** domain of I272A had a more rigid backbone than the **b'** of the WT protein, presenting with an S^2 of 0.8 ± 0.04 for I272A and an S^2 of 0.74 ± 0.07 for WT **b'xa'c**. Whereas the **a'** domain appeared to have similar backbone flexibility in both the WT and I272A **b'xa'c**. The **b'** domain of I272A had a higher S^2 value than that of its **a'** domain by 0.1 order parameter units. Conversely, the **b'** domain in WT **b'xa'c** had a similar S^2 value to that of the **a'** domain. Interestingly, the **x** linker region in WT **b'xa'c** presented with a slightly higher S^2 of 0.79 ± 0.08 than in the I272A mutant where **x** had an S^2 value of 0.71 ± 0.04 .

Deviations from the average S^2 were observed in residues L300 – E320 of WT **b'xa'c** which stretch from the end of strand 4 to the start of helix 6. This region had an average S^2 of 0.57 ± 0.07 . This same region did not appear to deviate from average values in I272A **b'xa'c**. Whereas in I272A **b'xa'c**, the region with reduced S^2 values of 0.65 ± 0.05 consisted of residues V352 – F372, which in the structure make up strand 6, helix 7 and part of strand 8.

A number of residues in WT **b'xa'c**, the majority of which are located in the **b'** domain, displayed contribution from R_{ex} that supports conformational exchange. However, I272A **b'xa'c** had more residues with R_{ex} contribution evenly distributed between the **b'** and **a'** domains, highlighting the fact that the mutation in **b'** also affects the **a'** domain.

5.4 Discussion

5.4.1 The alternative conformations of **b'xa'c** studied by NMR

It was evident from initial $^{15}\text{N}/^1\text{H}$ HSQC spectra that WT **b'xa'c** was a conformationally active protein displaying line broadened and poorly resolved peaks at 25°C and pH 6.5 (figure 5.4). The low conformational stability of the **a'** domain, determined by fluorescence, may have contributed to the poor quality of the NMR spectra. Similarly to **a'c** and **xa'c**, the unstructured acidic region at the C-terminus of WT **b'xa'c** presented with sharp peaks between 8.0 and 8.5 ppm, adding to the poor quality of NMR spectra. Chemical reduction of the **a'** active site significantly improved the quality of NMR spectra, suggesting that the redox state of the **a'** domain influenced a conformational change in **b'xa'c** leading to this improvement. This is in agreement with recently published data on **bb'xa'** and **abb'xa'** where it was shown that reduction of the **a'** active site turns the **b'xa'** fragment of PDI into a more compact structure where **a'** forms interactions with the **b'** domain and the **x** linker region (Wang et al., 2012b, Wang et al., 2012a). A more compact structure would tumble more isotropically and have potential to produce higher quality NMR data. In contrast, oxidation of the active site on the **a'** domain of PDI favours a more conformationally variable structure where **a'** moves away from **b'** and **x** and exposes a larger hydrophobic binding pocket, presumably to accommodate larger ligands. As the reduced protein forms a more compact conformation, as opposed to the more flexible oxidised structure, NMR spectra appear less line broadened with better resolved peaks. However, NMR data for **b'xa'c** was still challenging as, although the reduced and oxidised crystal structures of **abb'xa'** describe the change in conformation that occurs due to the redox switch in **a'**, it is not known how this conformational change affects the protein in solution. Even in the reduced state, **b'xa'c** was still presented with line broadened $^{15}\text{N}/^1\text{H}$ HSQC spectra.

b'xa'c mutants, I272A, L343A and 2DA, were generated to investigate if capping of the ligand binding site by the **x** linker region occurred in the presence of the **a'** domain. Initial NMR spectra acquired at 25°C and pH 6.5 showed that, like WT **b'xa'c**, these mutants produced spectra that were line broadened due to conformational exchange. NMR spectra acquired at a range of temperatures provided sharper peaks at the higher temperature of 40°C. This is because temperature can affect NMR spectra in two ways. First, increasing temperature reduces viscosity and increases correlation time, which in turn narrows NMR lines. Second, temperature increases conformational rates of exchange

and moves NMR time regimes from intermediate line broadening into fast exchange. An increase in pH also proved to have a positive effect on the quality of the NMR spectra with the best spectra acquired at pH 7.0 and 8.0. This is not surprising as the pH of the endoplasmic reticulum is around 7.3 so at pH 7.0, PDI fragments are closer to their physiological pH and therefore more likely to be at their most stable conformation.

Once it was established that addition of DTT and an increase in temperature and pH improved the quality of NMR spectra and decreased line broadening, it was possible to compare the effects of the different mutants on **b'xa'c**. The I272A mutant had a similar stabilising effect on **b'xa'c** as it does in **b'x**. The quality of the NMR spectra improved and line broadening was reduced due to the I272A mutation in the **b'** domain. This finding supports the denaturation data discussed in chapter 3, which showed an increase in the conformational stability of both the **b'** and **a'** domain of **b'xa'c**. As the conformational stability of both domains was affected by the I272A mutation in the core β sheet of the **b'** domain, it is likely that there is “cross-talk” between the domains in **b'xa'c**, which means that **b'** and **a'** do not act independently but influence each other's conformational behaviour. The I272A mutant of the **b'x** fragment of PDI favours capping of the ligand binding site by the **x** linker region and has been described as a **b'** stabilising mutant (Nguyen et al., 2008), but in I272A **b'xa'c** there does not appear to be any conclusive evidence that capping occurs in this protein. However, the indole NH of W347 in the **x** linker of I272A **b'xa'c** does not shift to the same extent as the indole NH of W347 in **b'x**, but figure 5.17 suggests that W347 is in at least two different environments demonstrated by the presence of two peaks for the indole NH of W347. This implies that, even though **x** may not be fully capping the ligand binding site, it is possible that it could be undergoing conformational changes that bring it closer to the **b'** domain resulting in a more compact structure for I272A **b'xa'c** than that of the WT protein. This is supported by gel filtration data discussed in chapter 2, where the I272A mutant of **b'xa'c** has a smaller hydrodynamic volume than the WT protein.

The double mutation at the C-terminus of the **x** linker region, in the 2DA mutant, has also been found to be a capping mutant in **b'x**, but in **b'xa'c** it has a detrimental effect in the conformational stability of the **a'** domain leading to line broadened $^{15}\text{N}/^1\text{H}$ HSQC spectra. This highlights the differences in the **b'x** capping mutants and suggests that although both I272A and 2DA favour capping of the ligand binding site in **b'x**, they must do so by different mechanisms. It is also important to mention that 2DA (D346A/D348A)

is a double mutation at either side of W347 in the **x** linker region, in which two negatively charged aspartic acids are substituted with two hydrophobic alanine residues. Therefore hydrophobic interactions between the **x** linker region and the **b'** ligand binding site are increased due the presence of three hydrophobic residues A346, W347 and A348 in **x**. W347 has been shown to interact with M307 in **b'** and the two alanine residues either side of W347 could strengthen this interaction. It is possible that **x** could be trying to form these interactions in 2DA **b'xa'c**, but the presence of the **a'** domain could obstruct these interactions. So, even though the indole of W347 in 2DA may present as two peaks in the line broadened NMR spectra, it is not necessarily indicative of capping.

The L343A mutant of **b'xa'c** appears very similar to its **b'x** homologue, in that it also produces line broadened NMR spectra and shows no evidence of capping of the ligand binding site by **x**. The indole NH of W347 of L343A appears as a single peak in HSQC spectra shown in figure 5.17, unlike that of I272A or 2DA, suggesting that **x** must be free in solution.

So, even though W347 in I272A and 2DA **b'xa'c** appears in at least two conformational states, represented by two peaks in NMR spectra, it does not necessarily mean that capping of the ligand binding site occurs in these proteins. This is because capped **b'x** mutants produce high quality NMR spectra, whereas the 2DA **b'xa'c** mutant does not. Therefore splitting of the W347 indole peak must not be due to capping. So, if splitting of W347 in 2DA is not necessarily indicative of capping of the ligand binding site by the **x** linker region, there is no conclusive evidence to show that capping occurs in I272A **b'xa'c**. Splitting of the W347 indole NH simply means that this residue in **x** is in two or more different conformations.

The NMR data for WT **b'xa'c** and its mutants supports the denaturation studies discussed in chapter 3, as the conformational stability if these proteins is reflected in the quality of the NMR spectra that they produce. The I272A mutant showed increased conformational stability in fluorescence and gave better resolved peaks in $^{15}\text{N}/^1\text{H}$ HSQC spectra. L343A and 2DA especially appear to be destabilising mutants in **b'xa'c**. However, the most important finding is that, although all of these mutations are in the **b'x** fragment of the protein, their effect on the **a'** domain is noticeable, suggesting that the domains have an influence in each other's conformational stability and behaviour. Although capping of the ligand binding site in **b'xa'c** does not appear to be as certain as in **b'x**, evidence of at least two alternative conformations in the **x** linker region is obvious.

5.4.2 Assignment of I272A **b'xa'c** and mapping assignments on WT **b'xa'c**

The percentage of the completed NMR backbone assignments for I272A **b'xa'c** was reported in section 5.3.3. Assignment of I272A **b'xa'c** was even more challenging than the assignment of **xa'c** as in addition to the sharp resonances from the unstructured **c** extension observed between 8.0 and 8.5 ppm and the challenging nature of the **a'** domain, I272A **b'xa'c** is nearly twice the molecular mass of **xa'c** at 32.5 kDa and NMR data suggests that this protein is very conformationally active and prone to line broadening due to conformational exchange. For these reasons triple resonance experiments CBCA(CO)NH and CBCANH were of such poor quality that they could not be used for the assignment of this protein. Instead, the HN(CO)CA and HNCA pair of triple resonance experiments were used which correlate the ^1H and ^{15}N with only the $^{13}\text{C}_\alpha$ of the preceding amino acid and the ^1H and ^{15}N with the $^{13}\text{C}_\alpha$ nuclei of the same and preceding amino acid, respectively. As no information could be obtained for the $^{13}\text{C}_\beta$ nuclei, the assignments obtained from **xa'c** and WT **b'x** were used to confirm the new I272A **b'xa'c** assignments. Similarly to **xa'c**, most of the residues in the unstructured **c** extension could not be assigned. A number of residues at the N-terminus of the **b'** domain could not be assigned due to differences in chemical shifts and ambiguous triple resonance data. Residues of strand 2 could not be assigned because of chemical shifts brought about by the mutation. Interestingly, this strand forms part of the ligand binding site on the **b'** domain and changes in chemical shifts or line broadening could also be brought about by changes in conformation due to possible interactions with the **x** linker region.

Figure 5.30 shows the amino acid sequence of I272A **b'xa'c** with the residues that could not be assigned due to line broadening, peak shifts or missing triple resonance data, highlighted in red.

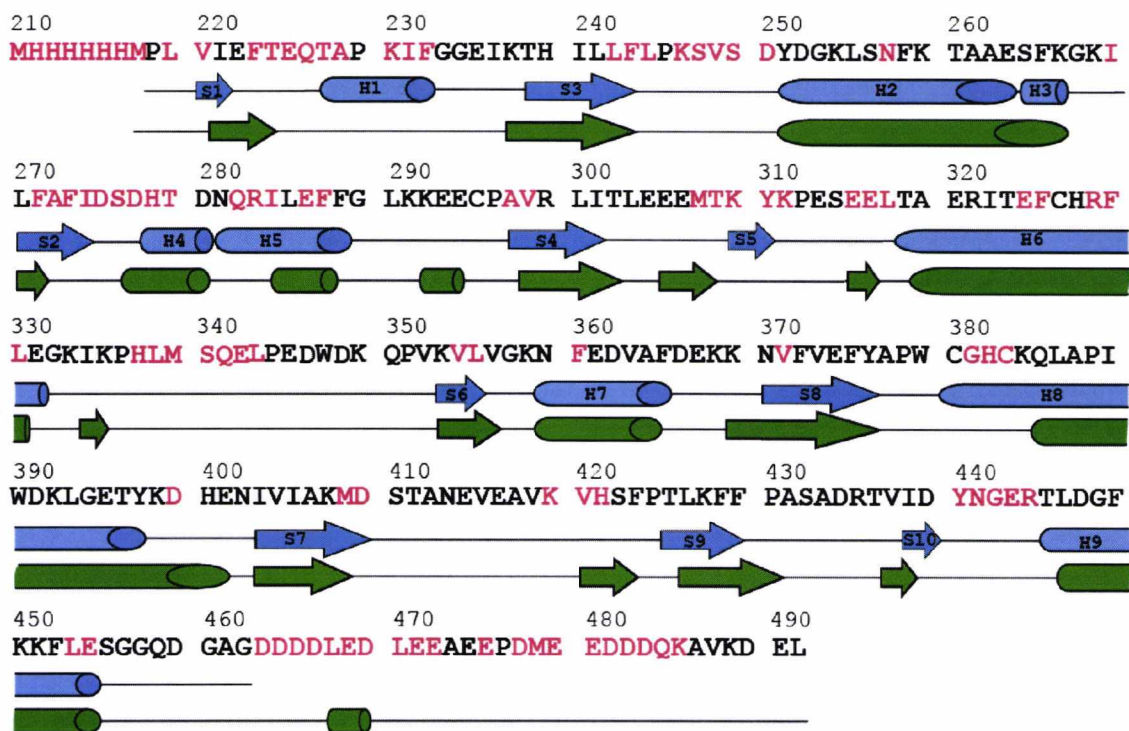


Figure 5.30 Amino acid sequence of I272A **b'xa'c** showing unassigned residues. Residues that could not be assigned from the triple resonance experiments are shown in red. The secondary structure elements from the **bb'xa'** crystal structure (3UEM.pdb) and those from the DANGLE prediction are shown in blue and green, respectively. Proline residues were unassignable as they do not contain a backbone NH.

Assignments of **b'xa'c** were essential for further NMR characterisation of the construct and although 100% assignments were not possible for this protein due to reasons already mentioned, 66% of amide ^1H and ^{15}N assignments were invaluable for backbone relaxation dynamics and ligand binding studies. Additionally, the majority of I272A assignments could be mapped on $^{15}\text{N}/^1\text{H}$ HSQC spectra of WT **b'xa'c**. Residues that could not be mapped from the I272A mutant to the WT protein, such as L241, H239, L300, E305 and R435, have been previously shown to be involved in ligand binding or capping of the ligand binding site by the **x** linker region and are therefore prone to conformational exchange and line broadening (Byrne et al., 2009).

5.4.3 Secondary Structure Prediction using DANGLE

DANGLE (Dihedral Angles from Global Likelihood Estimates) was able to predict the secondary structure of WT and I272A **b'xa'c** using the sequence information, ^1H and

¹⁵N chemical shifts for both **b'xa'c** proteins and the C_α chemical shifts from the triple resonance assignment of I272A **b'xa'c**. Overall, DANGLE was successful in its prediction of the secondary structure elements for both WT and I272A **b'xa'c** with few deviations from the recently solved crystal structure of **bb'xa'** (3UEM.pdb), as shown in figures 5.16 and 5.17. The discrepancies that are seen are mainly in regions where there are missing assignments and therefore there are no chemical shifts available for the software to use in the interrogation of chemical shifts from published structures.

5.4.4 Ligand binding

The **b'** domain has been described as the ligand binding domain of hPDI for several years now and its ligand binding site has been mapped by chemical shift perturbations using NMR spectroscopy (Pirneskoski et al., 2004, Byrne et al., 2009). Interactions of the ligand binding site with ligands have been shown to be primarily hydrophobic in nature. The recent crystal structures of **bb'xa'** and **abb'xa** show the hydrophobic binding site of PDI extending to the surfaces of the **b** and **a'** domains on the inside cleft of the “U”-shaped PDI molecule (Wang et al., 2012b, Wang et al., 2012a). Binding of the peptide ligand Δ -somatostatin to the WT **b'xa'c** fragment of PDI displayed chemical shift perturbations in both the **b'** and **a'** domains, seen in figure 5.22, suggesting that both domains contribute to ligand binding. Chemical shifts could be mapped to the core β -sheet of the **b'** domain, consisting of the primary ligand binding site, as well as residues in the **a'** domain. The changes in chemical shifts were mainly distributed across the interface between the **b'** and **a'** domains suggesting conformational changes in that part of the protein upon ligand binding. This is in agreement with published data showing a contribution from both domains in ligand binding (Wang et al., 2012a, Wang et al., 2012b).

The I272A mutant of **b'xa'c** also presented with chemical shift perturbations on ligand binding, but the majority of these shifts were in the **b'** domain of the protein with few shifts on the surface of the **a'** domain. The chemical shift perturbations in the β sheet of the **b'** ligand binding site were greater in the I272A mutant than in the WT protein but the overall chemical shift perturbations were greater in WT **b'xa'c** due to shifts in the **a'** domain. This suggests that ligand binding in the mutant protein has a greater effect on **b'** than the **a'** domain, possibly because the **b'** domain could be partly occluded by the **x** linker region and therefore it needs to undergo a greater conformational change in order to

expose the ligand binding site. Whereas, WT **b'xa'c** forms a more extended conformation, as it has a larger hydrodynamic volume on gel filtration, and more of its hydrophobic surfaces are exposed for ligand binding.

However, the magnitude of the chemical shifts in both the WT and I272A **b'xa'c** is very small compared to published data on **b'x** (Byrne et al., 2009), casting doubt on the significance of these shifts. Also, the W347 indole NH has been shown to shift towards the uncapped state in **b'x** as **x** is displaced from the ligand binding site with increasing concentrations of ligand. This is seen neither with the WT or I272A **b'xa'c**, but the indole NH of W379 appears affected by the addition of ligand. There are two possible explanations for this. First, **b'xa'c** could be binding ligand but **x** is not capping the ligand binding site, therefore its environment does not change and there is no chemical shift difference in the W347 indole NH. Second, **b'xa'c** is not binding Δ -somatostatin, therefore the indole NH of W347 does not shift as it does in **b'x**.

5.4.5 ^{15}N NMR Relaxation Dynamics of WT and I272A **b'xa'c**

Differences in the conformational stability of WT **b'xa'c** and the I272A mutant have already been shown by intrinsic fluorescence in the presence of denaturant as described in chapter 3. In addition to this, the backbone dynamics of WT and I272A **b'x** as well as **xa'c** were discussed in chapter 4 and highlighted the effect of the I272A mutation on the **b'** domain. Therefore it was expected that the dynamic properties of WT **b'xa'c** would be different to those of its I272A mutant. Also, recently published crystallography data on **bb'xa'** and **abb'xa'** have shown the protein to be very conformationally active, especially in the **b'xa'** region.

5.4.5.1 Relaxation data

The three relaxation parameters used to describe the molecular motions of WT and I272A **b'xa'c** were ^{15}N T_1 , T_2 and hetNOE. From the graphs in figure 5.25 there do not appear to be any major differences in the HetNOE values for both WT and I272A, suggesting similar internal motions. However, subtle differences were noticed when plotting deviating values onto the **bb'xa'** structure, shown in figure 5.31.

seen for WT and I272A **b'x** and **xa'c** in chapter 4, but this is possibly as a result of the poorer spectral quality and lower peak resolution due to the large size and line broadening of **b'xa'c**.

The estimated τ_m value for **b'xa'c**, both WT and I272A, at 40°C is 12.29 ns, only 1.0 ns higher than the experimental value of 11.30 ns for WT **b'xa'c** and 1.30 ns higher than the experimental τ_m of 11.00 ns for I272A **b'xa'c**. However, as the former is only an estimate value, there are bound to be discrepancies from the experimental value. It is also likely that this shift is caused by inter-domain motions. T_1 versus T_2 plots show the WT and I272A **b'xa'c** to have similar τ_m values suggesting that they tumble in solution at a comparable rate and therefore display similar sizes in the nanosecond timescale.

5.4.5.2 Model-free analysis

In order to further characterise the backbone dynamics of WT and I272A **b'xa'c**, model-free was carried out utilising all three relaxation parameters ^{15}N T_1 , T_2 and hetNOE. The I272A mutant had a higher average S^2 value than WT **b'xa'c** suggesting a more flexible backbone for the WT protein. I272A **b'x** also displayed a more rigid backbone than WT **b'x**, suggesting that the I272A mutation creates a more compact and rigid **b'** domain. The **x** linker region in the I272A mutant displayed a slightly more flexible backbone than **x** in the WT protein. This could be because the **x** linker of I272A needs to be more flexible in order to undergo the necessary conformational changes to form a more compact I272A **b'xa'c** structure. The **a'** domain was more flexible than the **b'** domain in both WT and mutant **b'xa'c**, highlighting the fact that **a'** is the more conformationally active domain in **b'xa'c**.

Model-free analysis highlighted areas of the WT and I272A **b'xa'c** proteins undergoing R_{ex} contributions. These regions are displayed on the structure of **bb'xa'** in figure 5.32.

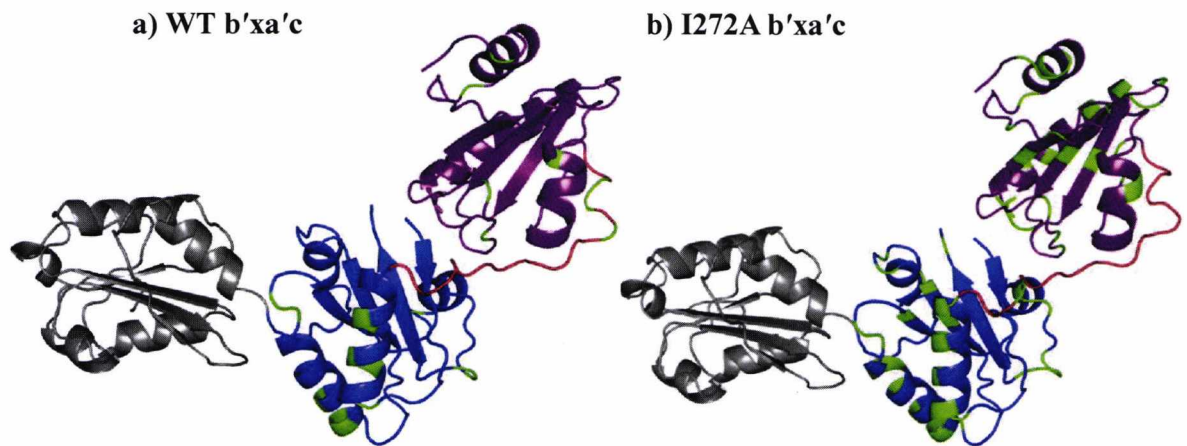


Figure 5.32 The crystal structure of **bb'xa'** (3UEM.pdb) showing regions in conformational exchange. Residues in green show contribution from R_{ex} from the model-free analysis of **a)** WT **b'xa'c** and **b)** I272A **b'xa'c**.

Similar to **b'x** data shown in chapter 4, the I272A mutant of **b'xa'c** shows a higher degree of R_{ex} contributions than the WT protein. This could be because the I272A mutation favours a more compact conformation due to capping of the ligand binding site. But, capping could prove more challenging in **b'xa'c** than it is in **b'x**, due to the presence of the **a'** domain. Also, residues of the **b'** β -sheet in the I272A mutant are highlighted by model-free as having R_{ex} contribution, which could be as a result of interactions with the **a'** domain while forming the compact structural module favoured by the I272A mutation.

CHAPTER 6

General Summary

Previous research into human PDI has shown the C-terminal fragment, **b'xa'c**, to be the minimal unit exhibiting isomerase activity (Pirneskoski et al., 2001). Most importantly, recent crystal structures of **bb'xa'** and **abb'xa'** have elucidated the critical role of the **a'** domain in the redox-dependent conformational changes that occur in PDI (Wang et al., 2012b, Wang et al., 2012a). The **b'x** fragment has also been shown to undergo significant conformational changes in which the **x** linker occludes, and consequently controls access of ligands to the hydrophobic ligand binding site on the **b'** domain, an event called “capping” (Nguyen et al., 2008). Therefore understanding the conformational changes that occur in **b'xa'c** could shed light on the structure/function relationship of full length PDI.

The main aim of this thesis was to characterize the **b'xa'c** fragment, and consequently to determine if capping of the **b'** ligand binding site by the **x** linker region occurred when **x** was tethered by both of its neighbouring domains **b'** and **a'**. Generation of the **b'xa'c** mutants I272A, L343A and 2DA (D346A/D348A) by site-directed mutagenesis was essential to this study as they have been shown to favour capped (I272A and 2DA) or uncapped (L343A) conformations of **b'x**.

Expression of **b'xa'c** in minimal media produced ~45 mg/L of protein, which consisted of a mixture of monomer and dimer species. Expression in rich media, such as LB, produced protein that was exclusively in the dimer form, possibly due to the high expression rate producing higher concentrations of protein. However, as sufficient quantities of monomeric protein could be purified away from the dimer form by gel filtration, protein expression was carried out in minimal medium. A number of other PDI fragments, such as **a'**, **a'c**, **xa'c** and WT and I272A **b'x** were also expressed to aid in the characterisation of **b'xa'c**. Several methods were used for the characterisation of the PDI constructs such as gel filtration, intrinsic fluorescence (including GdmCl denaturation) and NMR spectroscopy to follow conformational changes, ligand binding and backbone dynamics.

Initial biophysical characterisation by gel filtration showed that **b'xa'c** proteins presented with a hydrodynamic volume that was ~1.5 times larger than expected for globular proteins of similar sizes, thought to be due to the non-spherical conformation adopted by the PDI fragment. However, differences between the WT **b'xa'c** protein and the mutants were immediately obvious (figure 2.13). The I272A mutant presented with a smaller hydrodynamic volume than all of the **b'xa'c** proteins, followed by WT, L343A and lastly 2DA. This variation in hydrodynamic volume must be attributed to differences in the conformation of the proteins. So, I272A **b'xa'c** presented with a more compact conformation than the WT protein, supporting previous **b'x** data that shows this mutant to favour capping of the ligand binding site. L343A had a more “open” conformation, therefore supporting uncapping of the ligand binding site by the **x** linker region. However, the larger hydrodynamic volume of 2DA **b'xa'c** showed that this mutant must have a capping mechanism in **b'x** that is different in detail than that exhibited by the I272A mutant. This data was supported by intrinsic fluorescence which showed the I272A mutant to be more blue-shifted than WT, L343A and 2DA **b'xa'c**. 2DA and L343A appeared red-shifted in comparison with the WT protein, indicative of an uncapped conformation (figure 3.4).

Studies on the conformational stability of WT **b'xa'c** and the I272A, L343A and 2DA mutants in the presence of GdmCl revealed a biphasic transition, in which the first phase was attributed to the unfolding of the **a'** domain, and the second phase was due to the unfolding of **b'x**. This was confirmed by measuring the conformational stability of the

individual domains. The **a'** domain had a much lower conformational stability than **b'** as it appeared to be unfolding even at low denaturation concentrations. This made the analysis challenging, as calculations of $\Delta G^{(H_2O)}_U$ assumed that the **a'** domain was fully folded in the absence of denaturant. However, the **a'** domain started to unfold immediately on addition of GdmCl, consequently the denaturation curve lacked a plateau at the low denaturant concentrations. Therefore it is difficult to know if the **a'** domain was fully folded in the absence of denaturant. If indeed **a'** was partly unfolded, the unfolding of this domain did not fit a 2-state model and the fraction of the protein that was folded could not be accurately calculated. However, differences in stability between WT **b'xa'c** and I272A, L343A and 2DA mutants were obvious. The I272A mutation in the **b'** domain had a stabilising effect not only on **b'**, but also the **a'** domain, indicating that the domains have an effect on each other's conformational stability and that I272A has a similar stabilising role in **b'xa'c** as it did in **b'x**. Both L343A and 2DA presented with lower conformational stabilities than WT and I272A **b'xa'c**, suggesting the loss of some stabilising interactions.

Addition of DTT caused a blue-shift in fluorescence and increased the conformational stability of all of the proteins, suggesting a conformational change in the folded state of the reduced proteins making them more resistant to unfolding in the presence of a denaturant. This is in agreement with recently published data, which shows that reduction of the **a'** active site forms a more compact structural module in **bb'xa'** and **abb'xa'** (Wang et al., 2012b, Wang et al., 2012a). This finding was supported by proteolysis data, which showed that addition of DTT reduced the flexibility of the proteins rendering them less susceptible to proteolytic cleavage, in agreement with previously published data (Wang et al., 2010). Proteolysis also showed that the 2DA mutant was more susceptible to proteolytic cleavage, therefore it had an increased flexibility compared to the other **b'xa'c** proteins, suggesting that this mutant has a more open conformation, unlike the capped conformer of 2DA in **b'x**.

WT **b'xa'c** and its mutants appear to be conformationally active proteins displaying a high degree of NMR line broadening. NMR analysis also proved challenging due to the low conformational stability of the **a'** domain and sharp peaks due to the unstructured C-terminal extension, **c**. Reduction of the **a'** active site improved the quality of $^{15}\text{N}/^1\text{H}$ HSQC spectra and lowered the degree of line broadening but the quality of the spectra was

too poor for further NMR investigations until an increase in temperature and pH reduced line broadening further and so improved spectral quality.

Both NMR and denaturation studies showed an improvement in NMR spectra and conformational stability of the **a'** domain in the presence of the **x** linker region. $^{15}\text{N}/^1\text{H}$ HSQC spectra of the **a'c** fragment in isolation consisted of poorly resolved peaks, some of which were line broadened beyond detection. Addition of the **x** linker region in the **xa'c** construct resulted not only in increased conformational stability but also in better resolved and less line broadened NMR spectra, suggesting that the **x** linker interacts with the **a'** domain to make it more conformationally stable. This was also evident in the hetNOE data of **xa'c**, which showed that **x** behaved as a structured part of the protein.

Characterisation of WT and I272A **b'x** as well as **xa'c** proved valuable in investigations of **b'xa'c**. Backbone assignments of I272A **b'xa'c** could only be achieved using assignments from WT **b'x** and **xa'c** in combination with triple resonance experiments. NMR data for **b'xa'c** supported the denaturation studies in confirming the effect of the mutants on **b'xa'c**. As L343A and 2DA had lower conformational stabilities, especially in the **a'** domain, they produced NMR spectra that could not be used for further investigations.

Backbone relaxation studies showed that I272A **b'xa'c** had a more rigid backbone than WT **b'xa'c**, but the mutant had a greater number of residues in conformational exchange than the WT protein. This was in agreement with **b'x** data which showed the I272A mutant to have contribution from R_{ex} for the majority of its residues. Therefore, the I272A mutation, which in **b'x** favours capping of the ligand binding site, creates a protein that is more compact but conformationally active. **b'x** and **xa'c** had a more rigid backbone than the **b'x** and **a'c** halves of **b'xa'c** showing that **b'xa'c** is a flexible protein, in support of published proteolysis data (Wang et al., 2010).

T_1 vs T_2 plots showed the WT and I272A **b'xa'c** to have similar τ_m values, suggesting that they tumble in solution at a comparable rate. However, gel filtration showed the I272A mutant to be 3 kDa lighter in molecular weight than WT **b'xa'c**, indicating a more compact conformation. This suggests that WT **b'xa'c** undergoes slow inter-domain motions on the second to minute timescale which can be detected by gel filtration, giving the WT protein a larger hydrodynamic volume than the I272A mutant. A

number of experiments could be used to further investigate this slow inter-domain motion such as multi-angle light scattering (MALS), hydrogen exchange (HX) and NMR CPMG experiments which look at molecular motions on a longer timescale than backbone relaxation dynamics (Arifin and Palmer, 2003, Wales and Engen, 2006, Mittermaier and Kay, 2009). In **b'x**, I272A mutant experienced motions on a faster timescale than WT **b'x** which was detected by relaxation experiments and shown by its lower global correlation time, suggesting that I272A **b'x** formed a more compact conformation than the WT protein due to capping of the ligand binding site. This difference in size was not identified by gel filtration due to the longer detection timescale of this method. This variation in molecular motion seen in **b'xa'c**, when compared to **b'x**, is due to the presence of the **a'** domain which is affecting the behaviour of **b'x**. This is further evidence of the influence that PDI domains have on each other, suggesting that in order for PDI function, its domains must be able to “communicate” with each other.

So, a key question in this thesis was : Does capping occur in **b'xa'c**?

The I272A mutant had a smaller hydrodynamic volume than WT **b'xa'c** and the other two mutants. It also showed a blue-shift in fluorescence compared to the WT protein. Its conformational stability was higher than that of the WT and as a consequence $^{15}\text{N}/^1\text{H}$ HSQC spectra were less line broadened. The indole NH of W347 presented as two peaks in NMR spectra, representing different environments for the side chain of this tryptophan residue. This mutant also had a more rigid backbone than WT **b'xa'c** as shown by the relaxation data. All of this data supports a more compact and possibly capped conformation. However, ligand binding data was very similar to that of the WT protein. In both cases, the chemical shift perturbations measured were very small compared to published data of **b'x** in the presence of varying concentrations of Δ -somatostatin (Byrne et al., 2009). Most importantly, if Δ -somatostatin was binding to I272A **b'xa'c**, and if this mutant was in favour of the capped conformation, there should have been an upfield shift in the downfield indole NH of W347 as a result of **x** being displaced from the ligand binding site by Δ -somatostatin. This shift was not seen in $^{15}\text{N}/^1\text{H}$ HSQC spectra of I272A in the presence of peptide ligand. This could have been because **x** was not capping the ligand binding site in the first place and therefore its chemical environment would not change upon ligand binding, or **x** was capping the ligand binding site but Δ -somatostatin was not binding and therefore **x** could not be displaced.

Also, 2DA in **b'x** is a stabilising mutant favouring capping of the ligand binding site by the **x** linker region. However, this mutation had the opposite effect in **b'xa'c** as it had a larger hydrodynamic volume than WT, I272A and even L343A. Its conformational stability, especially that of the **a'** domain, was lower than that of the other **b'xa'c** proteins and its high susceptibility to proteolytic cleavage showed it to be a very flexible protein. It also produced $^{15}\text{N}/^1\text{H}$ spectra with poorly resolved and line broadened peaks and showed a red-shift in fluorescence compared not only to the WT but also the L343A mutant. All of these features are not to be expected from a mutant that is supposed to favour capping of the ligand site. But, its W347 indole NH presented as two peaks, indicative of two conformations for W347 in the **x** linker. However, as splitting of the W347 indole peak is the only evidence to suggest that W347 in 2DA is in at least two different environments, it does not necessarily mean it is capping the ligand binding site. This suggests that even though the two peaks may be seen for the W347 indole, it does not mean that they represent the capped and uncapped conformations.

The L343A mutant does support an uncapped conformation as it appeared with a larger hydrodynamic volume on gel filtration than WT **b'xa'c**. It displayed a red-shift in fluorescence and its conformational stability was lower than that of the WT protein. NMR spectra of L343A were poorly resolved and very line broadened. Also, there was no evidence of W347 being in more than one state as there was no splitting of the indole peak.

In summary, data shows that the I272A mutant had a positive effect in the conformational stability of **b'xa'c**, whereas the 2DA and L343A mutations in the **x** linker region had a negative effect not only on the **b'** domain but also **a'**. Although, I272A has stabilised **b'xa'c** through a more compact conformation, there is no conclusive evidence to prove that **x** is capping the ligand binding site in this mutant.

E. References

ABRAGAM, A. 1961. *Principles of Nuclear Magnetism*, Oxford, OUP.

ALANEN, H. I., WILLIAMSON, R. A., HOWARD, M. J., HATAHET, F. S., SALO, K. E., KAUPPILA, A., KELLOKUMPU, S. & RUDDOCK, L. W. 2006. ERp27, a new non-catalytic endoplasmic reticulum-located human protein disulfide isomerase family member, interacts with ERp57. *J Biol Chem*, 281, 33727-38.

ALANEN, H. I., WILLIAMSON, R. A., HOWARD, M. J., LAPPI, A. K., JANTTI, H. P., RAUTIO, S. M., KELLOKUMPU, S. & RUDDOCK, L. W. 2003. Functional characterization of ERp18, a new endoplasmic reticulum-located thioredoxin superfamily member. *J Biol Chem*, 278, 28912-20.

ANDRE, P., KIM, A., KHRAPKO, K. & THILLY, W. G. 1997. Fidelity and mutational spectrum of Pfu DNA polymerase on a human mitochondrial DNA sequence. *Genome Research*, 7, 843-852.

ANELLI, T., ALESSIO, M., MEZGHRANI, A., SIMMEN, T., TALAMO, F., BACHI, A. & SITIA, R. 2002. ERp44, a novel endoplasmic reticulum folding assistant of the thioredoxin family. *EMBO J*, 21, 835-44.

ARIFIN, D. R. & PALMER, A. F. 2003. Determination of size distribution and encapsulation efficiency of liposome-encapsulated hemoglobin blood substitutes using asymmetric flow field-flow fractionation coupled with multi-angle static light scattering. *Biotechnol Prog*, 19, 1798-811.

BARAK, N. N., NEUMANN, P., SEVVANA, M., SCHUTKOWSKI, M., NAUMANN, K., MALESEVIC, M., REICHARDT, H., FISCHER, G., STUBBS, M. T. & FERRARI, D. M. 2009. Crystal structure and functional analysis of the protein disulfide isomerase-related protein ERp29. *J Mol Biol*, 385, 1630-42.

BARBATO, G., IKURA, M., KAY, L. E., PASTOR, R. W. & BAX, A. 1992. Backbone dynamics of calmodulin studied by nitrogen-15 relaxation using inverse detected two-dimensional NMR spectroscopy: the central helix is flexible. *Biochemistry*, 31, 5269-5278.

BROADHURST, R. W., HARDMAN, C. H., THOMAS, J. O. & LAUE, E. D. 1995. Backbone dynamics of the A-domain of HMG1 as studied by ¹⁵N NMR Spectroscopy. *Biochemistry*, 34, 16608-16617.

BYRNE, L. J., SIDHU, A., WALLIS, A. K., RUDDOCK, L. W., FREEDMAN, R. B., HOWARD, M. J. & WILLIAMSON, R. A. 2009. Mapping of the ligand-binding site on the b' domain of human PDI: interaction with peptide ligands and the x-linker region. *Biochem J*, 423, 209-17.

CAI, H., WANG, C. C. & TSOU, C. L. 1994. Chaperone-like activity of protein disulfide isomerase in the refolding of a protein with no disulfide bonds. *J Biol Chem*, 269, 24550-2.

CAVANAGH, J., FAIRBROTHER, W., PALMER, A. G., RANCE, M. & SKELTON, N. 2007. *Protein NMR Spectroscopy: Principles and Practice*, Academic Press.

CHALTON, D. A. & LAKEY, J. H. 2010. Simple detection of protein soft structure changes. *Anal Chem*, 82, 3073-3076.

CHEN, Y., ZHANG, Y., YIN, Y., GAO, G., LI, S., JIANG, Y., GU, X. & LUO, J. 2005. SPD—a web-based secreted protein database. *Nucleic Acids Res*, 33, D169 - D173.

CHEUNG, M. S., MAGUIRE, M. L., STEVENS, T. J. & BROADHURST, R. W. 2010. DANGLE: A Bayesian inferential method for predicting protein backbone dihedral angles and secondary structure. *J Magn Reson*, 202, 223-33.

- CORNILESCU, G., DELAGLIO, F. & BAX, A. 1999. Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *J Biomol NMR*, 13, 289-302.
- CREIGHTON, T. E., HILLSON, D. A. & FREEDMAN, R. B. 1980. Catalysis by protein-disulphide isomerase of the unfolding and refolding of proteins with disulphide bonds. *J Mol Biol*, 142, 43-62.
- CUNNEA, P. M., MIRANDA-VIZUETE, A., BERTOLI, G., SIMMEN, T., DAMDIMOPOULOS, A. E., HERMANN, S., LEINONEN, S., HUIKKO, M. P., GUSTAFSSON, J. A., SITIA, R. & SPYROU, G. 2003. ERdj5, an endoplasmic reticulum (ER)-resident protein containing DnaJ and thioredoxin domains, is expressed in secretory cells or following ER stress. *J Biol Chem*, 278, 1059-66.
- DAI, Y. & WANG, C. C. 1997. A mutant truncated protein disulfide isomerase with no chaperone activity. *J Biol Chem*, 272, 27572-27576.
- DARAGAN, V. A. & MAYO, K. H. 1997. Motional model analyses of protein and peptide dynamics using ^{13}C and ^{15}N NMR relaxation. *Prog Nucl Magn Reson Spec*, 31, 63-105.
- DARBY, N. J. & CREIGHTON, T. E. 1995. Characterization of the active site cysteine residues of the thioredoxin-like domains of protein disulfide isomerase. *Biochemistry*, 34, 16770-80.
- DARBY, N. J., PENKA, E. & VINCENTELLI, R. 1998. The multi-domain structure of protein disulfide isomerase is essential for high catalytic efficiency. *J Mol Biol*, 276, 239-247.
- DENISOV, A. Y., MÄÄTTÄNEN, P., DABROWSKI, C., KOZLOV, G., THOMAS, D. Y. & GEHRING, K. 2009. Solution structure of the bb' domains of human protein disulfide isomerase. *FEBS J*, 276, 1440-1449.
- DIJKSTRA, K., KARVONEN, P., PIRNESKOSKI, A., KOIVUNEN, P., KIVIRIKKO, K. I., DARBY, N. J., VAN STRAATEN, M., SCHEEK, R. M. & KEMMINK, J. 1999. Letter to the Editor: Assignment of ^1H , ^{13}C and ^{15}N resonances of the a' domain of protein disulfide isomerase. *J Biomol NMR*, 14, 195-196.
- DOBSON, C. M. 2003. Protein folding and misfolding. *Nature*, 426, 884-90.
- EDMAN, J. C., ELLIS, L., BLACHER, R. W., ROTH, R. A. & RUTTER, W. J. 1985. Sequence of protein disulphide isomerase and implications of its relationship to thioredoxin. *Nature*, 317, 267-70.
- ELLGAARD, L. & RUDDOCK, L. W. 2005. The human protein disulphide isomerase family: substrate interactions and functional properties. *EMBO Reports*, 6, 28-32.
- ESSEX, D. W. 2009. Redox control of platelet function. *Antiox Red Signal*, 11, 1191-225.
- FARROW, N., ZHANG, O., SZABO, A., TORCHIA, D. & KAY, L. 1995. Spectral density function mapping using ^{15}N relaxation data exclusively. *J Biomol NMR*, 6, 153-162.
- FENOUILLET, E., BARBOUCHE, R. & JONES, I. M. 2007. Cell entry by enveloped viruses: redox considerations for HIV and SARS-coronavirus. *Antiox Red Signal*, 9, 1009-34.
- FERRARI, D. M. & SÖLING, H. D. 1999. The protein disulphide-isomerase family: unravelling a string of folds. *Biochem J*, 339 (Pt 1), 1-10.
- FRANCOIS, B. 1999. Recombinant protein expression in *Escherichia coli*. *Curr Opin Biotech*, 10, 411-421.
- FRAND, A. R. & KAISER, C. A. 1998. The ERO1 gene of yeast is required for oxidation of protein dithiols in the endoplasmic reticulum. *Mol Cell*, 1, 161-170.
- FREEDMAN, R. B. 1989. Protein disulfide isomerase: multiple roles in the modification of nascent secretory proteins. *Cell*, 57, 1069-72.

- FREEDMAN, R. B., HIRST, T. R. & TUIITE, M. F. 1994. Protein disulphide isomerase: building bridges in protein folding. *Trends Biochem Sci*, 19, 331-6.
- FRICKEL, E. M., FREI, P., BOUVIER, M., STAFFORD, W. F., HELENIUS, A., GLOCKSHUBER, R. & ELLGAARD, L. 2004. ERp57 is a multifunctional thiol-disulfide oxidoreductase. *J Biol Chem*, 279, 18277-87.
- GARBI, N., TANAKA, S., MOMBURG, F. & HAMMERLING, G. J. 2006. Impaired assembly of the major histocompatibility complex class I peptide-loading complex in mice deficient in the oxidoreductase ERp57. *Nat Immunol*, 7, 93-102.
- GEOGHEGAN, K. F., DIXON, H. B. F., ROSNER, P. J., HOTH, L. R., LANZETTI, A. J., BORZILLERI, K. A., MARR, E. S., PEZZULLO, L. H., MARTIN, L. B., LEMOTTE, P. K., MCCOLL, A. S., KAMATH, A. V. & STROH, J. G. 1999. Spontaneous [alpha]-N-6-Phosphogluconoylation of a "His Tag" in *Escherichia coli*: The Cause of Extra Mass of 258 or 178 Da in Fusion Proteins. *Anal Biochem*, 267, 169-184.
- GESS, B., HOFBAUER, K. H., WENGER, R. H., LOHAUS, C., MEYER, H. E. & KURTZ, A. 2003. The cellular oxygen tension regulates expression of the endoplasmic oxidoreductase ERO1-Lalpha. *Eur J Biochem*, 270, 2228-35.
- GOLDBERGER, R. F., EPSTEIN, C. J. & ANFINSEN, C. B. 1963. Acceleration of reactivation of reduced bovine pancreatic ribonuclease by a microsomal system from rat liver. *J Biol Chem*, 238, 628-35.
- GRODBERG, J. & DUNN, J. J. 1988. ompT encodes the *Escherichia coli* outer membrane protease that cleaves T7 RNA polymerase during purification. *J Bacter*, 170, 1245-1253.
- HATAHET, F. & RUDDOCK, L. W. 2009. Protein disulfide isomerase: a critical evaluation of its function in disulfide bond formation. *Antiox Red Signal*, 11, 2807-50.
- HECK, A. J. R. 2008. Native mass spectrometry: a bridge between interactomics and structural biology. *Nature Methods*, 5, 927-933.
- HIGUCHI, R., KRUMMEL, B. & SAIKI, R. K. 1988. A general-method of invitro preparation and specific mutagenesis of DNA fragments - study of protein and DNA interactions. *Nucl Acid Res*, 16, 7351-7367.
- HINIKER, A. & BARDWELL, J. C. 2004. Disulfide relays between and within proteins: the Ero1p structure. *Trends Biochem Sci*, 29, 516-9.
- HO, S. N., HUNT, H. D., HORTON, R. M., PULLEN, J. K. & PEASE, L. R. 1989. Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene*, 77, 51-59.
- HWANG, C., SINSKEY, A. J. & LODISH, H. F. 1992. Oxidized redox state of glutathione in the endoplasmic reticulum. *Science*, 257, 1496-1502.
- IBARRA-MOLERO, B. & SANCHEZ-RUIZ, J. M. 1997. Are there equilibrium intermediate states in the urea-induced unfolding of hen egg-white lysozyme? *Biochemistry*, 36, 9616-24.
- INOUE, M., ARNHEIM, N. & STERNGLANZ, R. 1973. Bacteriophage T7 lysozyme is an N-acetylmuramyl-l-alanine amidase. *J Biol Chem*, 248, 7247-7252.
- KAY, L. E. 2005. NMR studies of protein structure and dynamics. *J Magn Reson*, 173, 193-207.
- KEELER, J. 2005. *Understanding NMR Spectroscopy*, Wiley.
- KEMMINK, J., DARBY, N. J., DIJKSTRA, K., NILGES, M. & CREIGHTON, T. E. 1996. Structure determination of the N-terminal thioredoxin-like domain of protein disulfide isomerase using multidimensional heteronuclear ¹³C/¹⁵N NMR spectroscopy. *Biochemistry*, 35, 7684-91.

- KEMMINK, J., DIJKSTRA, K., MARIANI, M., SCHEEK, R. M., PENKA, E., NILGES, M. & DARBY, N. J. 1999. The structure in solution of the b domain of protein disulfide isomerase. *J Biomol NMR*, 13, 357-68.
- KLAPPA, P., FREEDMAN, R. B. & ZIMMERMANN, R. 1995. Protein disulphide isomerase and a luminal cyclophilin-type peptidyl prolyl cis-trans isomerase are in transient contact with secretory proteins during late stages of translocation. *Eur J Biochem*, 232, 755-64.
- KLAPPA, P., KOIVUNEN, P., PIRNESKOSKI, A., KARVONEN, P., RUDDOCK, L. W., KIVIRIKKO, K. I. & FREEDMAN, R. B. 2000. Mutations that destabilize the a' domain of human protein-disulfide isomerase indirectly affect peptide binding. *J Biol Chem*, 275, 13213-13218.
- KLAPPA, P., RUDDOCK, L. W., DARBY, N. J. & FREEDMAN, R. B. 1998. The b' domain provides the principal peptide-binding site of protein disulfide isomerase but all domains contribute to binding of misfolded proteins. *EMBO J*, 17, 927-935.
- KNOBLACH, B., KELLER, B. O., GROENENDYK, J., ALDRED, S., ZHENG, J., LEMIRE, B. D., LI, L. & MICHALAK, M. 2003. ERp19 and ERp46, new members of the thioredoxin family of endoplasmic reticulum proteins. *Mol Cell Proteomics*, 2, 1104-19.
- KOIVU, J., MYLLYLÄ, R., HELAAKOSKI, T., PIHLAJANIEMI, T., TASANEN, K. & KIVIRIKKO, K. I. 1987. A single polypeptide acts both as the beta-subunit of Prolyl 4-Hydroxylase and as a protein disulfide isomerase. *J Biol Chem*, 262, 6447-6449.
- KOIVUNEN, P., SALO, K. E., MYLLYHARJU, J. & RUDDOCK, L. W. 2005. Three binding sites in protein-disulfide isomerase cooperate in collagen prolyl 4-hydroxylase tetramer assembly. *J Biol Chem*, 280, 5227-35.
- KORTEMME, T. & CREIGHTON, T. E. 1995. Ionisation of cysteine residues at the termini of model alpha-helical peptides. Relevance to unusual thiol pKa values in proteins of the thioredoxin family. *J Mol Biol*, 253, 799-812.
- KOZLOV, G., MAATTANEN, P., SCHRAG, J. D., POLLOCK, S., CYGLER, M., NAGAR, B., THOMAS, D. Y. & GEHRING, K. 2006. Crystal structure of the bb' domains of the protein disulfide isomerase ERp57. *Structure*, 14, 1331-9.
- KURTIN, W. E. & LEE, J. M. 2002. The free energy of denaturation of lysozyme - An undergraduate experiment in biophysical chemistry. *Biochem Mol Biol Education*, 30, 244-247.
- LAEMMLI, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, 227, 680-685.
- LAPPI, A. K., LENSINK, M. F., ALANEN, H. I., SALO, K. E., LOBELL, M., JUFFER, A. H. & RUDDOCK, L. W. 2004. A conserved arginine plays a role in the catalytic cycle of the protein disulphide isomerases. *J Mol Biol*, 335, 283-95.
- LI, S. J., HONG, X. G., SHI, Y. Y., LI, H. & WANG, C. C. 2006. Annular arrangement and collaborative actions of four domains of protein-disulfide isomerase: a small angle X-ray scattering study in solution. *J Biol Chem*, 281, 6581-8.
- LIPARI, G. & SZABO, A. 1982a. Model-Free Approach to the Interpretation of Nuclear Magnetic-Resonance Relaxation in Macromolecules .1. Theory and Range of Validity. *J Am Chem Soc*, 104, 4546-4559.
- LIPARI, G. & SZABO, A. 1982b. Model-Free Approach to the Interpretation of Nuclear Magnetic-Resonance Relaxation in Macromolecules .2. Analysis of Experimental Results. *J Am Chem Soc*, 104, 4559-4570.

- LUNDSTRÖM, J. & HOLMGREN, A. 1990. Protein disulfide-isomerase is a substrate for thioredoxin reductase and has thioredoxin-like activity. *J Biol Chem*, 265, 9114-9120.
- MANDEL, A. M., AKKE, M. & PALMER, A. G. 1995. Backbone dynamics of *Escherichia coli* Ribonuclease HI: Correlations with structure and function in an active enzyme. *J Mol Biol*, 246, 144-163.
- MARCINIAK, S. J., YUN, C. Y., OYADOMARI, S., NOVOA, I., ZHANG, Y., JUNGREIS, R., NAGATA, K., HARDING, H. P. & RON, D. 2004. CHOP induces death by promoting protein synthesis and oxidation in the stressed endoplasmic reticulum. *Genes Dev*, 18, 3066-77.
- MAZZARELLA, R. A., SRINIVASAN, M., HAUGEJORDEN, S. M. & GREEN, M. 1990. ERp72, an abundant luminal endoplasmic reticulum protein, contains three copies of the active site sequences of protein disulfide isomerase. *J Biol Chem*, 265, 1094-101.
- MCCONNELL, H. M. 1958. Reaction rates by Nuclear Magnetic Resonance. *J Chem Phys*, 28, 431-431.
- MITTERMAIER, A. K. & KAY, L. E. 2009. Observing biological dynamics at atomic resolution using NMR. *Trends Biochem Sci*, 34, 601-611.
- MKRTCHIAN, S., FANG, C., HELLMAN, U. & INGELMAN-SUNDBERG, M. 1998. A stress-inducible rat liver endoplasmic reticulum protein, ERp29. *Eur J Biochem*, 251, 304-13.
- MYERS, J. K., PACE, C. N. & SCHOLTZ, J. M. 1995. Denaturant m values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding. *Protein Sci*, 4, 2138-48.
- NGUYEN, V. D., WALLIS, K., HOWARD, M. J., HAAPALAINEN, A. M., SALO, K. E., SAARANEN, M. J., SIDHU, A., WIERENGA, R. K., FREEDMAN, R. B., RUDDOCK, L. W. & WILLIAMSON, R. A. 2008. Alternative conformations of the x region of human protein disulphide-isomerase modulate exposure of the substrate binding b' domain. *J Mol Biol*, 383, 1144-55.
- NOIVA, R. 1999. Protein disulfide isomerase: the multifunctional redox chaperone of the endoplasmic reticulum. *Semin Cell Dev Biol*, 10, 481-93.
- OHMAN, A., OMAN, T. & OLIVEBERG, M. 2010. Solution structures and backbone dynamics of the ribosomal protein S6 and its permutant P(54-55). *Protein Sci*, 19, 183-9.
- OKUBARA, P. A., TIBBOT, B. K., TARUN, A. S., MCALPIN, C. E. & HUA, S. S. 2003. Partial retrotransposon-like DNA sequence in the genomic clone of *Aspergillus flavus*, pAF28. *Mycol Res*, 107, 841-6.
- OLIVER, J. D., RODERICK, H. L., LLEWELLYN, D. H. & HIGH, S. 1999. ERp57 functions as a subunit of specific complexes formed with the ER lectins calreticulin and calnexin. *Mol Biol Cell*, 10, 2573-2582.
- PAGANI, M., FABBRI, M., BENEDETTI, C., FASSIO, A., PILATI, S., BULLEID, N. J., CABIBBO, A. & SITIA, R. 2000. Endoplasmic reticulum oxidoreductin 1-lbeta (ERO1-Lbeta), a human gene induced in the course of the unfolded protein response. *J Biol Chem*, 275, 23685-92.
- PEAPER, D. R. & CRESSWELL, P. 2008. Regulation of MHC class I assembly and peptide binding. *Annu Rev Cell Dev Biol*, 24, 343-68.
- PALMER, A. G. 1997. Probing molecular motion by NMR. *Curr Opin Struct Biol*, 7, 732-7.
- PALMER, A. G. 2001. NMR probes of molecular dynamics: overview and comparison with other techniques. *Annu Rev Biophys Biomol Struct*, 30, 129-55.
- PALMER, A. G., RANCE, M. & WRIGHT, P. E. 1991. Intramolecular motions of a zinc finger DNA-binding domain from Xfin characterized by proton-detected natural abundance carbon-13 heteronuclear NMR spectroscopy. *J Am Chem Soc*, 113, 4371-4380.

- PELHAM, H. R. B. 1990. The retention signal for soluble proteins of the endoplasmic reticulum. *Trends Biochem Sci*, 15, 483-486.
- PIOTTO, M., SAUDEK, V. & SKLENÁŘ, V. 1992. Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *J Biomol NMR*, 2, 661-665.
- PIRNESKOSKI, A., KLAPPA, P., LOBELL, M., WILLIAMSON, R. A., BYRNE, L., ALANEN, H. I., SALO, K. E., KIVIRIKKO, K. I., FREEDMAN, R. B. & RUDDOCK, L. W. 2004. Molecular characterization of the principal substrate binding site of the ubiquitous folding catalyst protein disulfide isomerase. *J Biol Chem*, 279, 10374-81.
- PIRNESKOSKI, A., RUDDOCK, L. W., KLAPPA, P., FREEDMAN, R. B., KIVIRIKKO, K. I. & KOIVUNEN, P. 2001. Domains b' and a' of protein disulfide isomerase fulfill the minimum requirement for function as a subunit of prolyl 4-hydroxylase. The N-terminal domains a and b enhances this function and can be substituted in part by those of ERp57. *J Biol Chem*, 276, 11287-93.
- PRICE, N. C. & NAIRN, J. 2009. *Exploring Proteins: a student's guide to experimental skills and methods*, Oxford University Press.
- PUIG, A. & GILBERT, H. F. 1994. Protein disulfide isomerase exhibits chaperone and anti-chaperone activity in the oxidative refolding of lysozyme. *Journal of Biological Chemistry*, 269, 7764-7771.
- RAMACHANDRAN, G. N., RAMAKRISHNAN, C. & SASISEKHARAN, V. 1963. Stereochemistry of polypeptide chain configurations. *J Mol Biol*, 7, 95-9.
- ROWE, M. L., RUDDOCK, L. W., KELLY, G., SCHMIDT, J. M., WILLIAMSON, R. A. & HOWARD, M. J. 2009. Solution structure and dynamics of ERp18, a small endoplasmic reticulum resident oxidoreductase. *Biochemistry*, 48, 4596-606.
- RULE, G. S. & HITCHENS, T. K. 2006. *Fundamentals of Protein NMR Spectroscopy*, Springer.
- SAMBROOK, J. & RUSSELL, D. W. 2001. *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press.
- SARGSYAN, E., BARYSHEV, M., SZEKELY, L., SHARIPO, A. & MKRTCHIAN, S. 2002. Identification of ERp29, an endoplasmic reticulum luminal protein, as a new member of the thyroglobulin folding complex. *J Biol Chem*, 277, 17009-15.
- SCHELLMAN, J. A. 2002. Fifty years of solvent denaturation. *Biophys Chem*, 96, 91-101.
- SENANAYAKE, S. D., BRIAN, D. A. 1995. Precise large deletions by the PCR-based overlap extension method. *Mol Biotechnol*, 4, 13-15.
- SERVE, O., KAMIYA, Y., MAENO, A., NAKANO, M., MURAKAMI, C., SASAKAWA, H., YAMAGUCHI, Y., HARADA, T., KURIMOTO, E., YAGI-UTSUMI, M., IGUCHI, T., INABA, K., KIKUCHI, J., ASAMI, O., KAJINO, T., OKA, T., NAKASAKO, M. & KATO, K. 2010. Redox-dependent domain rearrangement of protein disulfide isomerase coupled with exposure of its substrate-binding hydrophobic surface. *J Mol Biol*, 396, 361-74.
- SONG, J. L. & WANG, C. C. 1995. Chaperone-like activity of protein disulfide-isomerase in the refolding of rhodanese. *Eur J Biochem*, 231, 312-6.
- STUDIER, F. W. & MOFFATT, B. A. 1986. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J Mol Biol*, 189, 113-130.
- SULLIVAN, D. C., HUMINIECKI, L., MOORE, J. W., BOYLE, J. J., POULSOM, R., CREAMER, D., BARKER, J. & BICKNELL, R. 2003. EndoPDI, a novel protein-disulfide isomerase-like protein that is preferentially expressed in endothelial cells acts as a stress survival factor. *J Biol Chem*, 278, 47079-88.

- SUN, X. X., DAI, Y., LIU, H. P., CHEN, S. M. & WANG, C. C. 2000. Contributions of protein disulfide isomerase domains to its chaperone activity. *Biochim Biophys Acta*, 1481, 45-54.
- TAO, H., BAUSCH, C., RICHMOND, C., BLATTNER, F. R. & CONWAY, T. 1999. Functional Genomics: Expression analysis of *Escherichia coli* growing on minimal and rich media. *J Bacter*, 181, 6425-6440.
- TAVENDER, T. J. & BULLEID, N. J. 2010. Molecular Mechanisms Regulating Oxidative Activity of the Ero1 Family in the Endoplasmic Reticulum. *Antiox Red Signal*, 13, 1177-1187.
- TIAN, G., KOBER, F. X., LEWANDROWSKI, U., SICKMANN, A., LENNARZ, W. J. & SCHINDELIN, H. 2008. The catalytic activity of protein-disulfide isomerase requires a conformationally flexible molecule. *J Biol Chem*, 283, 33630-40.
- TIAN, G., XIANG, S., NOIVA, R., LENNARZ, W. J. & SCHINDELIN, H. 2006. The crystal structure of yeast protein disulfide isomerase suggests cooperativity between its active sites. *Cell*, 124, 61-73.
- TSAI, B. & RAPOPORT, T. A. 2002. Unfolded cholera toxin is transferred to the ER membrane and released from protein disulfide isomerase upon oxidation by Ero1. *J Cell Biol*, 159, 207-16.
- UEHARA, T., NAKAMURA, T., YAO, D., SHI, Z. Q., GU, Z., MA, Y., MASLIAH, E., NOMURA, Y. & LIPTON, S. A. 2006. S-nitrosylated protein-disulphide isomerase links protein misfolding to neurodegeneration. *Nature*, 441, 513-7.
- VAN LITH, M., KARALA, A. R., BROWN, D., GATEHOUSE, J. A., RUDDOCK, L. W., SAUNDERS, P. T. K. & BENHAM, A. M. 2007. A developmentally regulated chaperone complex for the endoplasmic reticulum of male haploid germ cells. *Mol Biol Cell*, 18, 2798-2804.
- VANZI, F., MADAN, B. & SHARP, K. 1998. Effect of the protein denaturants urea and guanidinium on water structure: A structural and thermodynamic study. *J Am Chem Soc*, 120, 10748-10753.
- VIVIAN, J. T. & CALLIS, P. R. 2001. Mechanisms of tryptophan fluorescence shifts in proteins. *Biophys J*, 80, 2093-109.
- VRANKEN, W. F., BOUCHER, W., STEVENS, T. J., FOGH, R. H., PAJON, A., LLINAS, M., ULRICH, E. L., MARKLEY, J. L., IONIDES, J. & LAUE, E. D. 2005. The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins*, 59, 687-96.
- WALES, T. E. & ENGEN, J. R. 2006. Hydrogen exchange mass spectrometry for the analysis of protein dynamics. *Mass Spectrom Rev*, 25, 158-70.
- WALLIS, A. K., SIDHU, A., BYRNE, L. J., HOWARD, M. J., RUDDOCK, L. W., WILLIAMSON, R. A. & FREEDMAN, R. B. 2009. The ligand-binding b' domain of human protein disulphide-isomerase mediates homodimerization. *Protein Sci*, 18, 2569-77.
- WANG, C., CHEN, S., WANG, X., WANG, L., WALLIS, A. K., FREEDMAN, R. B. & WANG, C. C. 2010. Plasticity of human protein disulfide isomerase: evidence for mobility around the x-linker region and its functional significance. *J Biol Chem*, 285, 26788-97.
- WANG, C., LI, W., REN, J., FANG, J., KE, H., GONG, W., FENG, W. & WANG, C. 2012a Structural insights into the redox-regulated dynamic conformations of human protein disulfide isomerase *Antiox Red Signal*, In Press.
- WANG, C., YU, J., HUO, L., WANG, L., FENG, W. & WANG, C. C. 2012b. Human protein-disulfide isomerase is a redox-regulated chaperone activated by oxidation of domain a'. *J Biol Chem*, 287, 1139-49.
- WANG, L. K., WANG, L., VAVASSORI, S., LI, S., KE, H. M., ANELLI, T., DEGANO, M., RONZONI, R., SITIA, R., SUN, F. & WANG, C. C. 2008. Crystal structure of human ERp44 shows a dynamic functional modulation by its carboxy-terminal tail. *EMBO Reports*, 9, 642-647.

- WANG, Z. V., SCHRAW, T. D., KIM, J. Y., KHAN, T., RAJALA, M. W., FOLLENZI, A. & SCHERER, P. E. 2007. Secretion of the adipocyte-specific secretory protein adiponectin critically depends on thiol-mediated protein retention. *Mol Cell Biol*, 27, 3716-31.
- WETTERAU, J. R., COMBS, K. A., MCLEAN, L. R., SPINNER, S. N. & AGGERBECK, L. P. 1991. Protein disulfide isomerase appears necessary to maintain the catalytically active structure of the microsomal triglyceride transfer protein. *Biochemistry*, 30, 9728-9735.
- WILSON, R., LEES, J. F. & BULLEID, N. J. 1998. Protein disulfide isomerase acts as a molecular chaperone during the assembly of procollagen. *J Biol Chem*, 273, 9637-9643.
- WINTER, J., KLAPPA, P., FREEDMAN, R. B., LILIE, H. & RUDOLPH, R. 2002. Catalytic activity and chaperone function of human protein-disulfide isomerase are required for the efficient refolding of proinsulin. *J Biol Chem*, 277, 310-7.
- WISHART, D. S., BIGAM, C. G., HOLM, A., HODGES, R. S. & SYKES, B. D. 1995. ¹H, ¹³C and ¹⁵N random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects. *J Biomol NMR*, 5, 67-81.
- WISHART, D. S. & SYKES, B. D. 1994. [12] Chemical shifts as a tool for structure determination. *Method Enzymol*, Volume 239, 363-392.
- WÜTHERICH, K. 1986. *NMR of Proteins and Nucleic Acids*, Wiley.
- YAO, Y., ZHOU, Y. & WANG, C. 1997. Both the isomerase and chaperone activities of protein disulfide isomerase are required for the reactivation of reduced and denatured acidic phospholipase A2. *EMBO J*, 16, 651-658.

F. Appendix Contents

Chapter 2

- 2.1 DNA sequencing of **b'xa'c** mutants

Chapter 4

- 4.1 Minimal shift map values of **xa'c** vs **a'c**
- 4.2 **xa'c** peak heights for temperature experiments at pH 6.5
- 4.3 **xa'c** peak heights for pH experiments at 40°C
- 4.4 Resonance assignments for **xa'c**
- 4.5 Relaxation and Model-free data for **xa'c**
- 4.6 Relaxation and Model-free data for WT **b'x**
- 4.7 Relaxation and Model-free data for I272A **b'x**

Chapter 5

- 5.1 **b'xa'c** 15N/1H HSQC spectra for the temperature and pH experiments
- 5.2 I272A **b'xa'c** peak heights for temperature experiments at pH 6.5
- 5.3 I272A **b'xa'c** peak heights for pH experiments at 40°C
- 5.4 Resonance assignments for I272A **b'xa'c**
- 5.5 Relaxation and Model-free data for WT **b'xa'c**
- 5.6 Relaxation and Model-free data for I272A **b'xa'c**

Appendix 2.1 DNA sequencing of b'xa'c mutants

WT b'xa'c was used as the template for the generation of the I272A, L343A, 2DA and WT/W390F mutants. I272A/W390F, L343A/W390F and 2DA/W390F were generated using I272A, L343A and 2DA mutants, respectively, as templates. So, only the DNA and amino acid sequences of the W390F mutants are shown. The b'xa'c sequence is highlighted in grey and the mutation is shown in green.

WT/W390F

Frame 5' – 3'

```

nnmnggttctctagatattttgtttactttaagaaggagatataatgcatcaccatcac
X X V L - I F C L L - E G D I M H H H H
caccatatgcccttgtcatcgagttcaccgagcagacagccccgaagatTTTTGGAGGT
H H M P L V I E F T E Q T A P K I F G G
gaaatcaagactcacatcctgctgttcttggcccaagagtgtgtctgactatgacggcaaa
E I K T H I L L F L P K S V S D Y D G K
ctgagcaacttcaaaacagcagccgagagcttcaagggcaagatcctgttcatcttcac
L S N F K T A A E S F K G K I L F I F I
gacagcgaccacaccgacaaccagcgcacctcctcgagttcttggcctgaagaaggaagag
D S D H T D N Q R I L E F F G L K K E E
tgcccgccgtgcgccctcatcaccctggaggaggagatgaccaagtacaagcccgaatcg
C P A V R L I T L E E E M T K Y K P E S
gaggagctgacggcagagaggatcacagagttctgccaccgcttctggaggggcaaaatc
E E L T A E R I T E F C H R F L E G K I
aagccccacctgatgagccaggagctgccggaggactgggacaagcagcctgtcaaggtg
K P H L M S Q E L P E D W D K Q P V K V
cttgttgggaagaactttgaagacgtggcttttgatgagaaaaaaaacgtctttgtggag
L V G K N F E D V A F D E K K N V F V E
ttctatgccccatggtggtgctactgcaaacagttggctcccattttcgataaaactggga
F Y A P W C G H C K Q L A P I F D K L G
gagacgtacaaggaccatgagaacatcgtcatcgccaagatggactcgactgccaacgag
E T Y K D H E N I V I A K M D S T A N E
gtggaggccgtcaaagtgcacagcttccccacactcaagttcttctcctgccagtgccgac
V E A V K V H S F P T L K F F P A S A D
aggacggtcatcgattacaacggggaacgcacgctggatggtttaagaaatcctggag
R T V I D Y N G E R T L D G F K K F L E
agcgggtggccaggatggggcaggggatgatgacgatctcgaggacctggaagaagcagag
S G G Q D G A G D D D D L E D L E E A E
gagccagacatggaggaagacgatgatcagaaaagctgtgaaagatgaactgtaataagga
E P D M E E D D D Q K A V K D E L - - G
tccgaattcactagtgtgagctccgtcgacaagcttgcggccgactcgagcaccaccacca
S E F T S E L R R Q A C G R T R A P P P
ccaccactgagatccggctgctaacaagcccgaaggaagctgagttggctgctgccac
P P L R S G C - Q S P K G S - V G C C H
cgctgaacaataactagcataacccttggggcctctaaacgggtcttgaggggtttttt
R - T I T S I T P W G L - T G L E G F F
gctgaaaggaggaactatatccggattggcgaatgggacgcgcctgtagcggcgatta
A E R R N Y I R I G E W D A P C S G A L
agcgcggcggnnnnggttggttacgcgcagcgtgaccgctacacttgccagcgccttagcg
S A A X X V V T R S V T A T L A S A L A
ccennnctttnnntttnttcccttctttnnnnacgcttcgcccgnntttccccgtcaag

```

P X P X X X F P S X X X R S P X F P V K
 ctctaaatcggggnntnccctttnnggnccaattnnnnnnttanggnentnnccccaaa
 L - I G X X L X X X N X X X X X X X P K
 ancttgnnnaggggnaggnntnnnnnatgggnccnccccctgaaaancgggtttnccctt
 X L X R X R X X X G X X P - K X V X P L
 tgnntggannccnntcttnnaaagggnncttnnncnaatggnnnnnacnnnccn
 X X G X X X L X K X X X X X M X X X X X
 nnnngnnnt
 X X X X

WT/W390F
Frame 3' - 5'

nnnnnnnnntnnnggnntttnnnnnngggnnnnnnnnnnnnnnnnnnnnnnntncc
 X
 cnnncnggtnnnnnnaannnnnncnntnnnnnancnnnnngnnnnnnnnnnnnnn
 X X G X X X X X X X X X X X X X X X X X X X
 nnnnnnnnggncccnnnaannnnnnnnngnncannncagcgtnnnnnnnncnnnn
 X X X X P X X X X X X X X X Q R X X X X X
 nnggnttcccnnggnaccnncaccncccttggnnnnnnntccggannnnnatggtnnn
 X X S X X X X X X X P W X X X R X X M X X
 ggnngtgattnnnntccagncttnggaacacggaanccgaagcccattcatgtngt
 X X D X X X P X X X N T E X E A H S C X
 gttcagnnncagacgtttgcagcagcagnttcacgttcgcncggntatcggtgattc
 V Q X Q T F A A A V X S R S X X Y R - F
 attttgtaaccagtaaggcaaccccgccagcctagccgggtcctcaacgacaggagcac
 I L L T S K A T P P A - P G P Q R Q E H
 gatcatgcgcaccctggccaggacccaacgctgcccagatctcgatcccgcgaaatta
 D H A H P W P G P N A A R D L D P A K L
 atacgactcactataggagaccacaacggtttccctctagaataattttgtttaactt
 I R L T I G R P Q R F P S R N N F V - L
 taagaaggagatataatgcatcaccatcaccaccatagccccttgtcatcgagttcacc
 - E G D I M H H H H H M P L V I E F T
 gagcagacagcccgaagatttttggagggtgaaatcaagactcacatcctgctgttcttg
 E Q T A P K I F G G E I K T H I L L F L
 cccaagagtgtgtctgactatgacggcaaaactgagcaacttcaaacagcagccgagagc
 P K S V S D Y D G K L S N F K T A A E S
 ttcaagggcaagatcctgttcatcttcatcgacagcagaccacaccgacaaccagcgcac
 F K G K I L F I F I D S D H T D N Q R I
 ctcgagttcttggcctgaagaaggaagagtgccggcgcctcatcaccctggag
 L E F F G L K K E E C P A V R L I T L E
 gaggagatgaccaagtacaagcccgaatcggaggagctgacggcagagaggatcacagag
 E E M T K Y K P E S E E L T A E R I T E
 ttctgccaccgcttctggaggggcaaaatcaagcccacctgatgagccaggagctgccg
 F C H R F L E G K I K P H L M S Q E L P
 gaggactgggacaagcagcctgtcaaggtgcttgttgggaagaactttgaagacgtggct
 E D W D K Q P V K V L V G K N F E D V A
 tttgatgagaaaaaacgtcttggaggttctatgccccatgggtgtggctcactgcaa
 F D E K K N V F V E F Y A P W C G H C K
 cagttggctcccatttgcataaactgggagagacgtacaaggaccatgagaacatcgctc
 Q L A P I F D K L G E T Y K D H E N I V
 atcgccaagatggactcgactgccaacgaggtggaggccgtcaaagtgcacagcttcccc
 I A K M D S T A N E V E A V K V H S F P
 aactcaagttcttctgcccagtgccgacaggacggatcatcgattacaacggggaacgc
 T L K F F P A S A D R T V I D Y N G E R
 acgctggatgggtttaagaaatctctggagcgggtggccaggatggggcaggggatgat
 T L D G F K K F L E S G G Q D G A G D D
 gacgatctcgaggacctggaagaagcagaggagccagacatggaggaagacgatgatcag
 D D L E D L E E A E E P D M E E D D D Q

aaagctgtgaaagatgaactgtaataaggatccgaattcactagtgagctccgctcgacaa
 K A V K D E L - - G S E F T S E L R R Q
 gcttgccggccgactcgagcaccaccaccaccactgagatccggctgctaacaagcc
 A C G R T R A P P P P L R S G C - Q A
 nnaagagnnnnnnnn
 X K X X X X

I272A/W390F**Frame 5' - 3'**

nnnnnnngttctcnagatattttgtttactttaagaaggagatataatgcatcaccatcac
 X X V L X I F C L L - E G D I M H H H H
 caccatagccccctgtcatcgagttcaccgagcagacagccccgaagatttttggaggt
 H H M P L V I E F T E Q T A P K I F G G
 gaaatcaagactcacatcctgctgttcttgcccaagagtgtgtctgactatgacggcaaa
 E I K T H I L L F L P K S V S D Y D G K
 ctgagcaacttcaaaacagcagccgagagcttcaagggcaagatcctgttgccttcac
 L S N F K T A A E S F K G K I L F A F I
 gagcgcgaccacaccgacaaccagcgcacatcctcgagttctttggcctgaagaaggaagag
 D S D H T D N Q R I L E F F G L K K E E
 tgcccgccgctgcgccctcatcaccctggaggaggagatgaccaagtacaagcccgaatcg
 C P A V R L I T L E E E M T K Y K P E S
 gaggagctgacggcagagaggatcacagagttctgccaccgcttctcctggagggcaaaatc
 E E L T A E R I T E F C H R F L E G K I
 aagccccacctgatgagccaggagctgcccggaggactgggacaagcagcctgtcaaggtg
 K P H L M S Q E L P E D W D A K Q P V K V
 cttgttgggaagaacttgaagcvtggcttttgatgagaaaaaacgtctttgtggag
 L V G K N F E D V A F D E K K N V F V E
 ttctatgccccatgggtgtggctcactgcaaacagttggctcccattttgataaacggga
 F Y A P W C G H C K Q L A P I E D K L G
 gagacgtacaaggaccatgagaacatcgatcgccaagatggactcgactgccaacgag
 E T Y K D H E N I V I A K M D S T A N E
 gtggaggccgtcaaagtgcacagcttccccacactcaagttctttcctgccagtgccgac
 V E A V K V H S F P T L K F F P A S A D
 aggacggctcatcgattacaacggggaacgcacgctggatggttttaagaaattcctggag
 R T V I D Y N G E R T L D G F K K F L E
 agcgggtggccaggatggggcaggggatgatgacgatctcgaggacctggaagaagcagag
 S G G Q D G A G D D D D L E D L E E A E
 gagccagacatggaggaagacgatgatcagaaagctgtgaaagatgaactgtaataagga
 E P D M E E D D D Q K A V K D E L - - G
 tccgaattcactagtgagctccgctcgacaaagcttgccggccgactcgagcaccaccacca
 S E F T S E L R R Q A C G R T R A P P P
 ccaccactgagatccggctgctaacaagccccgaaaggaagctganttggctgctgccac
 P P L R S G C - Q S P K G S - X G C C H
 cgctgancaataactagcataacccttggggcctctaaacgggtcttgaggggtttttt
 R - X I T S I T P W G L - T G L E G F F
 gctgaaaggaggaactatatccggattggcgaatgggacgcgcctgtannggcgcatta
 A E R R N Y I R I G E W D A P C X G A L
 agcggcggcggnnnnggtggttacgcgcaccgtgaccgctacacttgccannncctaacg
 S A A X X V V T R T V T A T L A X X L T
 ccngctcctttcnntttcttcccttctttnnnncangttcnccggnnttcccngtcaa
 X A P F X F F P S F X X X F X X F P X Q
 gnnnaaatcggggnntnccctnanggttcnnattnannnnttannnncntnnnccccaa
 X X N R X X P X X F X X X X X X X P Q
 aannttgnnaaggnnnggttnnnnaaggnccctnccctnnaaacggtttnccct
 X X X K X X V X X R X X P X X N G X X P
 tnnncntnganncnntttnnnannnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
 X
 nnnnngnnnnt

I272A/W390F**Frame 3' – 5'**

nnggnannnnnnnnnnnnntnnnnnnnnnnnggncennnnnnaannnnnttcnggnnna
 X
 nnnncennnnnnnnnnnaaccennnnnnanggnnttcccnanggnnnncngcennnnncct
 X X X X X X X N X X X X X F X X X X X X X X X P
 nngnnnncannnnccggaannnaannnnngncaggnnnnttgatttccnggnnttcennncttn
 X X X X X E X X X X R X X I S X X X X X X
 ngnaacacggnnaaccgennnnccattcatgttntgttcaggtnncagacgttttgcanca
 X N T X T X X H S C X C S G X R R F A X
 gcagtcgcttcacgttcgctcgcgtatcggtgattcattctgctaaccagtaaggcaacc
 A V A S R S L A Y R - F I L L T S K A T
 ccgccagcctagccgggtcctcaacgacaggagcagcatcatgcgcaccctggccagga
 P P A - P G P Q R Q E H D H A H P W P G
 cccaacgctgcccagatctcgatcccgcgaaattaatacgactcactatagggagacca
 P N A A R D L D P A K L I R L T I G R P
 caacggtttccctctagaataattttgttttaactttaagaaggagatataatgcatcac
 Q R F P S R N N F V - L - E G D I M H H
 catcaccaccatgatgcccttgtcatcgagttcaccgagcagacagccccgaagattttt
 H H H H H M P L V I E F T E Q T A P K I F
 ggaggtgaaatcaagactcacatcctgctgttcttgccaagagtgtgtctgactatgac
 G G E I K T H I L L F L P K S V S D Y D
 ggcaactgagcaacttcaaacagcagccgagagcttcaagggcaagatcctgttcgccc
 G K L S N F K T A A E S F K G K I L F A
 ttcacgcagacgaccacaccgacaaccagcgcacccctcgagttctttggcctgaagaag
 F I D S D H T D N Q R I L E F F G L K K
 gaagagtccccggcctgctgcctcatcaccctggaggaggagatgaccaagtacaagccc
 E E C P A V R L I T L E E E M T K Y K P
 gaatcggaggagctgacggcagagaggatcacagagttctgccaccgcttctctggagggc
 E S E E L T A E R I T E F C H R F L E G
 aaaatcaagccccacctgatgagccaggagctgccggaggactgggacaagcagcctgtc
 K I K P H L M S Q E L P E D W D K Q P V
 aaggtgcttgttgggaagaactttgaagacgtggctttttgatgagaaaaaaacgtcttt
 K V L V G K N F E D V A F D E K K N V F
 gtggagttctatgccccatggtgtggtcactgcaaacagttggctcccattttcgataaa
 V E F Y A P W C G H C K Q L A P I F D K
 ctgggagacgtacaaggaccatgagaacatcgatcgccaagatggactgcactgccc
 L G E T Y K D H E N I V I A K M D S T A
 aacgaggtggaggccgtcaaagtgcacagcttccccacactcaagttctttcctgccagt
 N E V E A V K V H S F P T L K F F P A S
 gccgacaggacggtcatcgattacaacggggaacgcacgctggatggttttaagaaattc
 A D R T V I D Y N G E R T L D G F K K F
 ctggagagcgggtggccaggatggggcaggggatgatgacgatctcgaggacctggaagaa
 L E S G G Q D G A G D D D D L E D L E E
 gcagaggagccagacatggaggaagacgatgatcagaaagctgtgaaagatgaactgtaa
 A E E P D M E E D D D Q K A V K D E L -
 taaggatccgaattcactagtgagctccgtcgacaagcttgccggccgactcgagcacca
 - G S E F T S E L R R Q A C G R T R A P
 ccaccaccactgagatccggctgtaacaagccngnaagancnnngn
 P P P P L R S G C N K X X R X X X

L343A/W390F**Frame 5' – 3'**

nnnnnnnnnnnccggttctctagaatattttgtttactttaagaaggagatataatgcat
 X X X X R F S R I F C L L - E G D I M H
 caccatcaccaccatgatgcccttgtcatcgagttcaccgagcagacagccccgaagatt
 H H H H H M P L V I E F T E Q T A P K I
 tttggaggtgaaatcaagactcacatcctgctgttcttgccaagagtgtgtctgactat

F G G E I K T H I L L F L P K S V S D Y
 gacggcaaaactgagcaacttcaaaacagcagccgagagcttcaagggaagatcctgttc
 D G K L S N F K T A A E S F K G K I L F
 atcttcatcgacagcgaccacaccgacaaccagcgcacccctcgagttctttggcctgaag
 I F I D S D H T D N Q R I L E F F G L K
 aaggaagagtgcccgccgtgcgccctcatcaccctggaggaggagatgaccaagtacaag
 K E E C P A V R L I T L E E E M T K Y K
 cccgaatcggaggagctgacggcagagaggatcacagagttctgccaccgcttcctggag
 P E S E E L T A E R I T E F C H R F L E
 ggcaaaatcaagccccacctgatgagccaggaggccgaggaggactgggacaagcagcct
 G K I K P H L M S Q E A P E D W D K Q P
 gtcaaggtgcttgttgggaagaactttgaagacgtggcttttgatgagaaaaaacgctc
 V K V L V G K N F E D V A F D E K K N V
 tttgtggagtctatgccccatgggtggtcactgcaaacagttggctcccattttgat
 F V E F Y A P W C G H C K Q L A P I E D
 aaactgggagagacgtacaaggaccatgagaacatcgatcgccaagatggactcgact
 K L G E T Y K D H E N I V I A K M D S T
 gccaacgaggtggaggccgtcaaagtgcacagcttccccacactcaagttcttctcctgcc
 A N E V E A V K V H S F P T L K F F P A
 agtggcgacaggacgggtcatcgattacaacggggaacgcacgctggatggtttaagaaa
 S A D R T V I D Y N G E R T L D G F K K
 ttctggagagcgggtggccaggatggggcaggggatgatgacgatctcgaggacctggaa
 F L E S G G Q D G A G D D D D L E D L E
 gaagcagaggagccagacatggaggaagacgatgatcagaaagctgtgaaagatgaactg
 E A E E P D M E E D D D Q K A V K D E L
 taataaggatccgaattcactagtgtgagctccgtcgacaagcttgccggccgactcgagca
 - - G S E F T S E L R R Q A C G R T R A
 ccaccaccaccactgagatccgggtgctaacaagcccgaaggaagctgagttggc
 P P P P P L R S G C - Q S P K G S - V G
 tgctgccaccgctgancaataactagcataacccttggggcctctaaccgggtcttggaa
 C C H R - X I T S I T P W G L - T G L G
 ggggttttttgtgaaanggaggaactatatcccggattggcgaatgggacgcgcccctgg
 G V F C - X G G T I S R I G E W D A P W
 aaccgnnncattaanncngcnggnntgnnggttaccnnaacctggaccgctaaccn
 N X X I X X R X X X X V T X N L D R - X
 ttgccagcncnnaancngncnnnccttncnntttntcccncctttncnccgcnnt
 L P X P X X X X X X X F X P X L X P X X
 tnnccggnnnttccccgtcannnnntnaatccgggggnnccttnnnnggttcnncn
 X A X X S P S X X X I R G X P X X G S X
 ttnnnnnntttnnnggnncnncnnaannntnnnnnaanggnnganggttnnn
 X X X X X X X X X X K X X X X X X V X
 nntnnnngn
 X X X X

L343A/W390F

Frame 3' - 5'

agcaaccnccagcctagccgggtcctnaacgacaggagcacgatcatgcgccaccgtggc
 Q P X Q P S R V X N D R S T I M R T R G
 caggacccaacgctgccccgagatctcgatcccnggaaattaatacgactcactataggg
 Q D P T L P E I S I X X K L I R L T I G
 agaccacaacggtttccctctagaaataattttgtttaactttaagaaggagatataatg
 R P Q R F P S R N N F V - L - E G D I M
 catcaccatcaccaccatagccccttgtcatcgagttcaccgagcagacagccccgaag
 H H H H H M P L V I E F T E Q T A P K

atttttggaggtgaaatcaagactcacatcctgctgttcttggccaagagtgtgtctgac
 I F G G E I K T H I L L F L P K S V S D
 tatgacggcaactgagcaacttcaaacagcagccgagagcttcaagggcaagatcctg
 Y D G K L S N F K T A A E S F K G K I L
 ttcattctcatcgacagcgaccacaccgacaaccagcgcacccctcgagttctttggcctg
 F I F I D S D H T D N Q R I L E F F G L
 aagaaggaagagtgcccgccgtgcgctcatcaccctggaggaggagatgaccaagtac
 K K E E C P A V R L I T L E E E M T K Y
 aagccccgaatcggaggagctgacggcagagaggatcacagagttctgccaccgcttctg
 K P E S E E L T A E R I T E F C H R F L
 gagggcaaaatcaagccccacctgatgagccaggagggcggaggactgggacaagcag
 E G K I K P H L M S Q E A P E D W D K Q
 cctgtcaaggtgctgttgggaagaacttgaagacgtggcttttgatgagaaaaaac
 P V K V L V G K N F E D V A F D E K K N
 gtctttgtggagttctatgccccatggtgtggtcactgcaaacagttggctccattttc
 V F V E F Y A P W C G H C K Q L A P I F
 gataaactgggagagacgtacaaggaccatgagaacatcgatcatcgccaagatggactcg
 D K L G E T Y K D H E N I V I A K M D S
 actgccaacgaggtggaggccgtcaaagtgcacagcttccccacactcaagttctttct
 T A N E V E A V K V H S F P T L K F F P
 gccagtgccgacaggacggtcatcgattacaacggggaacgcacgctggatggtttaag
 A S A D R T V I D Y N G E R T L D G F K
 aaattcctggagagcgggtggccaggatggggcaggggatgatgacgatctcgaggacctg
 K F L E S G G Q D G A G D D D D L E D L
 gaagaagcagaggagccagacatggaggaagacgatgatcagaaagctgtgaaagatgaa
 E E A E E P D M E E D D D Q K A V K D E
 ctgtaataaggatccgaattcactagtgtgactccgtcgacaagcttggggccgactcga
 L - - G S E F T S E L R R Q A C G R T R
 gcaccaccaccaccactgagatccggctgtaacaagcccaagancgannnnnnnn
 A P P P P P L R S G C N K P K X R X X
 nnn
 X

2DA/W390F


Frame 5' – 3'

nnnnccggttctcagatattttgttactttaagaaggagatataatgcatcaccatcac
 X X G S Q I F C L L - E G D I M H H H H
 caccatattgcccctgtcatcgagttcaccgagcagacagccccgaagatttttggaggt
 H H M P L V I E F T E Q T A P K I F G G
 gaaatcaagactcacatcctgctgttcttggccaagagtgtgtctgactatgacggcaaa
 E I K T H I L L F L P K S V S D Y D K K
 ctgagcaacttcaaacagcagccgagagcttcaagggcaagatcctgttcatcttcatc
 L S N F K T A A E S F K G K I L F I F I
 gagacgaccacaccgacaaccagcgcacccctcgagttctttggcctgaagaaggaagag
 D S D H T D N Q R I L E F F G L K K E E
 tgcccgccgtgcgctcatcaccctggaggaggagatgaccaagtacaagcccgaatcg
 C P A V R L I T L E E E M T K Y K P E S
 gaggagctgacggcagagaggatcacagagttctgccaccgcttctggaggggcaaatc
 E E L T A E R I T E F C H R F L E G K I
 aagccccacctgatgagccaggagctgccggaggcctgggcaagcagcctgtcaaggtg
 K P H L M S Q E L P E A W A K Q P V K V
 cttgttgggaagaacttgaagacgtggcttttgatgagaaaaaacgcttttgggag
 L V G K N F E D V A F D E K K N V F V E
 ttctatgccccatggtgtggtcactgcaaacagttggctccattttcgataaactggga
 F Y A P W C G H C K Q L A P I F D K L G
 gagacgtacaaggaccatgagaacatcgatcgccaagatggactcgactgccaacgag
 E T Y K D H E N I V I A K M D S T A N E
 gtggaggccgtcaaagtgcacagcttccccacactcaagttctttctgcccagtgccgac
 V E A V K V H S F P T L K F F P A S A D

F I F I D S D H T D N Q R I L E F F G L
aagaaggaagagtgcccgccgtgcgccctcatcaccctggaggaggagatgaccaagtac
K K E E C P A V R L I T L E E E M T K Y
aagcccaatcggaggagctgacggcagagaggatcacagagttctgccaccgcttctctg
K P E S E E L T A E R I T E F C H R F L
gagggcaaatcaagccccacctgatgagccaggagctgccggaggcctgggccaagcag
E G K I K P H L M S Q E L P E A W A K Q
cctgtcaaggtgcttggttgggaagaactttgaagacgtggcttttgatgagaaaaaac
P V K V L V G K N F E D V A F D E K K N
gtctttgtggagttctatgccccatgggtgtggtcactgcaaacagttggctcccatttc
V F V E F Y A P W C G H C K Q L A P I F
gataaactgggagagacgtacaaggaccatgagaacatcgtcatcgccaagatggactcg
D K L G E T Y K D H E N I V I A K M D S
actgccaacgaggtggaggccgtcaaagtgcacagcttccccacactcaagttcttctct
T A N E V E A V K V H S F P T L K F F P
gccagtgccgacaggacgggtcatcgattacaacggggaacgcacgctggatggtttaag
A S A D R T V I D Y N G E R T L D G F K
aaattcctggagagcgggtggccaggatggggcaggggatgatgacgatctcgaggacctg
K F L E S G G Q D G A G D D D D L E D L
gaagaagcagaggagccagacatggaggaagacgatgatcagaaagctgtgaaagatgaa
E E A E E P D M E E D D D Q K A V K D E
ctgtaataaggatccgaattcactagtgagctccgtcgacaagcttgcgcccgactcga
L - - G S E F T S E L R R Q A C G R T R
gcaccaccaccaccactgagatccgggtgctaacaagccngnaagancnnngnn
A P P P P P L R S G C - Q A X K X X X

Appendix 4.1 Minimal shift map values of xa'c vs a'c

Residue	Minimal shift
350Gln	0.105302454
351Pro	
352Val	0.235969206
353Lys	0.061952593
354Val	-
355Leu	-
356Val	0.118135566
357Gly	0.401743424
358Lys	0.073344916
359Asn	0.160227521
360Phe	0.116001642
361Glu	0.11221198
362Asp	0.105900584
363Val	0.124712531
364Ala	0.03873699
365Phe	0.150549281
366Asp	0.502386317
367Glu	0.153124887
368Lys	0.02749247
369Lys	0.064855802
370Asn	0.163612192
371Val	-
372Phe	0.169631509
373Val	0.033568439
374Glu	0.115434244
375Phe	0.012804765
376Tyr	-
377Ala	0.138174481
378Pro	
379Trp	0.029013092
380Cys	0.038212885
381Gly	-
382His	0.078461413
383Cys	0.082831225
384Lys	0.089913032
385Gln	0.039129115
386Leu	0.063939925
387Ala	0.039949264
388Pro	
389Ile	0.034492744

 Resonances not present in amino acid

- Resonances could not be assigned

390Trp	0.057523396
391Asp	0.066213438
392Lys	0.173527887
393Leu	0.043794901
394Gly	0.141101565
395Glu	0.057627043
396Thr	0.018414661
397Tyr	0.040517154
398Lys	0.056462071
399Asp	-
400His	0.025243531
401Glu	0.021952335
402Asn	1.636305756
403Ile	0.151952688
404Val	0.104571984
405Ile	0.016150901
406Ala	0.138213238
407Lys	0.138189405
408Met	-
409Asp	0.189154747
410Ser	0.047484281
411Thr	0.243929218
412Ala	0.0753587
413Asn	0.130391597
414Glu	0.158118835
415Val	0.127034391
416Glu	0.110928204
417Ala	0.121703328
418Val	0.195725608
419Lys	0.049809923
420Val	-
421His	-
422Ser	0.124588649
423Phe	0.03089
424Pro	
425Thr	-
426Leu	0.037853459
427Lys	0.061119214
428Phe	0.017949444
429Phe	0.070312486
430Pro	
431Ala	0.054389177
432Ser	0.102344213
433Ala	0.117432677

434Asp	0.030725269
435Arg	0.079300019
436Thr	0.200877487
437Val	0.055221008
438Ile	0.050048983
439Asp	0.06162254
440Tyr	-
441Asn	-
442Gly	-
443Glu	0.018118877
444Arg	0.016634211
445Thr	0.037884681
446Leu	0.007932276
447Asp	0.016277653
448Gly	0.013508472
449Phe	0.007303671
450Lys	0.02524802
451Lys	0.014029203
452Phe	0.033262837
453Leu	0.057904712
454Glu	0.052961478
455Ser	0.031965349
456Gly	0.011366032
457Gly	0.022912771
458Gln	0.04346048
459Asp	-
460Gly	0.044650269
461Ala	0.041260103
462Gly	0.019633986
463Asp	0.015563756
464Asp	-
465Asp	-
466Asp	-
467Leu	-
468Glu	-
469Asp	-
470Leu	-
471Glu	-
472Glu	-
473Ala	0.024385605
474Glu	0.018510179
475Glu	-
476Pro	-
477Asp	-

478Met	-
479Glu	
480Glu	-
481Asp	-
482Asp	-
483Asp	-
484Gln	-
485Lys	0.048254078
486Ala	0.007248743
487Val	0.012852175
488Lys	0.016167359
489Asp	0.05224063
490Glu	0.018961069
491Leu	0.016087497

Appendix 4.2 xa'c peak heights for temperature experiments at pH 6.5

Residue	Peak heights			
	15°C	25°C	35°C	40°C
332Gly	2389870	1457270	491282	214040
333Lys	2763400	2159460	822528	506369
334Ile	8804990	4866580	3342750	2661760
335Lys	3699670	4768140	3477910	2809800
336Pro				
337His	-	-	-	-
338Leu	1192640	2217470	1918620	1639640
339Met	793738	1647320	2150150	2136650
340Ser	840182	1198950	1039850	893141
341Gln	372932	810316	1254570	1452120
342Glu	-	-	-	-
343Leu	-	-	-	-
344Pro				
345Glu	1284490	1992460	472376	740254
346Asp	540235	1493170	2291150	2555880
347Trp	655862	1598370	2234320	2784230
348Asp	1035670	2073840	3100060	1263630
349Lys	855592	1798580	2663160	3128050
350Gln	1289810	2451250	3291690	3566750
351Pro				
352Val	763434	1898700	2825540	3237920
353Lys	212301	602150	1039760	1417010
354Val	-	-	-	-
355Leu	-	-	-	-
356Val	538689	1234390	1766180	2022220
357Gly	276763	701388	1090520	1342870
358Lys	485224	1128920	1813700	194179
359Asn	851142	1581620	2120640	2383710
360Phe	557990	1298080	1968230	2365300
361Glu	834222	1684160	2405060	2779280
362Asp	765604	1743510	2791990	3325390
363Val	758052	1592290	2196040	2319170
364Ala	151228	255063	654827	993612
365Phe	55024.9	167951	446991	787305
366Asp	-9027.73	3346.653	117554	211863
367Glu	361586	335838	837493	1182530
368Lys	170838	319799	968805	1470220
369Lys	210867	151651	427458	682860
370Asn	151420	201862	73177.1	76049.8
371Val	-	-	-	-

372Phe	54194.6	765322	1184030	1288780
373Val	417097	975943	1398060	1508190
374Glu	272553	712597	1067740	1191130
375Phe	378239	821385	1179850	1228480
376Tyr	-	-	-	-
377Ala	479666	886579	1251620	1384650
378Pro				
379Trp	420522	953108	1338220	1596430
380Cys	563173	1148390	1600730	1771700
381Gly	-	-	-	-
382His	312125	281837	118717	77134.4
383Cys	223113	342345	282724	213575
384Lys	815097	1416400	1906000	2181790
385Gln	924714	1829970	2567580	2887390
386Leu	665248	1319440	1979210	2300700
387Ala	924476	1918380	2718380	3125150
388Pro				
389Ile	598694	1201470	1888120	2167220
390Trp	454451	1087950	1643920	1985610
391Asp	1063820	2093370	2944920	3132210
392Lys	917446	2100800	2862110	3380720
393Leu	724825	1576900	2322700	2501270
394Gly	581115	1290290	1930780	2171250
395Glu	1183160	2738090	3374710	3745410
396Thr	370833	1070610	1803200	567231
397Tyr	663742	1520190	2035980	2297110
398Lys	197873	890252	1969500	2617340
399Asp	-	-	-	-
400His	315245	1037350	1975930	2385820
401Glu	731521	1627270	2099430	2032390
402Asn	386724	831361	1156690	1189270
403Ile	99568.2	253132	551938	701091
404Val	375821	952395	1430040	1635390
405Ile	359777	859811	1588320	1757390
406Ala	504941	1069440	1575250	1669990
407Lys	642784	1394590	1841970	1983150
408Met	-	-	-	-
409Asp	315107	702258	1137270	1369690
410Ser	403955	822364	1241350	1507140
411Thr	384889	825311	1287200	1454830
412Ala	436556	932689	1394770	1617410
413Asn	944955	1995150	2485290	2559810
414Glu	653909	1497180	2141230	2422260
415Val	963190	1767210	2230300	2579710

416Glu	737668	1718980	2391750	2510480
417Ala	830395	1927790	2325370	2463210
418Val	7832850	36065.5	411214	934168
419Lys	452305	1543450	1380880	530562
420Val	-	-	-	-
421His	-	-	-	-
422Ser	1435490	1679990	1480210	1191650
423Phe	1108730	2490420	1423520	1524960
424Pro				
425Thr	-	-	-	-
426Leu	541868	1097250	1443970	1591730
427Lys	647402	1334120	1799210	135670
428Phe	453542	999612	1386550	1481350
429Phe	249422	619620	1031250	1212320
430Pro				
431Ala	549860	1167460	1899450	2344440
432Ser	196176	450260	1087530	1610470
433Ala	336365	409662	248542	164293
434Asp	723182	1561510	2200260	2340360
435Arg	168731	539624	946240	1120820
436Thr	308371	827113	1657650	2219710
437Val	645882	1521560	2299500	2574850
438Ile	451938	864861	1385500	1565890
439Asp	676877	1603180	2153760	2372710
440Tyr	-	-	-	-
441Asn	-	-	-	-
442Gly	-	-	-	-
443Glu	1354730	1222270	2134530	1110300
444Arg	420572	558954	254823	168445
445Thr	996437	1863150	2370840	2567700
446Leu	869843	2440540	4152670	4600590
447Asp	1167640	2175960	3117950	3320660
448Gly	549654	1163470	1693260	1895840
449Phe	1188110	2467910	3875080	3496830
450Lys	897872	1674590	1901710	2269230
451Lys	933558	2012530	2726000	3162860
452Phe	562189	1206610	1740070	1840940
453Leu	1227700	901840	1368990	1520380
454Glu	536036	1048350	1413800	1558980
455Ser	3050.234	30826.3	392834	858356
456Gly	836720	1646650	2095790	2177280
457Gly	571017	1134810	1410330	1508170
458Gln	678888	1033630	1440020	1461790
459Asp	-	-	-	-

460Gly	1078160	1598090	1885220	1812130
461Ala	1435870	2117520	2185610	2165640
462Gly	2897780	4707840	4391460	3785080
463Asp	5292410	8470820	7184710	6164750
464Asp	-	-	-	-
465Asp	-	-	-	-
466Asp	-	-	-	-
467Leu	-	-	-	-
468Glu	-	-	-	-
469Asp	-	-	-	-
470Leu	-	-	-	-
471Glu	-	-	-	-
472Glu	-	-	-	-
473Ala	7493820	10069900	8751560	7933700
474Glu	225628	196776	112905	138118
475Glu	-	-	-	-
476Pro				
477Asp	-	-	-	-
478Met	-	-	-	-
479Glu	-	-	-	-
480Glu	-	-	-	-
481Asp	-	-	-	-
482Asp	-	-	-	-
483Asp	-	-	-	-
484Gln	-	-	-	-
485Lys	21366100	13529000	10091100	8901500
486Ala	14539500	12540200	7025110	5312580
487Val	15284100	15178600	10162400	8561390
488Lys	1061660	1549370	1100190	1893830
489Asp	7275020	13602500	8580010	6479070
490Glu	5261920	12032400	9174250	8646500
491Leu	2044410	2135910	1797690	1683890

Appendix 4.3 xa'c peak heights for pH experiments at 40°C

Residue	Peak heights			
	pH 6.0	pH 6.5	pH 7.0	pH 8.0
332Gly	332255	214040	592829	591794
333Lys	240112	506369	193202	223104
334Ile	1331430	2661760	1187770	429715
335Lys	1402200	2809800	1289970	289224
336Pro				
337His	-	-	-	-
338Leu	739941	1639640	782923	848440
339Met	1137500	2136650	1052220	1406650
340Ser	431921	893141	432888	135172
341Gln	704689	1452120	944236	1289860
342Glu	-	-	-	-
343Leu	-	-	-	-
344Pro				
345Glu	658690	740254	365096	831735
346Asp	1223790	2555880	1548980	1743510
347Trp	1343240	2784230	1839240	2682550
348Asp	1435860	1263630	2085950	1430780
349Lys	1387030	3128050	1989380	3302130
350Gln	1439500	3566750	2115890	3613910
351Pro				
352Val	1430710	3237920	1951900	3439300
353Lys	692718	1417010	998375	1592700
354Val	-	-	-	-
355Leu	-	-	-	-
356Val	932892	2022220	1298460	2220480
357Gly	609316	1342870	851784	1548560
358Lys	851318	194179	1242110	2164600
359Asn	1173500	2383710	1498940	2516930
360Phe	1070040	2365300	1531520	2610480
361Glu	1385620	2779280	1801040	2722270
362Asp	1490510	3325390	2151400	3492980
363Val	1100360	2319170	1542270	2595850
364Ala	1088940	993612	1030270	1034390
365Phe	908224	787305	735547	716909
366Asp	821455	211863	287844	1033290
367Glu	1113740	1182530	1023200	1223170
368Lys	1284130	1470220	1248450	1554920
369Lys	1007840	682860	677634	676578
370Asn	136002	76049.8	479020	149527
371Val	-	-	-	-

372Phe	705355	1288780	862574	1443330
373Val	677883	1508190	990904	1802890
374Glu	596086	1191130	803278	1480930
375Phe	585981	1228480	768696	1464130
376Tyr	-	-	-	-
377Ala	643846	1384650	914348	1635180
378Pro				
379Trp	403283	1596430	974044	1423290
380Cys	829406	1771700	1185820	1920580
381Gly	-	-	-	-
382His	434.7871	77134.4	59645.6	351483
383Cys	-29224.5	213575	145035	315907
384Lys	871322	2181790	1318850	2440910
385Gln	1355650	2887390	1944790	3423750
386Leu	1029390	2300700	1505330	2845020
387Ala	1333200	3125150	1969630	3564710
388Pro				
389Ile	948522	2167220	1383400	2532350
390Trp	925350	1985610	1281250	2341070
391Asp	1368010	3132210	2044120	3592430
392Lys	1543170	3380720	2120080	3667460
393Leu	1189340	2501270	1786640	3247510
394Gly	979855	2171250	1366860	2515230
395Glu	1678030	3745410	2346610	4173800
396Thr	1016830	567231	1420890	2461580
397Tyr	1069610	2297110	1415830	2547070
398Lys	1250930	2617340	1693510	2711930
399Asp	-	-	-	-
400His	1185440	2385820	1531350	2394780
401Glu	1148690	2032390	1340480	969681
402Asn	546732	1189270	708576	835454
403Ile	412269	701091	501392	796193
404Val	814788	1635390	1089440	1854120
405Ile	812280	1757390	1164620	2050980
406Ala	768037	1669990	1095310	1930070
407Lys	938108	1983150	1285600	2207460
408Met	-	-	-	-
409Asp	637316	1369690	877800	1526950
410Ser	687151	1507140	972394	1788910
411Thr	688582	1454830	984796	1697420
412Ala	731482	1617410	1022630	1787160
413Asn	1111870	2559810	1549850	2674070
414Glu	946451	2422260	1453310	2187790
415Val	1004700	2579710	1595840	2721920

416Glu	1242470	2510480	1475220	1141730
417Ala	1290040	2463210	1358190	946363
418Val	127654	934168	238798	961566
419Lys	1032240	530562	1045180	1057900
420Val	-	-	-	-
421His	-	-	-	-
422Ser	713204	1191650	685571	470689
423Phe	676836	1524960	974993	1714030
424Pro				
425Thr	-	-	-	-
426Leu	701903	1591730	1013230	1682000
427Lys	785241	135670	1117270	2229610
428Phe	663150	1481350	970355	1749120
429Phe	611218	1212320	792363	1508500
430Pro				
431Ala	1183370	2344440	1625590	2446630
432Ser	1265600	1610470	1246710	2395890
433Ala	344970	164293	94335.2	-8520.38
434Asp	1376440	2340360	1537020	1352830
435Arg	758721	1120820	804778	367983
436Thr	1355120	2219710	1570730	2540680
437Val	978768	2574850	1507590	2772650
438Ile	652022	1565890	948279	1933750
439Asp	921011	2372710	1282850	2655980
440Tyr	-	-	-	-
441Asn	-	-	-	-
442Gly	-	-	-	-
443Glu	927838	1110300	951123	3118840
444Arg	243593	168445	108250	60003.3
445Thr	963056	2567700	1487300	2814920
446Leu	2426320	4600590	2689740	1909150
447Asp	1661990	3320660	2177310	2874650
448Gly	924283	1895840	1280120	2098070
449Phe	1530080	3496830	2088680	3411330
450Lys	1118770	2269230	1169350	1926990
451Lys	1260870	3162860	1927060	3914410
452Phe	3525120	1840940	3477630	2617620
453Leu	2328320	1520380	-173773	2219970
454Glu	583507	1558980	833088	2797690
455Ser	976021	858356	882595	164013
456Gly	925149	2177280	1293930	1118370
457Gly	678293	1508170	790507	1105320
458Gln	7734.479	1461790	514864	1244000
459Asp	-	-	-	-

460Gly	741181	1812130	775998	11520.3
461Ala	155483	2165640	1075190	7667220
462Gly	1253800	3785080	1897280	875930
463Asp	4104540	6164750	4375620	4866930
464Asp	-	-	-	-
465Asp	-	-	-	-
466Asp	-	-	-	-
467Leu	-	-	-	-
468Glu	-	-	-	-
469Asp	-	-	-	-
470Leu	-	-	-	-
471Glu	-	-	-	-
472Glu	-	-	-	-
473Ala	1257830	7933700	3536080	3785530
474Glu	78914.2	138118	69568.8	-6964.83
475Glu	-	-	-	-
476Pro				
477Asp	-	-	-	-
478Met	-	-	-	-
479Glu	-	-	-	-
480Glu	-	-	-	-
481Asp	-	-	-	-
482Asp	-	-	-	-
483Asp	-	-	-	-
484Gln	-	-	-	-
485Lys	4377380	8901500	4562380	2623490
486Ala	3389990	5312580	2672030	508451
487Val	4830750	8561390	4368660	2608860
488Lys	2356910	1893830	146046	717784
489Asp	-121444	6479070	4208240	2674650
490Glu	2944060	8646500	3653030	3544270
491Leu	6261280	1683890	8164890	1215910

Appendix 4.4 Resonance assignment for xa'c

Residue	H _N	N ^H	C _α	C _β
324Met	-	-	-	-
325His	-	-	-	-
326His	-	-	-	-
327His	-	-	-	-
328His	-	-	-	-
329His	-	-	-	-
330His	-	-	-	-
331Met	-	-	54.76	33.8
332Gly	8.38	109.76	45.40	
333Lys	8.09	120.65	55.18	34.45
334Ile	8.10	122.12	59.38	39.39
335Lys	8.38	127.18	53.35	33.9
336Pro			-	-
337His	-	-	54.99	32.15
338Leu	7.97	123.87	53.38	42.78
339Met	8.73	122.43	55.29	35.21
340Ser	8.30	115.26	56.42	63.64
341Gln	8.46	126.87	55.21	33.67
342Glu	-	-	-	-
343Leu	-	-	-	-
344Pro			61.23	33.84
345Glu	9.01	122.46	57.57	31.17
346Asp	8.47	115.57	51.94	40.21
347Trp	7.60	119.59	58.81	28.85
348Asp	7.48	119.11	51.31	38.25
349Lys	7.40	118.69	55.94	34.66
350Gln	8.04	117.31	52.53	29.49
351Pro			64.14	32.80
352Val	7.19	112.25	59.69	32.95
353Lys	9.33	128.76	53.99	-
354Val	-	-	-	-
355Leu	-	-	51.22	43.87
356Val	8.54	111.96	57.08	35.7
357Gly	8.92	109.88	47.92	
358Lys	8.14	114.8	57.58	34.67
359Asn	7.45	115.71	51.40	40.30
360Phe	7.73	120.65	61.50	40.76
361Glu	9.00	117.10	59.25	31.02
362Asp	7.73	117.49	55.55	41.24
363Val	7.04	116.19	62.29	34.29
364Ala	8.18	118.70	54.34	19.86
365Phe	6.94	108.91	55.21	-

366Asp	6.51	119.17	53.57	42.22
367Glu	9.02	128.97	56.57	31.06
368Lys	8.72	115.78	54.51	34.25
369Lys	7.26	116.17	52.52	37.52
370Asn	9.12	122.96	-	-
371Val	-	-	59.88	34.38
372Phe	9.79	133.51	51.6	41.64
373Val	9.61	125.67	59.87	36.47
374Glu	7.81	125.40	53.00	27.91
375Phe	9.54	130.68	56.25	38.97
376Tyr	-	-	54.05	43.49
377Ala	7.17	120.02	47.35	-
378Pro			62.71	33.38
379Trp	5.95	109.99	53.05	31.02
380Cys	6.59	126.20	58.59	31.84
381Gly	-	-	47.40	
382His	9.54	127.59	57.97	32.34
383Cys	9.80	128.34	62.45	30.38
384Lys	8.63	123.75	58.85	33.75
385Gln	7.56	117.45	57.13	30.6
386Leu	7.45	117.64	54.5	42.95
387Ala	7.42	121.69	56.06	18.6
388Pro			63.73	32.68
389Ile	7.31	119.52	62.35	38.88
390Trp	8.57	124.41	58.76	31.39
391Asp	8.35	116.20	56.21	40.69
392Lys	7.56	121.73	57.18	33.30
393Leu	8.63	126.89	57.04	41.57
394Gly	8.25	105.71	47.33	
395Glu	8.14	121.16	58.53	31.26
396Thr	8.01	115.96	64.62	66.59
397Tyr	7.24	117.33	58.11	38.99
398Lys	7.30	123.19	-	-
399Asp	-	-	-	-
400His	7.90	122.42	58.27	34.05
401Glu	8.56	126.65	57.98	31.94
402Asn	10.63	115.74	54.06	40.88
403Ile	9.04	127.08	60.18	39.83
404Val	8.80	129.84	59.92	37.02
405Ile	9.43	129.28	56.67	36.52
406Ala	9.27	128.10	49.99	26.36
407Lys	9.12	119.27	53.83	37.98
408Met	-	-	54.54	39.86
409Asp	8.63	125.46	51.57	39.66

410Ser	8.82	128.26	60.16	63.23
411Thr	8.83	110.82	60.77	66.72
412Ala	6.66	122.75	50.10	24.10
413Asn	7.07	114.71	51.83	45.37
414Glu	8.79	123.06	54.43	34.8
415Val	8.68	116.67	57.56	35.68
416Glu	8.76	122.94	57.82	31.85
417Ala	8.66	116.61	53.44	23.28
418Val	6.78	112.4	58.85	36.18
419Lys	8.33	125.10	51.58	32.24
420Val	-	-	-	-
421His	-	-	54.74	33.52
422Ser	7.06	113.31	55.56	62.8
423Phe	8.60	120.00	53.37	42.24
424Pro			-	-
425Thr	-	-	61.78	-
426Leu	9.41	130.74	52.94	43.55
427Lys	8.90	120.37	53.68	39.45
428Phe	9.72	124.15	53.81	43.99
429Phe	9.16	127.19	53.07	-
430Pro			60.40	34.74
431Ala	7.70	125.90	51.57	19.84
432Ser	7.87	119.12	55.65	63.14
433Ala	8.98	127.81	52.43	21.32
434Asp	8.17	117.20	53.16	41.87
435Arg	8.24	116.71	55.28	32.33
436Thr	7.47	115.67	61.78	67.67
437Val	8.34	124.38	59.84	35.74
438Ile	9.45	129.26	58.50	41.84
439Asp	8.92	128.84	54.09	42.13
440Tyr	-	-	-	-
441Asn	-	-	-	-
442Gly	-	-	44.42	
443Glu	8.67	118.38	55.7	32.10
444Arg	9.32	130.02	54.62	29.65
445Thr	6.87	108.29	56.24	69.27
446Leu	8.99	122.46	57.52	41.95
447Asp	8.27	115.36	56.42	41.24
448Gly	8.23	110.18	47.01	
449Phe	8.54	121.24	57.10	40.44
450Lys	8.73	118.40	59.10	34.17
451Lys	8.04	118.45	58.20	33.74
452Phe	7.88	120.32	59.11	40.79
453Leu	8.48	121.35	56.66	42.19

454Glu	8.88	118.28	30.58	57.09
455Ser	7.41	111.49	56.96	62.78
456Gly	7.86	111.39	46.13	
457Gly	7.87	106.63	44.98	
458Gln	7.06	116.94	54.95	31.86
459Asp	-	-	53.45	42.03
460Gly	8.44	110.16	45.43	
461Ala	8.06	122.49	52.02	22.60
462Gly	8.38	107.74	45.07	
463Asp	8.27	120.61	53.44	41.81
464Asp	-	-	-	-
465Asp	-	-	-	-
466Asp	-	-	-	-
467Leu	-	-	-	-
468Glu	-	-	-	-
469Asp	-	-	-	-
470Leu	-	-	-	-
471Glu	-	-	-	-
472Glu	-	-	55.21	32.18
473Ala	8.34	126.02	51.69	22.51
474Glu	7.98	125.64	56.85	32.85
475Glu	-	-	-	-
476Pro			-	-
477Asp	-	-	-	-
478Met	-	-	-	-
479Glu	-	-	-	-
480Glu	-	-	-	-
481Asp	-	-	-	-
482Asp	-	-	-	-
483Asp	-	-	-	-
484Gln	-	-	53.56	32.19
485Lys	8.33	121.29	55.57	34.24
486Ala	8.10	124.30	51.78	22.06
487Val	8.02	119.65	60.74	34.06
488Lys	8.32	125.30	55.15	34.48
489Asp	8.28	121.94	53.74	41.71
490Glu	8.24	121.26	55.26	32.17
491Leu	7.88	129.00	55.69	43.55

Appendix 4.5 Relaxation and Model-free data for xa'c

Residue	T ₁ /s	T ₁ error/s	T ₂ /s	T ₂ error/s	hetNOE	S ²	S ² error	S _r ²	S _r ² error	S _s ²	S _s ² error	τ _c	τ _c error	R _{ex}	R _{ex} error
332Gly	0.753683	0.044878	0.143523	0.006187	-0.45	0.361	0.021					1282.043	79.092		
333Lys	0.950634	0.144729	0.808153	0.138136	-0.53564	0.081	0.012					1.649	0.359		
334Ile	0.785768	0.030916	0.444341	0.022743	-4.03353	0.062	0.008					348.407	29.621		
335Lys	0.63522	0.045718	0.161621	0.017525	-1.02159	0.299	0.043					959.316	30.044		
336Pro															
337His															
338Leu	0.70122	0.020457	0.142941	0.002041	-0.64041	0.329	0.007					1098.693	35.588		
339Met	0.729539	0.037551	0.075175	0.002643		0.768	0.031					1000	419.771		
340Ser	0.944354	0.023323	0.060728	0.000981	-0.28861	0.957	0.013					13.775	0		
341Gln	1.051467	0.026886	0.056492	0.000931	-0.18607	0.958	0.013								
342Glu															
343Leu															
344Pro															
345Glu			0.075694	0.000425											
346Asp	0.920493	0.021816	0.062209	0.000933	-0.16369	0.951	0.012								
347Trp	0.968783	0.02516	0.061478	0.001022	-0.16144	0.943	0	0.998	0	0.945	0				
348Asp	0.946118	0.008386	0.056775	0.00055	-0.14358	0.963	0.006								
349Lys	0.938093	0.026195	0.067861	0.000562	-0.13505	0.881	0.007					9.769	25.909		
350Gln	0.909855	0.012014	0.077046	0.00078	-0.27719	0.789	0.008					116.739	14.013		
351Pro															
352Val	1.04254	0.024885	0.054882	0.000564											
353Lys	0.937076	0.036079	0.061944	0.001234	-0.10806	0.953	0.074								
354Val															

381Gly																			
382His	1.004112	0.041881	0.066689	0.003951	-0.1067	0.868	0.03												
383Cys	0.968527	0.042193	0.069405	0.001873	-0.0662	0.867	0.02												
384Lys	0.890782	0.0204	0.071373	0.000787	-0.13885	0.847	0.009						58.022	21.213					
385Gln	1.054544	0.152763	0.070474	0.000948	-0.1232	0.846	0.011												
386Leu	0.969819	0.018142	0.068394	0.001245	-0.1328	0.879	0.011												
387Ala	0.979146	0.02063	0.066393	0.00088	-0.08836	0.892	0.01												
388Pro																			
389Ile	0.941793	0.026899	0.068871	0.001018	-0.11035	0.874	0.012						5.089	25.352					
390Trp	0.86632	0.019785	0.06998	0.001181	-0.1408	0.88	0.013						72.98	29.272					
391Asp	0.950552	0.015262	0.069948	0.000789	-0.15377	0.86	0.009						35.599	18.479					
392Lys	0.911584	0.017649	0.067751	0.000834	-0.16239	0.889	0.01						46.432	26.361					
393Leu	0.914406	0.01838	0.06876	0.000668	-0.12361	0.874	0.008						43.265	23.603					
394Gly	0.84931	0.011035	0.065624	0.000606	-0.1309	0.909	0.009						579.446	124.696					
395Glu	0.907834	0.015214	0.06804	0.00158	-0.19541	0.905	0.016						55.598	29.989					
396Thr	0.951507	0.035493	0.059122	0.000911	-0.13582	0.99	0	0.995	0	0.995	0								
397Tyr	0.985479	0.018351	0.071394	0.000939	-0.15431	0.842	0.01						17.285	16.794					
398Lys	0.998301	0.030794	0.049816	0.000828	-0.18486	0.861	0	0.955	0	0.902	0						5.63	0.61	
399Asp																			
400His	0.965204	0.02124	0.022968	0.001297	-0.15801	0.9	0.02											20	2.481
401Glu	1.00776	0.026166	0.068148	0.000738	-0.17627	0.872	0.009												
402Asn	0.969208	0.019556	0.059105	0.000941	-0.17029	0.955	0.012												
403Ile	0.971097	0.04686	0.076857	0.002254															
404Val	0.95084	0.014551	0.064599	0.000963	-0.1107	0.914	0.01												
405Ile	0.973023	0.019661	0.06908	0.00027	-0.15687	0.863	0.004						11.417	19.241					
406Ala	0.979287	0.027954	0.061787	0.000529	-0.14956	0.957	0.008												

407Lys	0.926726	0.013956	0.05985	0.000729	-0.10973	0.967	0.009								
408Met															
409Asp	0.924078	0.050347	0.063087	0.000834	-0.143	0.945	0.012								
410Ser	0.967515	0.026945	0.065375	0.001172	-0.18714	0.905	0.014								
411Thr	0.854579	0.009836	0.070417	0.000943	-0.12308	0.861	0.011					462.975	68.141		
412Ala	1.028096	0.031748	0.067742	0.001166	-0.16746	0.869	0.013								
413Asn	0.993989	0.021351	0.063914	0.00063											
414Glu	0.970786	0.023242	0.061776	0.00112											
415Val	0.847405	0.164178	0.059995	0.00219											
416Glu	0.874029	0.041746	0.071822	0.016693	-0.29682	0.947	0.044					133.888	168.857		
417Ala	0.914688	0.076917	0.067449	0.001099	-0.46569	0.879	0.014					127.4	40.778		
418Val	1.027089	0.018374	0.060426	0.00059											
419Lys	0.982941	0.007747	0.085712	0.001031	-0.56342	0.728	0.007					100.444	5.784		
420Val															
421His															
422Ser	0.967351	0.015037	0.084497	0.00117											
423Phe	0.936138	0.015087	0.071081	0.000615	-0.16695	0.844	0.007					56.715	17.089		
424Pro															
425Thr															
426Leu	0.994301	0.024767	0.066682	0.000697	-0.13082	0.89	0.009								
427Lys	1.015508	0.013118	0.065072	0.000337	-0.20345	0.905	0.004								
428Phe	0.941048	0.018547	0.069989	0.000967	-0.15062	0.849	0	0.884	0	0.96	0	1000	0		
429Phe	0.99673	0.029323	0.072062	0.001181	-0.14521	0.834	0.013					5.2	17.967		
430Pro															
431Ala			0.066563	0.001352											
432Ser	1.057021	0.029691	0.062695	0.001126	-0.39169	0.895	0.013					32.868	0		

433Ala	1.081248	0.072242	0.0661	0.002517	-0.29805	0.864	0.029					29.342	22.091		
434Asp	0.94017	0.01759	0.075464	0.001113	-0.29366	0.81	0.011					64.446	13.96		
435Arg	0.86384	0.022488	0.067377	0.000913	-0.27956	0.895	0.011					95.033	35.244		
436Thr	0.955232	0.013924	0.080131	0.001347	-0.00458	0.783	0.012					71.672	13.284		
437Val	0.931401	0.016605	0.083013	0.002774	-0.24965	0.82	0.016					67.111	15.446		
438Ile	0.966842	0.024845	0.068285	0.000694	-0.15173	0.875	0.009					3.418	23.55		
439Asp	1.002737	0.015968	0.073763	0.000536	-0.1798	0.81	0.006					29.691	12.311		
440Tyr															
441Asn															
442Gly															
443Glu	0.937369	0.023779	0.072756	0.001049	-0.27461	0.83	0.011					48.215	16.577		
444Arg	1.009909	0.055729	0.077077	0.002832	-0.23076	0.79	0.025					14.438	13.668		
445Thr	0.95034	0.017808	0.07776	0.00043	-0.14529	0.767	0.004					59.251	12.187		
446Leu			0.073753	0.000779	-0.32601	0.806	0.009					30.807	14.551		
447Asp	0.975942	0.019604	0.073326	0.00099	-0.18146	0.824	0.01					30.859	15.204		
448Gly	0.918606	0.028083	0.069692	0.001109	-0.14071	0.868	0.013					20.14	24.116		
449Phe	0.88492	0.016193	0.066122	0.001238											
450Lys	0.887046	0.013561	0.06974	0.001217											
451Lys	0.895224	0.021216	0.069558	0.001686	-0.16439	0.893	0.017					46.502	29.789		
452Phe	0.902499	0.018674	0.068894	0.001069	-0.16024	0.883	0.012					52.716	26.211		
453Leu	0.893346	0.013338	0.073365	0.002319											
454Glu	0.881533	0.027571	0.071581	0.001502	-0.19085	0.861	0.016					42.74	23.306		
455Ser	0.94481	0.014402	0.073384	0.000831											
456Gly	0.85835	0.010367	0.076147	0.001568	-0.20975	0.834	0.015					444.835	57.677		
457Gly	0.792769	0.016421	0.079515	0.000832	-0.26635	0.729	0.009					846.127	71.54		
458Gln	0.839201	0.026345	0.091219	0.002017	-0.39581	0.663	0.015					873.405	67.883		

459Asp																			
460Gly	0.849438	0.083244	0.126281	0.011103	-0.51657	0.548	0.038							22.401	4.698				
461Ala	0.773507	0.009352	0.44654	0.009239	-1.05643	0.027	0.003							1023.214	4.311				
462Gly	0.66483	0.10173	0.191682	0.000629	-0.81921	0.177	0.002							1243.718	24.303				
463Asp	0.761382	0.005611	0.254112	0.003712	-0.91831	0.211	0.004							931.421	5.494				
464Asp																			
465Asp																			
466Asp																			
467Leu																			
468Glu																			
469Asp																			
470Leu																			
471Glu																			
472Glu																			
473Ala	0.797248	0.006624	0.420339	0.006315	-1.37794	0.06	0.002							539.965	6.64				
474Glu	1.09002	0.025704	0.404196	0.027564	-2.085	0.148	0.01							208.63	9.612				
475Glu																			
476Pro																			
477Asp																			
478Met																			
479Glu																			
480Glu																			
481Asp																			
482Asp																			
483Asp																			
484Gln																			

Appendix 4.5

485Lys	0.75777	0.026033	0.425447	0.008139	-1.30319	0.019	0.004					843.97	14.646		
486Ala	0.561065	0.07359	0.14637	0.029571	-0.94419	0.302	0.088					1028.001	62.872		
487Val	0.603923	0.080918	0.176621	0.035989	-1.33353	0.241	0.076					785.098	49.733		
488Lys	0.62504	0.090743	0.152714	0.031498	-0.9341	0.32	0.086					1024.272	64.356		
489Asp	0.545033	0.053603	0.151476	0.034477	-1.43959	0.225	0.09					742.442	56.774		
490Glu	1.210261	0.03701	0.896622	0.030155	-2.05386	0.006	0.003					256.862	10.535		

Appendix 4.6 Relaxation and Model-free data for WT b'x

Residue	T ₁ /s	T ₁ error/s	T ₂ /s	T ₂ error/s	hetNOE	S ²	S ² error	S _i ²	S _i ² error	S _s ²	S _s ² error	τ _c	τ _c error	R _{ex}	R _{ex} error
213Lys	0.767382	0.026255	0.2040184	0.005173	-1.62876							941.729	47.172		
214His	0.80075	0.046944	0.1284862	0.003005	-0.79088	0.443	0.029	0.808	0.03	0.549	0.03				
215Asn					-0.62511							1000	59.85		
216Gln	0.681718	0.032714	0.1739383	0.004602	-0.66916	0.325	0.022	0.714	0.027	0.455	0.025	930.608	31.857		
217Leu	0.705516	0.012593	0.1349465	0.004461	-0.63415	0.498	0.015								
218Pro												1000	243.36		
219Leu	0.751632	0.016201	0.0784011	0.002572	-0.39428	0.795	0.031	0.966	0.026	0.823	0.023	17.701	27.362	14.244	2.66
220Val	0.905614	0.013861			-0.22132	0.881	0.021								
221Ile	0.790221	0.02477	0.0701609	0.004026	-0.18809							231.517	0	0.008	0
222Glu	0.816174	0.018304	0.0584411	0.00341	-0.27386	0.979	0.022							9.762	1.824
223Phe	0.845002	0.032797	0.040428	0.002805	-0.16633	0.954	0.039								
224Thr					-0.18803							2000	1.989	4.388	1.967
225Glu	0.682886	0.049963			-0.2161	0.862	0.076							0.006	0
226Gln	0.773418	0.024568	0.0605947	0.002636	-0.24375	1	0.076					62.642	86.837	1.789	0.626
227Thr	0.830241	0.018221	0.0592804	0.001637	-0.22316	0.96	0.028					1000	427.922		
228Ala	0.771435	0.01954	0.0763882	0.003036	-0.31225	0.831	0.038	0.96	0.032	0.865	0.028				
229Pro												1221.264	474.49	3.277	0.868
230Lys	0.763799	0.010218	0.0555313	0.002292	-0.26662	0.926	0.031					874.009	799.043		
231Ile	0.796942	0.039427	0.0671716	0.002006	-0.25833	0.942	0.042	0.989	0.03	0.953	0.031	728.819	655.33	4.746	1.803
232Phe	0.790843	0.035144			-0.287	0.946	0.045					1732.186	469.233	5.511	1.429
233Gly	0.677243	0.019289			-0.31176	0.814	0.04					1000	397.368		
234Gly	0.711935	0.034901			-0.37787	0.829	0.058					6.508	62.457		
235Glu	0.865595	0.025631	0.0606231	0.004831	-0.20236	0.942	0.03					25.063	40.966		

236Ile	0.872527	0.033618	0.0694608	0.000668	-0.22078	0.917	0.009						148.025	0		
237Lys	0.899179	0.059498	0.065079	0.001108	-0.28347	0.97	0.015						0.538	45.543		
238Thr	0.887517	0.019335	0.0630806	0.002888	-0.2231	0.924	0.023									
239His	0.887789	0.021886	0.0664738	0.001849	-0.09822	0.929	0									
240Ile			0.0638512	0.001501	-0.21502											
241Leu	0.777286	0.005739	0.0595743	0.001103	-0.12604								18.827	49.346		
242Leu	0.861594	0.026124			-0.20149	0.929	0.03									
243Phe	0.839215	0.044564	0.0638029	0.001678	-0.09955	0.991	0						31.832	37.619		
244Leu	0.866366	0.011851	0.0715494	0.003375	-0.21226	0.916	0.019									
245Pro													75.894	53.907		
246Lys	0.830892	0.018596	0.068442	0.000941	-0.23236	0.935	0.012						0	53.907	0.007	0
247Ser	0.788144	0.023451	0.0595856	0.001503	-0.25892	1	0.012						920.065	281.795		
248Val	0.796726	0.023575	0.0745263	0.003027	-0.36903	0.839	0.036	0.952	0.029	0.881	0.027		48.394	20.743		
249Ser	0.859234	0.040785	0.0814582	0.004218	-0.31897	0.837	0.03									
250Asp	0.902821	0.041831	0.0632934	0.002903	-0.19405	0.944	0.035						0.667	1.959		
251Tyr					-0.20076	0.323	0.08						5.37	34.25	4.381	1.725
252Asp	0.898282	0.013703	0.0542646	0.004988	-0.18693	0.895	0.022						1000	0.961	1.511	0.947
253Gly	0.779014	0.021809	0.0599141	0.002705	-0.19835	0.962	0.045						889.69	360.234		
254Lys	0.799187	0.028653	0.072279	0.002673	-0.23675	0.915	0.028						751.601	407.184		
255Leu	0.795891	0.012279	0.0669484	0.00255	-0.23363	0.961	0.024						614.39	230.336		
256Ser	0.801072	0.009882	0.0692843	0.001775	-0.1669	0.948	0.02									
257Asn	0.83652	0.018883	0.0624339	0.001027	-0.20392	1	0.02						1000	0		
258Phe	0.797366	0.014999	0.0649749	0.001761	-0.04662	0.992	0.017									
259Lys	0.879883	0.036643	0.0636154	0.000952	-0.20951	0.99	0	0.995	0	0.995	0		50.391	39.259		
260Thr	0.830413	0.020002	0.0709677	0.001117	-0.1618	0.911	0.013						569.4	110.377		
261Ala	0.801195	0.010693	0.0708756	0.000301	-0.19678	0.889	0.004						17.126	63.059		
262Ala	0.853352	0.023877	0.0670166	0.001768	-0.20338	0.946	0.02									

Appendix 4.6

263Glu	0.792204	0.009541	0.0689485	0.001672	-0.24063							47.386	11.278		
264Ser	0.847026	0.025039	0.0874072	0.001017	-0.29089	0.74	0.008					71.827	25.125		
265Phe	0.842703	0.017631	0.074719	0.001008	-0.21579	0.866	0.011					89.99	51.11		
266Lys	0.833016	0.01597	0.0694865	0.00182	-0.2553	0.932	0.018					36.157	25.031		
267Gly	0.8217	0.044725	0.0760786	0.002433	-0.24789	0.861	0.024					493.482	0		
268Lys	0.758095	0.041518	0.0669077	0.00093	-0.35863	0.943	0.012					9.515	42.939		
269Ile	0.87598	0.019862	0.0682948	0.001132	-0.21724	0.927	0.014					30.453	82.042	2.214	0.836
270Leu	0.838456	0.031977	0.0580886	0.001868	-0.20206	0.956	0.041					181.02	378.058	6.383	4.064
271Phe					-0.37022	0.928	0.068					1000	1.395	1.694	1.395
272Ile					-0.21722	0.973	0.061					100.668	233.228	2.097	0.457
273Phe	0.813649	0.012748	0.0569767	0.000885	-0.20264	0.984	0.025							2.26	0.934
274Ile	0.813349	0.020379	0.056129	0.002439	-0.14535	0.991	0.03					679.023	196.43		
275Asp	0.801655	0.015857	0.0744004	0.00244	-0.21099	0.91	0.023								
276Ser	0.839581	0.01401	0.0642568	0.001954	-0.14898	0.967	0.024					53.566	54.061		
277Asp	0.819507	0.025939	0.0692068	0.00133	-0.19875	0.931	0.016								
278His	0.785815	0.031536	0.0669469	0.001633	-0.20322							15.932	25.489		
279Thr	0.884152	0.059517			-0.21156	0.865	0.055					24.747	27.637		
280Asp	0.801342	0.036727	0.0765123	0.002347	-0.17067	0.87	0.023					654.454	450.548		
281Asn	0.809115	0.021298	0.0670945	0.00103	-0.21129	0.952	0.015					258.684	98.463	3.851	0.253
282Gln	0.832885	0.028609	0.0595939	0.000686	-0.75427	0.811	0.009					72.755	16.46		
283Arg			0.0777446	0.001178	-0.47487	0.813	0.012					1000	87.509		
284Ile	0.757186	0.043581			-0.63814	0.39	0.031	0.699	0.04	0.557	0.031	1000	0.732	2.167	0.548
285Leu	0.783479	0.021513	0.0570052	0.001466	-0.16142	0.976	0.043					49.577	35.276		
286Glu	0.854726	0.016765	0.0725134	0.002382	-0.21807	0.91	0.019					262.581	0		
287Phe	0.692511	0.014881	0.0643344	0.000819	-0.51683	1	0.01					105.876	13.658		
288Phe	0.936278	0.024247			-0.7311	0.744	0.019					150.105	66.545		
289Gly	0.817045	0.01237	0.072622	0.002353	-0.2329	0.934	0.019					115.906	27.252		

Appendix 4.6

290Leu	0.870409	0.019117	0.0723895	0.001327	-0.4619	0.871	0.013						39.487	20.889		
291Lys	0.905961	0.014126	0.0709166	0.002179	-0.31144	0.873	0.016						86.851	31.447		
292Lys	0.840074	0.014601	0.0735741	0.001527	-0.23311	0.895	0.015						655.431	202.472		
293Glu	0.803626	0.015846	0.0691673	0.00043	-0.21489	0.914	0.007						204.378	56.968		
294Glu	0.826245	0.015457	0.0787241	0.003829	-0.49817	0.877	0.016						77.441	25.735		
295Cys	0.834198	0.018968	0.0749577	0.001083	-0.24696	0.865	0.011									
296Pro													636.864	0		
297Ala	0.851842	0.020152	0.0640264	0.001845	-0.16927	0.956	0.018	0.956	0.018	1	0					
298Val	0.881193	0.032418	0.0620972	0.001518	-0.22707	0.987	0	0.993	0	0.993	0	59.546	0			
299Arg	0.886328	0.041316	0.0605468	0.002174	-0.24468	0.988	0.028									
300Leu	0.858874	0.019069	0.064134	0.002819	-0.14047										2.062	0
301Ile	0.807311	0.028373	0.0563884	0.001095	-0.15659	0.999	0								6.337	2.095
302Thr	0.874013	0.044181			-0.15941	0.923	0.05						43.401	43.487		
303Leu	0.826295	0.025981	0.0720524	0.002227	-0.18838	0.917	0.022						981.629	204.767		
304Glu	0.794344	0.014861	0.0790477	0.002883	-0.41024	0.785	0.03	0.931	0.024	0.843	0.023	990.84	43.988			
305Glu	0.679413	0.012428	0.109873	0.00185	-0.49228	0.535	0.011						770.615	39.363		
306Glu	0.779569	0.014475	0.0969692	0.000865	-0.53088	0.623	0.007						855.013	270.408		
307Met	0.764397	0.036171			-0.40249	0.866	0.045						203.299	0		
308Thr	0.840015	0.027194	0.0658684	0.001964	-0.39935	0.933	0.016						49.09	118.98		
309Lys	0.822399	0.026418	0.0664325	0.00307	-0.19807	0.969	0.028									
310Tyr	0.806931	0.014459	0.0459519	0.002657	-0.20875								41.896	77.782		
311Lys			0.0671471	0.001506	-0.21972	0.948	0.021									
312Pro													117.717	47.458		
313Glu	0.830616	0.016363	0.071332	0.002052	-0.30679	0.917	0.018						649.291	73.896		
314Ser	0.780458	0.022489			-0.75632	0.72	0.029						57.937	30.216		
315Glu	0.834848	0.035732	0.0752342	0.002842	-0.28734	0.879	0.026						719.742	147.877		
316Glu	0.813039	0.023814	0.0815931	0.004201	-0.54229	0.76	0.038	0.931	0.029	0.816	0.031	99.008	21.845			

Appendix 4.6

317Leu	0.837161	0.016515	0.078086	0.001313	-0.32433	0.842	0.012					82.518	20.486		
318Thr	0.891214	0.026808	0.0751333	0.001464	-0.43131	0.845	0.014					47.165	31.391		
319Ala	0.787663	0.04045	0.0726758	0.001169	-0.24315	0.884	0.014					29.439	53.927		
320Glu	0.814981	0.034342	0.0689577	0.001201	-0.17807	0.931	0.015					54.777	54.27		
321Arg	0.826188	0.026512	0.0684169	0.000636	-0.21339	0.932	0.009					1000	0	0.257	0
322Ile	0.776729	0.022683	0.0644019	0.000982	-0.16832	0.968	0								
323Thr	0.787046	0.022781	0.0663657	0.001076	-0.10246							25.537	59.222		
324Glu	0.824761	0.0285	0.069443	0.002061	-0.16749	0.938	0.022					30.893	124.466		
325Phe	0.813284	0.022837	0.0664096	0.001584	-0.16093	0.971	0.019					974.827	241.415		
326Cys	0.776049	0.016784	0.0699864	0.00073	-0.19278	0.901	0.011					74.377	99.288		
327His	0.815509	0.024512	0.066506	0.000835	-0.2069	0.96	0.012					90.161	105.481		
328Arg	0.810412	0.022793	0.0669565	0.001328	-0.20616	0.96	0.017								
329Phe	0.763928	0.031182	0.0698515	0.003063	-0.20483							1000	243.694		
330Leu					-0.43336	0.8	0.045					86.08	27.207		
331Glu	0.842084	0.014134	0.0728917	0.000658	-0.19665	0.879	0.008					1000	0	1.455	0
332Gly	0.760581	0.030171	0.0601229	0.001427	-0.1802	0.962	0								
333Lys	0.858599	0.014854	0.0663632	0.001023	-0.20647	0.951	0.014								
334Ile	0.933836	0.032948	0.0613793	0.002314	-0.21709	0.929	0.033					50.208	87.598		
335Lys	0.829839	0.022908	0.0666774	0.002179	-0.20917	0.959	0.023								
336Pro															
337His						0.944	0					168.776	57.835	5.984	1.168
338Leu	0.833762	0.015158	0.0492051	0.002725	-0.41413	0.909	0.021					959.165	15.525		
339Met	0.757813	0.023567	0.1296892	0.002081	-0.88332	0.427	0.009					1000	18.229		
340Ser	0.681554	0.018101	0.1746966	0.002069	-0.85821	0.295	0.011	0.795	0.016	0.371	0.012	22.975	38.99		
341Gln					-0.51412	0.551	0.413					938.8	72.168		
342Glu			0.0952045	0.003828	-0.72339	0.611	0.034					48.467	2.56		
343Leu	0.845666	0.031124	0.1127241	0.001254	-0.76142	0.569	0.006								

Appendix 4.6

344Pro												1000	6.082		
345Glu	0.60486	0.019266	0.2135437	0.008259	-0.96128	0.254	0.009	0.829	0.021	0.306	0.007	1000	11.131		
346Asp	0.658827	0.015644	0.2242941	0.010322	-0.9203	0.268	0.011	0.806	0.017	0.333	0.011	1000	29.804		
347Trp	0.630118	0.013336	0.2034166	0.005578	-0.77832	0.251	0.011	0.848	0.017	0.296	0.012	1000	3.845		
348Asp			0.2287559	0.004285	-1.02892	0.205	0.005					1000	11.811		
349Lys	0.675177	0.00857	0.2584409	0.004368	-0.89298	0.177	0.006	0.776	0.009	0.228	0.007	602.002	17.897		
350Gln	0.779836	0.022025	0.4451385	0.014516	-1.70343	0.043	0.006								
351Pro												832.661	80.562		
352Val					-0.83636	0.582	0.052								

Appendix 4.7 Relaxation and Model-free data for I272A b'x

Residue	T ₁ /s	T ₁ error/s	T ₂ /s	T ₂ error/s	hetNOE	S ²	S ² error	R _{ex}	R _{ex} error
213Lys									
214His	0.676495	0.032672			-0.82388	0.784	0.035		
215Asn	0.648511	0.018454	0.081036	0.001121	-0.42163				
216Gln	0.625995	0.008145	0.191222	0.007402	-0.72485	0.79	0.01		
217Leu	0.68298	0.005944	0.235127	0.018713	-0.80773	0.862	0.008		
218Pro									
219Leu	0.7364	0.014208	0.092877	0.001491	-0.62763	0.877	0.011		
220Val									
221Ile	0.680727	0.014109	0.078976	0.001205	-0.34059	0.999	0.012		
222Glu									
223Phe									
224Thr	0.620896	0.030737	0.065507	0.002605	-0.0803	0.989	0.049	3.606	0.838
225Glu									
226Gln	0.643789	0.020972	0.07386	0.001897	-0.18767				
227Thr	0.648749	0.024049	0.064777	0.00263	-0.21777	0.947	0.035	4.278	0.751
228Ala	0.618513	0.025126	0.07176	0.00052	-0.2303	0.993	0.04	2.231	0.486
229Pro									
230Lys			0.066421	0.000277	-0.27908	0.839	0.211	5.162	2.491
231Ile	0.62823	0.021469	0.06684	0.000515	-0.20222	0.978	0.033	3.437	0.41
232Phe	0.609949	0.0235	0.057419	0.002315	-0.20076			5.628	0
233Gly	0.605533	0.012353	0.076497	0.001223	-0.29784			1.284	0
234Gly	0.625757	0.030182	0.074114	0.004454	-0.31484	0.981	0.047	1.923	0.984
235Glu	0.699509	0.024704	0.061133	0.00095	-0.30157	0.878	0.031	6.008	0.445
236Ile	0.736997	0.012107	0.091746	0.003408	-0.59696	0.846	0.013		
237Lys	0.626977	0.024682	0.062259	0.002581	-0.25246	0.98	0.039	4.515	0.806
238Thr	0.675015	0.025684	0.06739	0.000582	-0.45961	0.91	0.035	4.114	0.428
239His	0.644414	0.026161	0.049548	0.000789	-0.17913	0.953	0.039	8.948	0.558
240Ile	0.566023	0.029846	0.063697	0.002779	-0.19669			3.911	0
241Leu									
242Leu	0.645787	0.041522	0.06934	0.001754	-0.19112	0.951	0.061	3.211	0.808
243Phe	0.637702	0.022577	0.068272	0.00207	-0.1893				
244Leu	0.652721	0.018509	0.065393	0.000843	-0.18813	0.941	0.027	4.201	0.371
245Pro									
246Lys	0.649995	0.024071	0.06313	0.00102	-0.23493	0.945	0.035	4.702	0.485
247Ser	0.620204	0.027759	0.062712	0.002962	-0.23289	0.99	0.044	4.273	0.917
248Val	0.637372	0.027594	0.061613	0.001354	-0.23148	0.964	0.042	4.872	0.608
249Ser	0.696705	0.026797	0.066059	0.002154	-0.24903	0.881	0.034	4.747	0.635
250Asp	0.58949	0.033191	0.065286	0.000917	-0.22304			3.529	0
251Tyr	0.625154	0.051528			-0.23306	0.957	0.074		
252Asp	0.683193	0.034769	0.058984	0.000332	-0.22027	0.899	0.046	6.357	0.548
253Gly	0.589129	0.019698	0.071356	0.000529	-0.17361			2.226	0

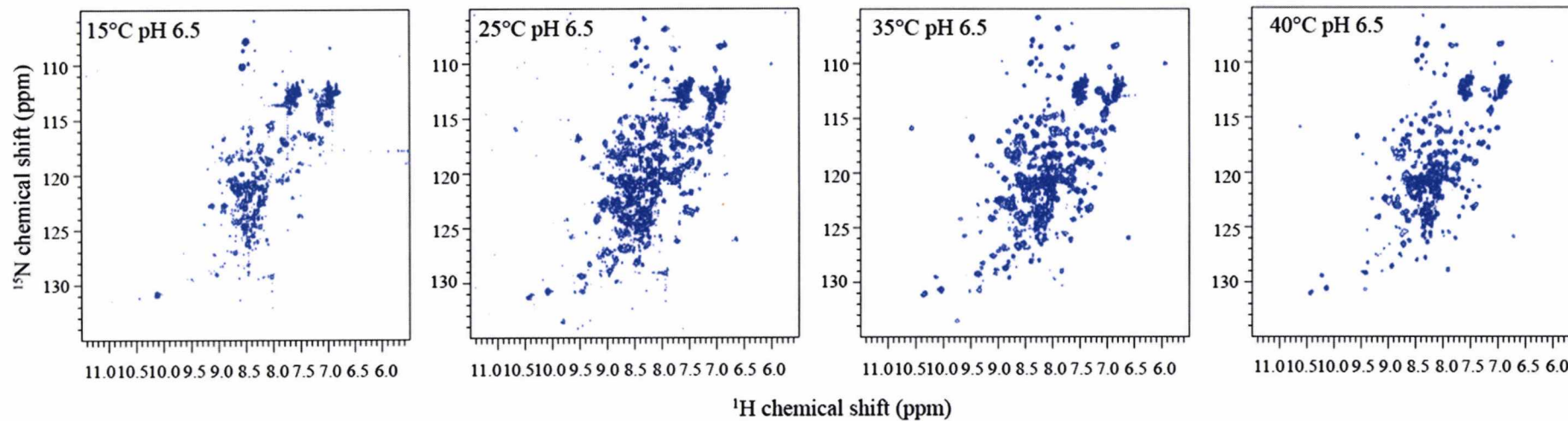
254Lys	0.612987	0.019578	0.068268	0.001419	-0.20573			2.86	0
255Leu	0.607057	0.022146	0.067997	0.001254	-0.18901			2.918	0
256Ser	0.652974	0.019867	0.069841	0.000992	-0.21472	0.941	0.029	3.231	0.394
257Asn	0.644198	0.018613	0.094878	0.004456	-0.30993	0.936	0.023		
258Phe	0.612583	0.017911	0.068677	0.00044	-0.19148			2.773	0
259Lys	0.629764	0.020033	0.070917	0.001018	-0.19883	0.975	0.031	2.605	0.418
260Thr	0.654973	0.022528	0.071236	0.001098	-0.19794	0.938	0.032	2.985	0.437
261Ala	0.609328	0.019662	0.070681	0.000719	-0.2108			2.36	0
262Ala	0.637368	0.016211	0.07015	0.0009	-0.16317	0.964	0.025	2.897	0.342
263Glu	0.63856	0.018261	0.071633	0.001077	-0.6954	0.962	0.028	2.623	0.386
264Ser									
265Phe	0.658125	0.020318	0.074061	0.000882	-0.22594	0.933	0.029	2.502	0.376
266Lys	0.660391	0.016752	0.071137	0.001421	-0.26369	0.93	0.024	3.095	0.395
267Gly	0.630038	0.03442	0.072903	0.002108	-0.23198	0.975	0.053	2.226	0.743
268Lys	0.62186	0.02785	0.066604	0.000607	-0.21935	0.988	0.044	3.372	0.539
269Ile	0.677219	0.030858	0.074523	0.005233	-0.23013	0.907	0.041	2.728	1.061
270Leu	0.568715	0.021984	0.053072	0.002334	-0.16726			7.054	0
271Phe	0.577897	0.015699	0.047231	0.000949	-0.18948				
272Ile	0.633335	0.025562	0.072535	0.001538	-0.23272	0.97	0.039	2.355	0.546
273Phe			0.081412	0.002439	-0.71029				
274Ile									
275Asp	0.699608	0.011442	0.087369	0.001466	-0.41104	0.919	0.011		
276Ser	0.626862	0.022658	0.066871	0.001354	-0.18991	0.98	0.035	3.405	0.516
277Asp	0.641721	0.026782	0.066595	0.000729	-0.25566	0.957	0.04	3.735	0.499
278His	0.640197	0.022599	0.068406	0.00103	-0.19322				
279Thr	0.62982	0.024894	0.071805	0.002254	-0.17447	0.975	0.039	2.432	0.631
280Asp	0.60521	0.028226	0.074695	0.002202	-0.22113			1.6	0
281Asn	0.625621	0.018617	0.067247	0.001009	-0.22364				
282Gln	0.649126	0.024268	0.062265	0.001361	-0.19966	0.946	0.035	4.907	0.545
283Arg	0.661118	0.019647	0.071488	0.000603	-0.26605	0.929	0.028	3.038	0.346
284Ile									
285Leu									
286Glu	0.659911	0.017393	0.069676	0.000636	-0.21638				
287Phe	0.645474	0.013544	0.075552	0.001262	-0.4233	0.951	0.02	2.02	0.323
288Phe									
289Gly	0.630418	0.016836	0.070378	0.002063	-0.23825	0.974	0.026	2.725	0.517
290Leu									
291Lys	0.674723	0.025176	0.078365	0.001185	-0.28892				
292Lys	0.635717	0.01988	0.06515	0.001131	-0.1883	0.966	0.03	3.961	0.445
293Glu	0.617936	0.015509	0.070821	0.000577	-0.19116	0.994	0.025	2.404	0.316
294Glu	0.655096	0.015351	0.065906	0.000829	-0.3589	0.937	0.022	4.122	0.322
295Cys	0.688697	0.017785	0.075588	0.000603	-0.23891				
296Pro									
297Ala	0.656393	0.017587	0.067636	0.001538	-0.23093	0.936	0.025	3.756	0.448

298Val	0.709347	0.020877	0.069159	0.001426	-0.2417	0.866	0.025	4.254	0.423
299Arg	0.673766	0.025463	0.065601	0.000619	-0.19294	0.912	0.034	4.499	0.431
300Leu	0.65481	0.022875	0.069618	0.001396	-0.18862	0.938	0.033	3.308	0.482
301Ile					-0.20176				
302Thr	0.610033	0.021129	0.038279	0.001376					
303Leu	0.639115	0.011852	0.056436	0.001403	-0.33905				
304Glu	0.631199	0.01679	0.108458	0.004148	-0.53429	0.891	0.02		
305Glu	0.594609	0.015228	0.116932	0.002516	-0.52718	0.805	0.013		
306Glu	0.707805	0.012987	0.1084	0.00125	-0.55077	0.803	0.008		
307Met	0.607333	0.017685	0.075941	0.000787	-0.36926			1.38	0
308Thr	0.632757	0.020056	0.055896	0.000703	-0.24152				
309Lys	0.671753	0.021587	0.073799	0.002716	-0.11152	0.914	0.029	2.773	0.607
310Tyr	0.561622	0.030965	0.069165	0.001034	-0.18508			2.67	0
311Lys	0.638319	0.029306	0.070145	0.000936	-0.24057	0.962	0.044	2.915	0.554
312Pro									
313Glu	0.671805	0.027508	0.074543	0.001302	-0.31861	0.914	0.037	2.639	0.5
314Ser									
315Glu	0.64618	0.028574	0.069421	0.001357	-0.20653	0.95	0.042	3.201	0.57
316Glu	0.704435	0.020479	0.079419	0.002693	-0.45743	0.936	0.021		
317Leu	0.661357	0.018994	0.066911	0.001306	-0.24321				
318Thr	0.771989	0.021396	0.084974	0.003524	-0.42056	0.84	0.019		
319Ala	0.607601	0.023317	0.07364	0.001126	-0.25797			1.791	0
320Glu	0.601883	0.025061	0.070319	0.002404	-0.20198			2.433	0
321Arg									
322Ile	0.608016	0.015143	0.066382	0.000782	-0.18201			3.276	0
323Thr	0.608195	0.022628	0.070303	0.001821	-0.20486			2.436	0
324Glu	0.618164	0.026678	0.067907	0.0012	-0.20655	0.993	0.043	3.014	0.569
325Phe	0.627114	0.021164	0.066908	0.000399	-0.20133	0.979	0.033	3.402	0.4
326Cys	0.632343	0.019952	0.068478	0.000919	-0.17734	0.971	0.031	3.154	0.411
327His	0.628847	0.017945	0.065128	0.001842	-0.16504	0.977	0.028	3.842	0.545
328Arg	0.624776	0.026788	0.068821	0.00094	-0.181	0.983	0.042	2.943	0.535
329Phe	0.623259	0.021608	0.066997	0.000403	-0.19602	0.985	0.034	3.31	0.413
330Leu	0.711111	0.032694	0.06423	0.000917	-0.21284	0.864	0.04	5.388	0.518
331Glu	0.664062	0.017175	0.075759	0.000854	-0.21064	0.925	0.024	2.298	0.319
332Gly	0.644217	0.019495	0.069854	0.00397	-0.18796	0.953	0.029	3.078	0.882
333Lys									
334Ile	0.686657	0.020829	0.069595	0.000607	-0.23317	0.894	0.027	3.825	0.343
335Lys	0.651021	0.017671	0.068983	0.003886	-0.22653	0.943	0.026	3.376	0.871
336Pro									
337His									
338Leu	0.650806	0.025324	0.062962	0.001431	-0.30852	0.944	0.037	4.758	0.564
339Met									
340Ser	0.646986	0.017446	0.060943	0.001102	-0.16922	0.949	0.026	5.219	0.423
341Gln									

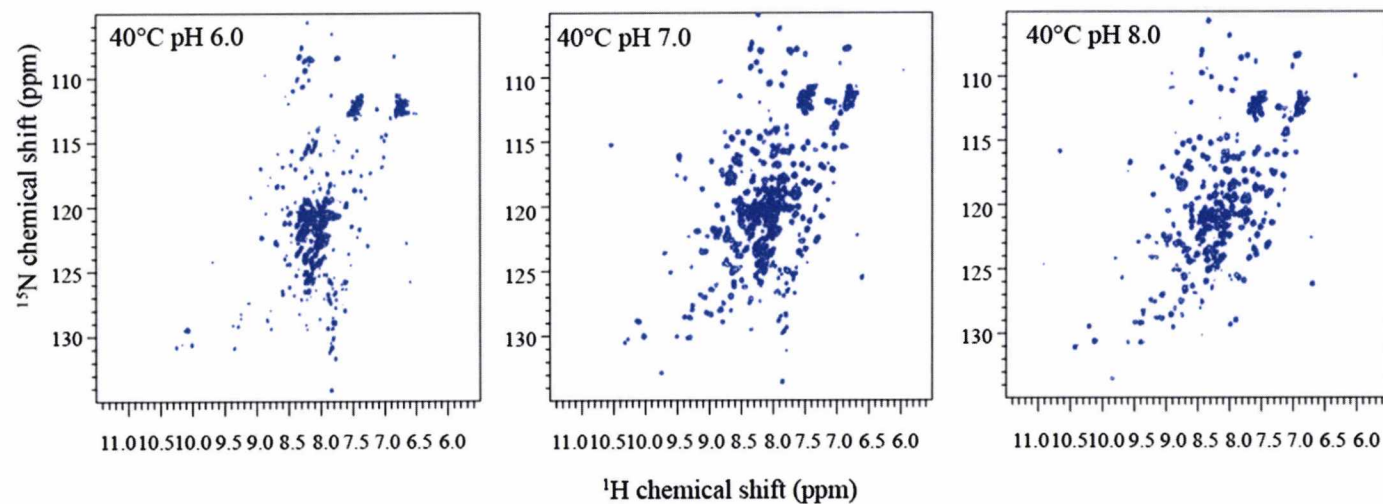
342Glu									
343Leu	0.742676	0.018427	0.122813	0.002842	-0.74836	0.742	0.013		
344Pro									
345Glu	0.670609	0.03025	0.191908	0.019012	-1.07915	0.693	0.03		
346Asp	0.640406	0.021469	0.063267	0.001148	-0.25096	0.959	0.032	4.501	0.475
347Trp									
348Asp			0.187619	0.008322	-0.46736	0.449	0.02		
349Lys	0.664345	0.019117	0.085049	0.00866	-0.29561	0.929	0.026		
350Gln			0.435568	0.025073	-1.97464	0.259	0.01		
351Pro									
352Val	0.651515	0.034384	0.176524	0.009261	-0.84738	0.575	0.022		

Appendix 5.1 b'xa'c $^{15}\text{N}/^1\text{H}$ HSQC spectra for the temperature and pH titrations

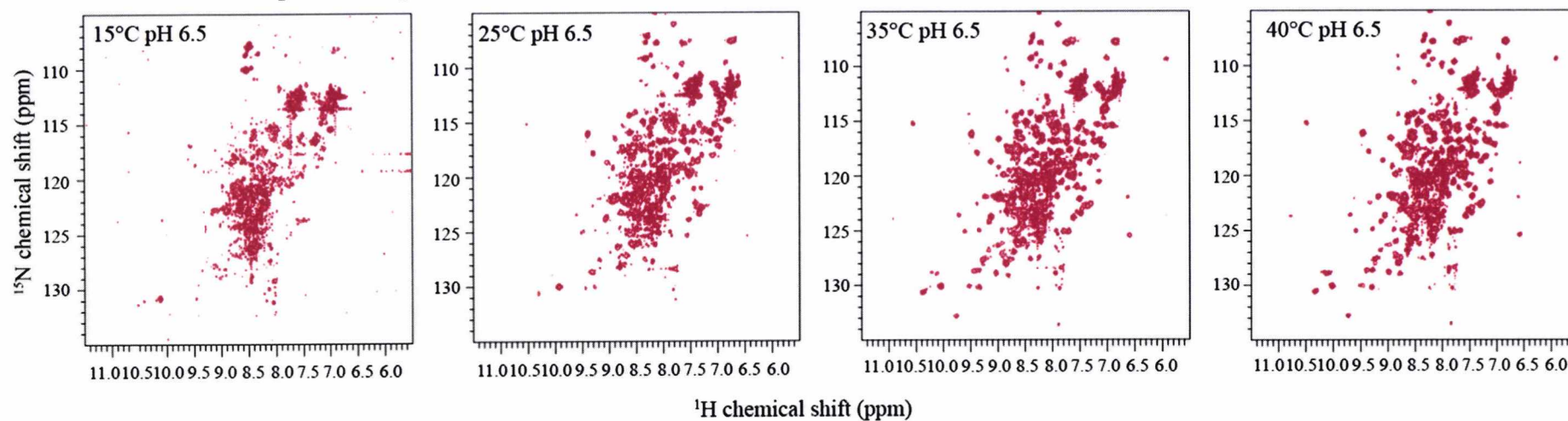
WT b'xa'c temperature experiments



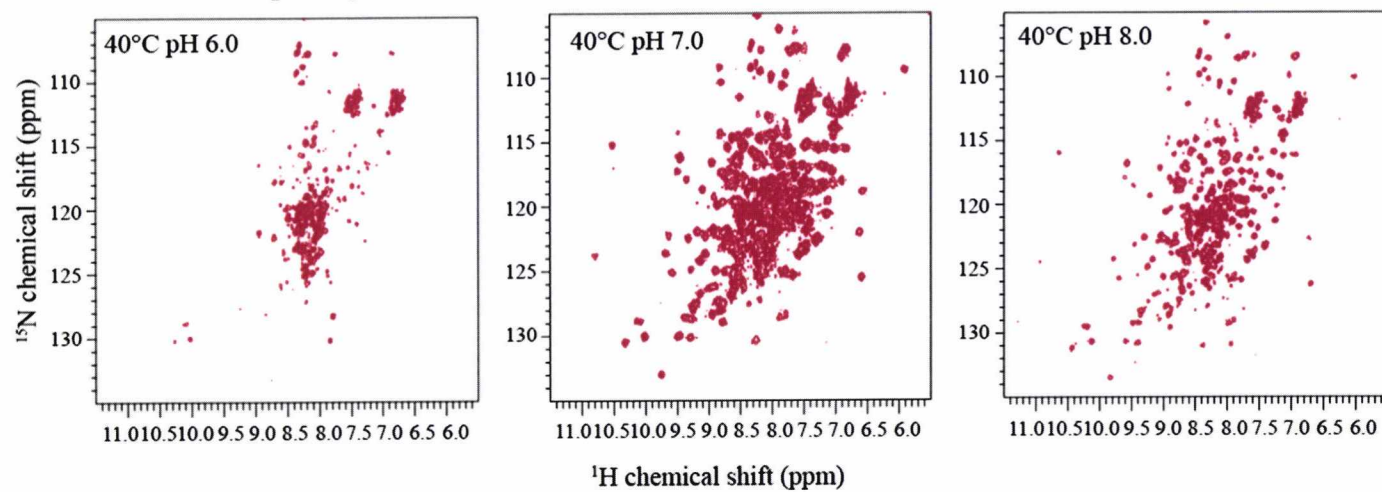
WT b'xa'c pH experiments



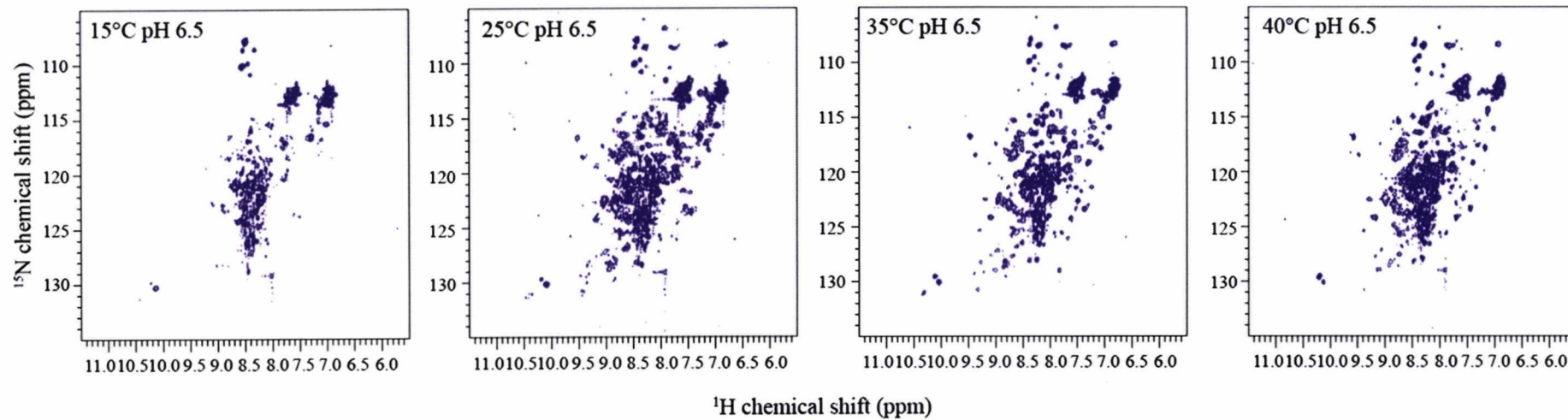
I272A b'xa'c temperature experiments



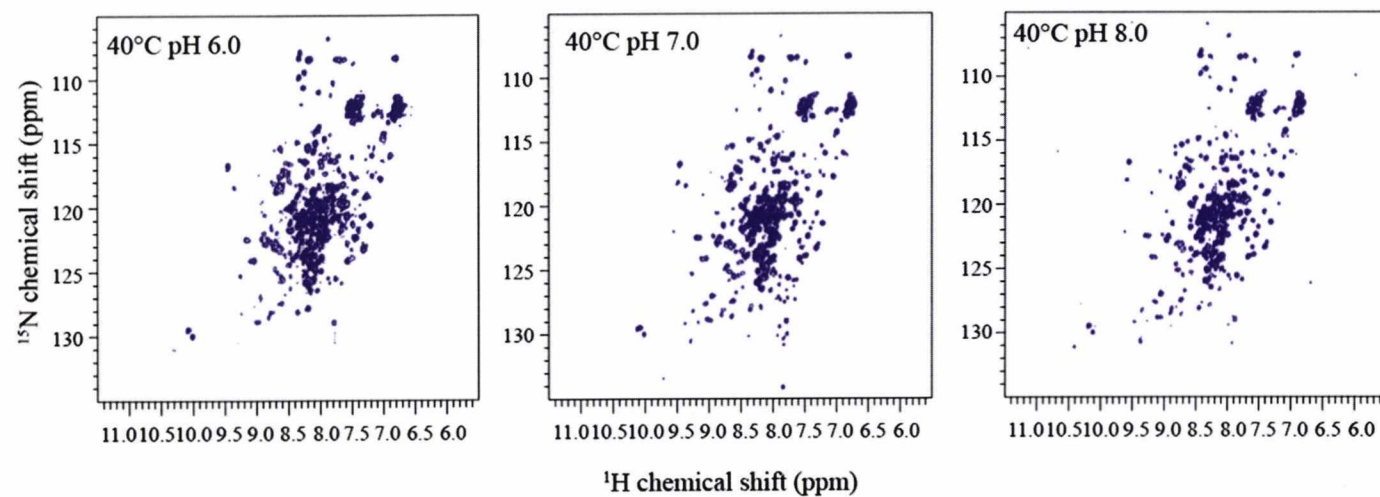
I272A b'xa'c pH experiments



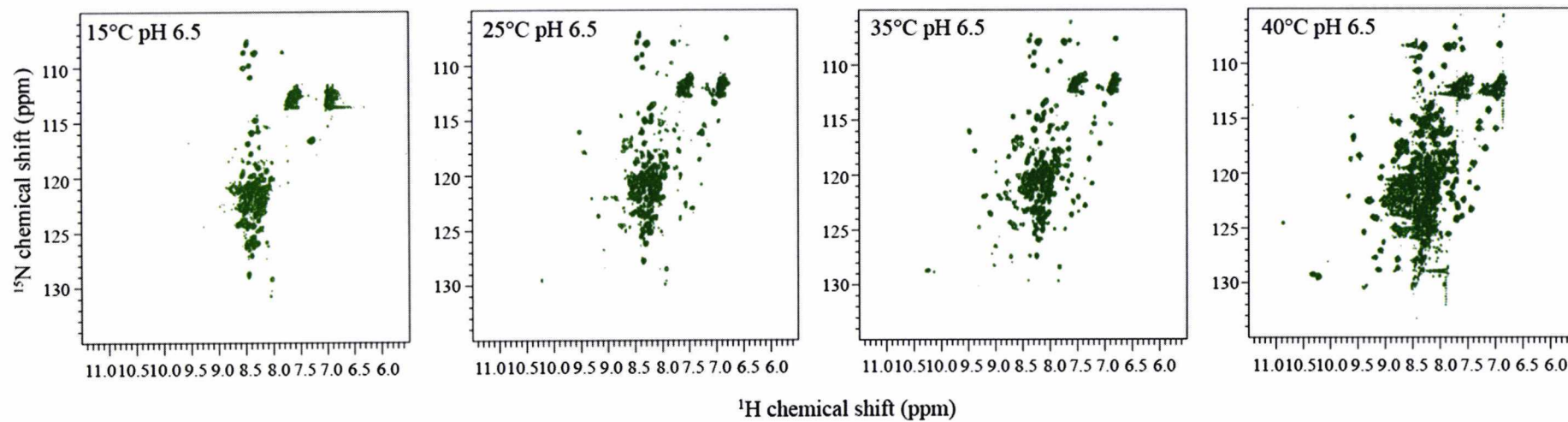
L343A b'xa'c temperature experiments



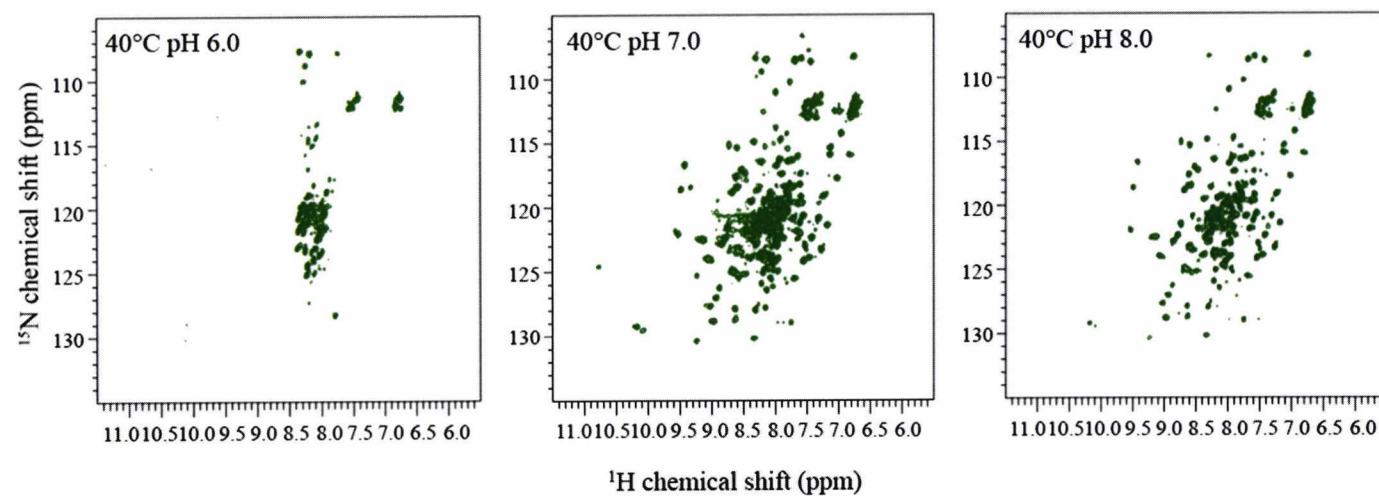
L343A b'xa'c pH experiments



2DA b'xa'c temperature experiments



2DA b'xa'c pH experiments



Appendix 5.2 I272A b'xa'c peak heights for temperature experiments at pH 6.5

Residue	Peak heights			
	15°C	25°C	35°C	40°C
218Pro	14788	36364.7	170378	198749
219Leu				
220Val				
221Ile				
222Glu				
223Phe				
224Thr				
225Glu				
226Gln				
227Thr				
228Ala				
229Pro				
230Lys				
231Ile				
232Phe				
233Gly	-4806.004	8557.7969	61617.4	23963.6
234Gly	78077.2	129759	171906	198841
235Glu	55525	153560	178882	115638
236Ile	21402.5	90251.7	76860.1	242663
237Lys				
238Thr	62518	120779	557863	385236
239His	-1577.086	58199.7	112256	131716
240Ile	13590.8	77102.5	141500	158409
241Leu	70679.9	73377.7	93579	141167
242Leu				
243Phe				
244Leu				
245Pro				
246Lys				
247Ser				
248Val				
249Ser				
250Asp				
251Tyr	56719.1	66458.7	300191	250430
252Asp	10050	160849	305116	365185
253Gly	2306.5928	136729	277538	324026
254Lys	68089.9	-57436.3	208786	234278
255Leu	74614.6	266771	46083.6	191157
256Ser	-2129.363	108964	202355	300391

257Asn				
258Phe	3502.2705	62839.8	132620	158339
259Lys	73600.5	106119	193651	257976
260Thr	128212	277194	277496	354877
261Ala	58591.2	149399	85127.1	315179
262Ala	-418.4941	134279	245978	271564
263Glu	29254.1	199141	390187	557419
264Ser	680631	859547	810728	2381410
265Phe	338191	400411	243890	277609
266Lys	88493.4	192818	172426	405285
267Gly	15258.6	55229.8	74759.4	48807.4
268Lys				
269Ile				
270Leu	15803.7	31716.6	79928.8	81059.4
271Phe				
272Ala				
273Phe				
274Ile				
275Asp				
276Ser				
277Asp				
278His				
279Thr				
280Asp	16325.5	48975	75123.5	77869
281Asn	-2990.107	12898.2	249007	252876
282Gln				
283Arg				
284Ile				
285Leu	-21076.1	2584.3584	39836.9	406015
286Glu				
287Phe				
288Phe	740820	-296344	421636	185786
289Gly	6866.0869	73619.5	166499	189787
290Leu	-30371.2	133788	201546	278191
291Lys	124761	200669	355095	438482
292Lys	47697.3	158495	325097	311874
293Glu	40665.1	147182	318716	368915
294Glu	-538.93701	165940	368365	409445
295Cys	8764.6406	172130	316889	366635
296Pro				
297Ala				
298Val				
299Arg	52206.1	110061	157075	153901
300Leu	22111.4	105445	119753	134319

301Ile	4994.9746	53549.8	62308.3	73835.3
302Thr	15020.6	23247.8	94307	77132.7
303Leu	64748	114025	133215	208362
304Glu	51287.9	60143.6	398096	437356
305Glu	102067	182841	390577	356929
306Glu	19709.9	10156	261696	178064
307Met				
308Thr				
309Lys				
310Tyr				
311Lys				
312Pro				
313Glu	15213.9	103628	154447	110677
314Ser	76267.2	149300	273013	834791
315Glu				
316Glu				
317Leu				
318Thr	72163.7	193067	453748	544982
319Ala	16393.7	67462.8	125387	161370
320Glu	17567.9	163840	221993	192528
321Arg				
322Ile	55281.6	106012	265491	342689
323Thr	77305.9	63664.1	209420	270014
324Glu				
325Phe				
326Cys	22736.8	118224	183460	214486
327His	130480	309648	464601	433373
328Arg				
329Phe				
330Leu				
331Glu	4096.6865	173610	268482	288787
332Gly	48761.3	120086	209461	326116
333Lys	38561.5	43965.2	147973	114762
334Ile	-115124	407488	473102	6102560
335Lys	54019.9	110170	115400	88489.2
336Pro				
337His				
338Leu				
339Met				
340Ser				
341Gln				
342Glu				
343Leu				
344Pro				

345Glu	108471	213661	518380	630088
346Asp	14031.7	183752	372452	389607
347Trp	46993.1	101161	300221	317682
348Asp	10145.4	47806.3	531630	558970
349Lys	8327.7227	205906	292565	274354
350Gln	101001	72232.1	556214	660900
351Pro				
352Val	68302.1	190336	438176	483274
353Lys	-17106.2	56497.9	93712.9	78467.8
354Val				
355Leu				
356Val	-9670.864	70541.1	137222	157653
357Gly	-2587.278	45884.6	101498	120935
358Lys	47154.1	76516.9	217308	215662
359Asn	19040.6	145923	258689	287062
360Phe				
361Glu	48794.8	112016	252066	240123
362Asp	75204	159135	355994	366324
363Val	20373.6	152820	241475	248567
364Ala	58769.2	113913	315181	245262
365Phe	56605	36411.3	149211	160364
366Asp	-24791.6	46981.9	49947.1	77626
367Glu	17519.3	45660.8	128191	168840
368Lys	-6083.778	73638.4	130094	-28990.3
369Lys	19906.8	74870.7	178103	201997
370Asn	6534.9063	69851.1	65916.9	21578.3
371Val				
372Phe	44255.8	55014	132236	182320
373Val	-6933.819	102399	119979	142728
374Glu	-648.3218	51221	139422	166176
375Phe	12730.2	76000.3	140574	127319
376Tyr				
377Ala	-7489.798	49409.4	100317	84364.4
378Pro				
379Trp	11526.8	69937.1	124005	164328
380Cys	18137.6	69770.7	156336	168475
381Gly				
382His				
383Cys				
384Lys	49655	56013.3	311464	265655
385Gln	25749	140037	259364	245578
386Leu	44803	94870.6	181235	193262
387Ala	30244.4	105112	256091	267380
388Pro				

389Ile	53703.3	139707	175145	175812
390Trp	-125.5215	97709.8	184202	254121
391Asp	158148	128878	300463	312179
392Lys	-19211.8	126378	239970	275712
393Leu	41267.8	147247	219537	284942
394Gly	-11536	-4860	191519	122807
395Glu	137967	439405	410013	426665
396Thr	13938.3	71043.9	196424	249970
397Tyr	273.8623	155883	241472	273861
398Lys	56166	241257	121842	321202
399Asp				
400His	-17864.1	130620	329035	410084
401Glu	67819.3	149004	253746	281831
402Asn	46557.2	70694.5	165482	166621
403Ile	2490.6675	64269.9	93831.9	117789
404Val	56527.4	65241.1	120062	138991
405Ile	5737.8652	119695	195796	181699
406Ala	8814.9131	109131	162410	185019
407Lys	68243.6	101365	175890	180118
408Met				
409Asp				
410Ser	64914.7	18207.3	125242	150590
411Thr	-14404.1	55825.7	95220.5	108585
412Ala	7128.917	43357.3	130436	93248.1
413Asn	65606.2	82585.9	393096	508623
414Glu	106376	264147	439091	314148
415Val	2335.8499	115858	264280	269032
416Glu	98763	277585	293882	341189
417Ala	59723.9	196479	422072	405192
418Val	524592	587326	997019	423602
419Lys				
420Val				
421His				
422Ser	40034	152967	260976	211236
423Phe	78378.2	158254	82826.7	297511
424Pro				
425Thr	66742.1	120338	167099	193875
426Leu	74952.6	92665.4	170677	178434
427Lys	-1685.3914	72328.5	239937	210106
428Phe	-2608.853	84372.2	111579	116787
429Phe	10196.6	49759.6	93105.1	134519
430Pro				
431Ala	7140.9053	100296	226879	243521
432Ser	43533.9	2249110	1543060	1284360

433Ala				
434Asp	59254.1	447516	369133	365258
435Arg	85012.5	69220	254481	286010
436Thr	51412.9	175063	472609	467308
437Val	75342.8	249542	183184	441746
438Ile	9802.0147	126333	157066	163914
439Asp	77449.9	177047	336662	349008
440Tyr				
441Asn				
442Gly				
443Glu				
444Arg				
445Thr	22561.9	208900	430591	360779
446Leu	46633.5	213917	507545	307603
447Asp	76531.4	347023	393122	514075
448Gly	47539	70401.1	150444	156241
449Phe	307620	162388	154870	425311
450Lys				
451Lys	119463	264380	377601	351858
452Phe	58498	103142	279474	309347
453Leu				
454Glu				
455Ser	32832.9	83963	189767	217728
456Gly	72188.9	204655	400214	395080
457Gly	58643	146278	336035	374942
458Gln	19835.5	294526	545298	602977
459Asp	10604700	738704	641623	853937
460Gly	674846	1708670	2012490	1654490
461Ala	4403340	7328620	9856780	9919290
462Gly	1212460	2994380	2984650	2352070
463Asp	4059810	7206410	7265680	7394120
464Asp				
465Asp				
466Asp				
467Leu				
468Glu				
469Asp				
470Leu				
471Glu				
472Glu				
473Ala	4725300	7830590	7824400	6611690
474Glu	-77149	281700	30708.5	402952
475Glu				
476Pro				

477Asp				
478Met				
479Glu				
480Glu				
481Asp				
482Asp				
483Asp				
484Gln				
485Lys				
486Ala	11032200	1749280	221412	4681460
487Val	11509700	10636400	9409720	7363360
488Lys	11402100	9364160	6896980	5304180
489Asp	12345300	9577290	6296160	4229060
490Glu	14039700	11990000	9733560	7596650
491Leu	13979900	13731200	14358400	13393400

Appendix 5.3 I272A b'xa'c peak heights for pH experiments at 40°C

Residue	Peak heights			
	pH 6.0	pH 6.5	pH 7.0	pH 8.0
218Pro	107482	198749	106310	125989
219Leu				
220Val				
221Ile				
222Glu				
223Phe				
224Thr				
225Glu				
226Gln				
227Thr				
228Ala				
229Pro				
230Lys				
231Ile				
232Phe				
233Gly	38675.3	23963.6	58474.2	12111.6
234Gly	47678.2	198841	148679	283742
235Glu	53203.3	115638	66766.5	9682.1904
236Ile	82574.6	242663	197271	251487
237Lys				
238Thr	141363	385236	130099	149333
239His	2692.7334	131716	154778	167536
240Ile	44543.1	158409	266853	216632
241Leu	46544.3	141167	180286	161935
242Leu				
243Phe				
244Leu				
245Pro				
246Lys				
247Ser				
248Val				
249Ser				
250Asp				
251Tyr	60845.7	250430	246099	303505
252Asp	61058.6	365185	445414	427566
253Gly	44659.9	324026	275160	275661
254Lys	24034.1	234278	137406	306097
255Leu	47712.8	191157	98901.7	79666.2
256Ser	43040.3	300391	334684	369050
257Asn				

258Phe	47702.1	158339	201473	242805
259Lys	57904.1	257976	300976	324386
260Thr	139636	354877	368391	414499
261Ala	-10913	315179	380406	420140
262Ala	69259.9	271564	283768	325241
263Glu	98661.6	557419	494578	611168
264Ser	1901460	2381410	1440900	1519970
265Phe	34079.5	277609	318314	312616
266Lys	57004.2	405285	375861	436903
267Gly	9844.1699	48807.4	30519.2	22198.3
268Lys				
269Ile				
270Leu	43131.2	81059.4	71360.6	81607.9
271Phe				
272Ala				
273Phe				
274Ile				
275Asp				
276Ser				
277Asp				
278His				
279Thr				
280Asp	-28580.6	77869	58516.1	-22494.7
281Asn	56872.9	252876	209895	240226
282Gln				
283Arg				
284Ile				
285Leu	79095.8	406015	415235	468745
286Glu				
287Phe				
288Phe	51000.3	185786	238515	241858
289Gly	43612.4	189787	188390	242406
290Leu	58836.4	278191	332007	385127
291Lys	244452	438482	333549	389981
292Lys	50707.4	311874	259884	246537
293Glu	45243.3	368915	317510	333168
294Glu	58592.9	409445	485491	499713
295Cys	68105.3	366635	401627	435682
296Pro				
297Ala				
298Val				
299Arg	30419.2	153901	143950	174744
300Leu	7739.5772	134319	106389	120278
301Ile	49136.8	73835.3	69233.8	49717

302Thr	57632.4	77132.7	52450.9	93205.4
303Leu	54679.2	208362	188869	200256
304Glu	42425.4	437356	374654	392211
305Glu	66882.6	356929	196616	121689
306Glu	37010.1	178064	238571	203359
307Met				
308Thr				
309Lys				
310Tyr				
311Lys				
312Pro				
313Glu	58463.8	110677	102905	52409.3
314Ser	158799	834791	561980	512168
315Glu				
316Glu				
317Leu				
318Thr	97883.2	544982	400522	469270
319Ala	58487.7	161370	139004	84454.9
320Glu	55492.3	192528	139898	110410
321Arg				
322Ile	64537.5	342689	295166	278014
323Thr	68403.7	270014	316507	333157
324Glu				
325Phe				
326Cys	54399.1	214486	211977	222451
327His	143955	433373	413179	498785
328Arg				
329Phe				
330Leu				
331Glu	48144.8	288787	310449	349979
332Gly	188887	326116	329994	360177
333Lys	37742.2	114762	221753	223175
334Ile	4808240	6102560	2348720	2446610
335Lys	-788.34375	88489.2	120315	128198
336Pro				
337His				
338Leu				
339Met				
340Ser				
341Gln				
342Glu				
343Leu				
344Pro				
345Glu	142550	630088	412451	342907

346Asp	136879	389607	329224	330606
347Trp	106384	317682	284857	338987
348Asp	161418	558970	425137	456860
349Lys	103479	274354	277870	313551
350Gln	170633	660900	597076	601656
351Pro				
352Val	144757	483274	441802	453973
353Lys	54825.3	78467.8	120734	150776
354Val				
355Leu				
356Val	47502.8	157653	154074	178381
357Gly	50506.4	120935	88198.2	111590
358Lys	146933	215662	194761	227706
359Asn	93181.3	287062	276235	284393
360Phe				
361Glu	115193	240123	239610	254454
362Asp	128949	366324	312442	319504
363Val	77324.6	248567	215264	263054
364Ala	95581.7	245262	280685	267410
365Phe	47656.2	160364	92848.1	115772
366Asp	39297.2	77626	- 16427.8	-19056.6
367Glu	51696.7	168840	126603	94162
368Lys	41136.3	-28990.3	146315	114849
369Lys	70441.1	201997	120398	102025
370Asn	18803.9	21578.3	80431.2	57891.1
371Val				
372Phe	23086.6	182320	130703	124293
373Val	58387.6	142728	172404	137971
374Glu	67983.5	166176	208196	280774
375Phe	61590.4	127319	126476	136096
376Tyr				
377Ala	55906.9	84364.4	105457	144572
378Pro				
379Trp	58663	164328	146139	162624
380Cys	44879.2	168475	150815	203412
381Gly				
382His				
383Cys				
384Lys	78281.8	265655	289088	307177
385Gln	155577	245578	268537	290060
386Leu	43580.1	193262	205816	275881
387Ala	126600	267380	279209	316132
388Pro				
389Ile	90377.2	175812	200879	218260

390Trp	108757	254121	284475	294700
391Asp	105790	312179	321003	342255
392Lys	122577	275712	291787	328546
393Leu	89510.8	284942	274409	277029
394Gly	88785	122807	224071	226872
395Glu	354203	426665	339994	462098
396Thr	105637	249970	241398	207101
397Tyr	114996	273861	238185	281422
398Lys	97246.7	321202	260837	310970
399Asp				
400His	171289	410084	268363	7493.0938
401Glu	113962	281831	204165	194781
402Asn	40928.4	166621	135343	122795
403Ile	39098.4	117789	98388	58341.8
404Val	45369.7	138991	124877	168236
405Ile	52511.9	181699	183475	234620
406Ala	80581.4	185019	204693	208020
407Lys	70111	180118	170182	189617
408Met				
409Asp				
410Ser	62506.5	150590	129618	135132
411Thr	56026.2	108585	97563.7	130586
412Ala	37633	93248.1	105022	87510.6
413Asn	89638.2	508623	506613	435584
414Glu	54629.2	314148	236095	223512
415Val	-2687.9648	269032	245318	254936
416Glu	155397	341189	357034	300062
417Ala	42781.1	405192	305421	277236
418Val	129619	423602	432600	546364
419Lys				
420Val				
421His				
422Ser	69937.1	211236	142230	149270
423Phe	51328.9	297511	209851	193215
424Pro				
425Thr	55829.5	193875	170564	174767
426Leu	70203.2	178434	164240	151812
427Lys	57734.9	210106	291205	255784
428Phe	37014.4	116787	139344	145181
429Phe	65452	134519	143774	114383
430Pro				
431Ala	78936.6	243521	213152	253356
432Ser	41491.5	1284360	253505	242762
433Ala				

434Asp	150595	365258	262339	291892
435Arg	92680	286010	37756.2	35766.9
436Thr	86758.3	467308	250959	257577
437Val	-4736.9883	441746	456325	481135
438Ile	37491.5	163914	159196	184226
439Asp	85256.9	349008	265522	303804
440Tyr				
441Asn				
442Gly				
443Glu				
444Arg				
445Thr	111689	360779	445929	451805
446Leu	153223	307603	373334	321453
447Asp	304649	514075	366120	375069
448Gly	82153.9	156241	182836	210563
449Phe	73969.2	425311	187624	117412
450Lys				
451Lys	89406.8	351858	384480	388132
452Phe	138243	309347	269879	278956
453Leu				
454Glu				
455Ser	256119	217728	124693	151950
456Gly	99194	395080	276299	292443
457Gly	56000.4	374942	246298	244822
458Gln	59462.7	602977	342599	346786
459Asp	-273116	853937	839362	900179
460Gly	333568	1654490	564522	526536
461Ala	3080590	9919290	4704740	5044430
462Gly	623338	2352070	705770	703871
463Asp	3685310	7394120	3281770	3502440
464Asp				
465Asp				
466Asp				
467Leu				
468Glu				
469Asp				
470Leu				
471Glu				
472Glu				
473Ala	2602470	6611690	3103720	2891680
474Glu	164800	402952	195417	219819
475Glu				
476Pro				
477Asp				

478Met				
479Glu				
480Glu				
481Asp				
482Asp				
483Asp				
484Gln				
485Lys				
486Ala	4261190	4681460	1390450	1300300
487Val	5297080	7363360	2658510	2713880
488Lys	4380680	5304180	1596260	1564610
489Asp	3252970	4229060	1154970	1242780
490Glu	6051560	7596650	2587250	2683730
491Leu	7001970	13393400	5764010	6469370

Appendix 5.4 Resonance Assignment for I272A b'xa'c

Residue	H _N	N _H	C _α	H _ε	N _ε
210Met	-	-	-		
211His	-	-	-		
212His	-	-	-		
213His	-	-	-		
214His	-	-	-		
215His	-	-	-		
216His	-	-	-		
217Met	-	-	-		
218Pro	-	-	-		
219Leu	-	-	-		
220Val	-	-	-		
221Ile	-	-	60.99		
222Glu	8.16	127.71	52.25		
223Phe	-	-	-		
224Thr	-	-	-		
225Glu	-	-	-		
226Gln	-	-	-		
227Thr	-	-	-		
228Ala	-	-	-		
29Pro					
230Lys	-	-	-		
231Ile	-	-	-		
232Phe	-	-	50.94		
233Gly	7.63	106.96	45.28		
234Gly	7.62	108.1	44.67		
235Glu	8.43	118.06	54.51		
236Ile	7.76	120.12	53.48		
237Lys	-	-	56.15		
238Thr	6.82	115.96	62.28		
239His	8.86	125.37	54.16		
240Ile	9.15	122.39	59.06		
241Leu	8.83	127.81	52.22		
242Leu	-	-	-		
243Phe	-	-	-		
244Leu	-	-	-		
45Pro			-		
246Lys	-	-	-		
247Ser	-	-	-		
248Val	-	-	-		
49Ser	-	-	-		
250Asp	-	-	52.11		

251Tyr	7.64	120.57	63.08		
252Asp	8.6	115.1	57.46		
253Gly	7.76	110.2	47.02		
254Lys	8	123.01	60		
255Leu	8.06	119.94	57.61		
256Ser	8.35	114.73	62.38		
257Asn	-	-	56.69		
258Phe	-	-	-		
259Lys	8.5	117.07	59.86		
260Thr	8.22	116.49	66.88		
261Ala	7.56	124.15	54.86		
262Ala	7.18	118.77	54		
263Glu	6.99	114.3	58.86		
264Ser	7.42	112	60.74		
265Phe	7.13	115.75	57.33		
266Lys	7.28	123.1	58.63		
267Gly	9.45	114.7	50.59		
268Lys	-	-	-		
269Ile	-	-	60.68		
270Leu	8.22	130.77	54.88		
271Phe	-	-	-		
272Ala	-	-	-		
273Phe	-	-	-		
274Ile	-	-	-		
275Asp	-	-	-		
276Ser	-	-	-		
277Asp	-	-	-		
278His	-	-	-		
279Thr	-	-	65.95		
280Asp	10.75	124.25	56.73		
281Asn	7.92	114.83	52.8		
282Gln	-	-	-		
283Arg	-	-	-		
284Ile	-	-	58.32		
285Leu	7.44	122.33	62.14		
286Glu					
287Phe	-	-	61.02		
288Phe	7.48	112.96	59.48		
289Gly	7.98	110.32	46.66		
290Leu	7.78	121.08	53.57		
291Lys	8.09	118.72	54.11		
292Lys	9.03	124.1	61.03		
293Glu	9.43	116.66	58.96		
294Glu	7.62	118.04	55.85		

295Cys	7.03	117.71	58.47		
296Pro			-		
297Ala	-	-	-		
298Val	-	-	59.61		
299Arg	8.81	122.66	53.24		
300Leu	8.72	125.49	54.15		
301Ile	9.6	122.72	58.48		
302Thr	8.79	114.76	59.19		
303Leu	8.56	125.34	54.31		
304Glu	7.45	123.88	56.65		
305Glu	8.45	122.38	59.41		
306Glu	7.88	116.39	62.43		
307Met	-	-	-		
308Thr	-	-	-		
309Lys	-	-	-		
310Tyr	-	-	-		
311Lys	-	-	-		
312Pro			62.16		
313Glu	8.89	119.76	57.88		
314Ser	7.49	111.7	56.55		
315Glu	-	-	-		
316Glu	-	-	-		
317Leu	-	-	53.14		
318Thr	6.79	108.3	59.61		
319Ala	9.27	125.34	56.17		
320Glu	9.33	118.36	60.49		
321Arg	-	-	58.06		
322Ile	8.48	121.75	65.44		
323Thr	8.54	116.95	68.14		
324Glu	-	-	-		
325Phe	-	-	59.54		
326Cys	7.7	115.97	64.54		
327His	8.57	118.08	60.54		
328Arg	-	-	-		
329Phe	-	-	-		
330Leu	-	-	57.48		
331Glu	7.71	116.34	56.36		
332Gly	7.68	108.41	46.22		
333Lys	7.93	116.53	56.18		
334Ile	7.93	120.69	56.68		
335Lys	8.95	126.74	50.88		
336Pro			-		
337His	-	-	-		
338Leu	-	-	-		

339Met	-	-	-		
340Ser	-	-	-		
341Gln	-	-	-		
342Glu	-	-	-		
343Leu	-	-	-		
344Pro			62.86		
345Glu	-	-	58.89		
346Asp	8.43	115.59	52.64		
347Trp	7.58	119.61	60.73	10.06	129.32
348Asp	7.46	118.61	51.93		
349Lys	7.36	118.69	56.94		
350Gln	7.88	117.32	53.34		
351Pro			66.31		
352Val	7.1	112.39	61.31		
353Lys	9.23	128.35	55.06		
354Val	-	-	-		
355Leu	-	-	51.8		
356Val	8.48	112.02	58.35		
357Gly	8.8	109.66	48.05		
358Lys	8	114.8	58.78		
359Asn	7.41	115.53	51.93		
360Phe	-	-	63.37		
361Glu	8.93	117.05	60.73		
362Asp	7.64	117.27	56.58		
363Val	6.99	115.99	64.39		
364Ala	8.1	119.01	55.39		
365Phe	6.87	108.86	56.19		
366Asp	6.54	119.33	54.54		
367Glu	8.92	128.77	57.82		
368Lys	8.68	115.8	55.52		
369Lys	7.23	116	53.38		
370Asn	9.01	122.83	53.47		
371Val	-	-	61.47		
372Phe	9.72	133.41	52.31		
373Val	9.56	125.56	61.53		
374Glu	7.76	125.51	53.83		
375Phe	9.46	130.48	57.43		
376Tyr	-	-	55.14		
377Ala	7.09	120.05	47.4		
378Pro			64.7		
379Trp	5.89	109.88	53.72	10.3	130.97
380Cys	6.56	125.93	60.24		
381Gly	-	-	-		
382His	-	-	-		

383Cys	-	-	54.86		
384Lys	8.54	123.5	60.36		
385Gln	7.5	117.26	58.4		
386Leu	7.4	117.53	55.41		
387Ala	7.38	121.68	57.25		
388Pro			65.91		
389Ile	7.28	119.23	64.38		
390Trp	8.53	123.85	60.31	9.99	130.48
391Asp	8.28	116.13	57.51		
392Lys	7.49	121.38	58.38		
393Leu	8.57	126.62	58.28		
394Gly	8.19	105.57	47.27		
395Glu	8.09	120.99	60.03		
396Thr	7.96	115.82	66.93		
397Tyr	7.19	117.32	59.56		
398Lys	7.25	122.91	59.85		
399Asp	-	-	58.21		
400His	7.86	122.24	59.57		
401Glu	8.45	126.21	59.3		
402Asn	10.49	115.69	54.81		
403Ile	8.88	126.68	-		
404Val	8.75	129.4	61.4		
405Ile	9.27	129.12	57.96		
406Ala	9.22	128.15	50.35		
407Lys	9.07	119.14	54.68		
408Met	-	-	-		
409Asp	-	-	52.25		
410Ser	8.74	127.94	61.8		
411Thr	8.77	110.77	62.58		
412Ala	6.6	122.48	50.51		
413Asn	7	114.38	52.38		
414Glu	8.67	122.3	54.99		
415Val	8.5	114.96	58.91		
416Glu	8.67	122.76	59.16		
417Ala	8.52	116.55	54.3		
418Val	6.73	112.26	60.42		
419Lys	-	-	-		
420Val	-	-	-		
421His	-	-	55.44		
422Ser	7	113.08	56.68		
423Phe	8.49	120.04	54.13		
424Pro			63.22		
425Thr	8.76	117.77	63.62		
426Leu	9.26	130.59	53.76		

427Lys	8.84	120.4	54.55		
428Phe	9.66	124.09	54.82		
429Phe	9.11	127.19	53.78		
430Pro			62.23		
431Ala	7.64	125.78	52.24		
432Ser	8.01	119.4	58.02		
433Ala	-	-	53.13		
434Asp	8.09	117.06	53.99		
435Arg	8.15	116.7	57.29		
436Thr	7.45	115.74	63.8		
437Val	8.21	124.13	61.43		
438Ile	9.36	129.04	59.93		
439Asp	8.78	128.47	54.93		
440Tyr	-	-	-		
441Asn	-	-	-		
442Gly	-	-	-		
443Glu	-	-	-		
444Arg	-	-	55.54		
445Thr	6.81	108.3	57.52		
446Leu	8.9	122.28	58.91		
447Asp	8.16	115.16	57.58		
448Gly	8.15	109.94	46.98		
449Phe	8.45	120.64	51.81		
450Lys	-	-	60.61		
451Lys	7.98	118.33	59.65		
452Phe	7.83	120.15	60.76		
453Leu	-	-	-		
454Glu	-	-	64.31		
455Ser	7.37	111.26	58.41		
456Gly	7.82	111.1	45.92		
457Gly	7.86	106.71	44.8		
458Gln	7.15	117.42	55.69		
459Asp	8.3	121.16	54.29		
460Gly	8.3	109.68	45.32		
461Ala	8.07	123.02	52.56		
462Gly	8.29	107.83	45		
463Asp	8.15	120.45	54.38		
464Asp	-	-	-		
465Asp	-	-	-		
466Asp	-	-	-		
467Leu	-	-	-		
468Glu	-	-	-		
469Asp	-	-	-		
470Leu	-	-	-		

471Glu	-	-	-		
472Glu	-	-	56.27		
473Ala	8.22	125.79	52.26		
474Glu	7.85	125.44	58.12		
475Glu	-	-	-		
476Pro					
477Asp	-	-	-		
478Met	-	-	-		
479Glu	-	-	-		
480Glu	-	-	-		
481Asp	-	-	-		
482Asp	-	-	-		
483Asp	-	-	-		
484Gln	-	-	-		
485Lys	-	-	56.65		
486Ala	8	124.03	52.5		
487Val	7.89	119.23	62.44		
488Lys	8.2	124.96	56.09		
489Asp	8.18	121.71	54.48		
490Glu	8.14	121.07	56.37		
491Leu	7.76	128.84	56.87		

Appendix 5.5 Relaxation and Model-free data for WT b'xa'c

Residues	T ₁	T ₁ error	T ₂	T ₂ error	hetNOE	S ²	S ² error	R _{ex}	R _{ex} error
218Pro	0								
219Leu									
220Val									
221Ile									
222Glu	0.99211	0.12105	0.10114	0.0075223	-1.3136	0.649	0.042		
223Phe									
224Thr									
225Glu									
226Gln									
227Thr									
228Ala									
229Pro									
230Lys									
231Ile									
232Phe									
233Gly	0.79356	0.23042			-0.36619				
234Gly	1.0005	0.079609	0.058706	0.0074703	-0.51051	0.871	0.059		
235Glu	1.0245	0.12528	0.028875	0.010193	-0.60437	0.81	0.099		
236Ile	0.75935	0.081579	0.086733	0.01259	-0.74002	0.879	0.078		
237Lys									
238Thr	0.79767	0.085209	0.075715	0.015559					
239His									
240Ile	1.0507	0.1318	0.13098	0.034672	-0.15379	0.667	0.078		
241Leu					-0.53112			2.594	0
242Leu									
243Phe									
244Leu									
245Pro									
246Lys									
247Ser									
248Val									
249Ser									
250Asp									
251Tyr	0.81686	0.11194	0.10252	0.0097837	-0.98247	0.664	0.053		
252Asp	0.98323	0.080002	0.051871	0.0078813	-0.33819	0.841	0.068	5.701	3.131
253Gly	0.96432	0.088563	0.046274	0.0064524	-0.072001	0.858	0.079	7.767	3.271
254Lys									
255Leu	0.78201	0.049803	0.10903	0.0092645	-1.0534	0.736	0.039		
256Ser	0.86188	0.039667	0.049723	0.0055461				4.623	2.354
257Asn									
258Phe	0.87995	0.14978	0.17144	0.020769					

259Lys	0.81612	0.042571	0.020848	0.010855	-0.32578				
260Thr	1.0102	0.058021	0.10309	0.012485	-0.57991	0.754	0.039		
261Ala	0.84647	0.044285	0.1198	0.0067194	-0.32993	0.851	0.043		
262Ala	0.96508	0.11195	0.087834	0.010767					
263Glu	0.86344	0.044061	0.026641	0.0037153	-0.22524	0.961	0.049		
264Ser	1.3291	0.064228			-1.3671				
265Phe	0.91342	0.064275	0.04197	0.0069279	-0.22276	0.905	0.064	5.214	4.463
266Lys	0.97846	0.056858	0.032474	0.0054217	-0.3417	0.845	0.049	17.151	5.202
267Gly					-1.1888				
268Lys									
269Ile									
270Leu	0.94595	0.11257	0.12293	0.047308	-0.35218				
271Phe					-0.49217				
272Ile									
273Phe									
274Ile									
275Asp					-0.5571				
276Ser									
277Asp									
278His									
279Thr									
280Asp	1.0649	0.1783	0.21957	0.092306		0.507	0.088		
281Asn	1.0229	0.17241	0.04554	0.0058454		0.808	0.136	8.91	3.575
282Gln									
283Arg									
284Ile									
285Leu	0.99525	0.060312	0.11056	0.010431					
286Glu									
287Phe									
288Phe	1.2768	0.065623	0.047797	0.0040579	-1.6333	0.702	0.032		
289Gly	1.0096	0.093147	0.090914	0.038053	-0.43443	0.81	0.073		
290Leu	0.83526	0.074327	0.10318	0.011013	-0.68087	0.735	0.052		
291Lys	0.8556	0.045865	0.12828	0.0077118	-1.2064	0.598	0.025		
292Lys	1.118	0.2344	0.059117	0.0056464	-0.096276	0.957	0.084		
293Glu	1.0954	0.080545	0.045421	0.0039921	-0.20724	0.755	0.056	9.829	2.132
294Glu	0.95736	0.049659	0.048727	0.0049742	-0.41181	0.864	0.045	6.578	2.216
295Cys	1.2096	0.068472	0.048729	0.0062528	-0.34965	0.715	0.038		
296Pro									
297Ala									
298Val									
299Arg	0.9324	0.1064	0.15515	0.060265	-0.18132	0.741	0.085		
300Leu									
301Ile									
302Thr	1.2183	0.30929	0.11532	0.045532	-0.25084	0.622	0.134		

303Leu	1.0778	0.077281	0.080436	0.03164					
304Glu	1.0666	0.16727	0.16039	0.012683	-0.29898	0.409	0.03		
305Glu									
306Glu	1.0977	0.094368	0.22114	0.054847	-0.38956	0.524	0.046		
307Met									
308Thr									
309Lys									
310Tyr									
311Lys									
312Pro									
313Glu					-0.14229				
314Ser	0.94407	0.043458	0.090095	0.0072319	-1.3267	0.81	0.033		
315Glu									
316Glu									
317Leu									
318Thr	1.0963	0.10815	0.039668	0.0036496	-0.75503	0.754	0.074	13.033	2.612
319Ala	0.70969	0.14485	0.34493	0.11714	-0.2418	0.378	0.104		
320Glu	1.1591	0.253	0.19288	0.060995	-0.32598	0.494	0.093		
321Arg									
322Ile	0.71348	0.35242						12.098	0
323Thr	0.8792	0.08328	0.040406	0.0068463	-0.22283	0.941	0.089	9.565	4.433
324Glu									
325Phe									
326Cys	1.1785	0.11715	0.042222	0.0051707	-0.33721	0.702	0.07	12.357	3.111
327His	1.0392	0.070064	0.04847	0.0016553	-0.38523	0.796	0.054	9.91	2.36
328Arg									
329Phe									
330Leu									
331Glu	1.1934	0.12377	0.051851	0.0037575	-0.37259	0.898	0.055		
332Gly	0.90307	0.040485	0.10299	0.0075739	-0.62689	0.77	0.03		
333Lys	0.8435	0.028492	0.063529	0.0065259	-0.59326	0.98	0.031		
334Ile	0.85594	0.028172	0.35393	0.018936	-1.9029	0.238	0.009		
335Lys	1.1968	0.38712	0.071571	0.013738	-0.26501	0.803	0.133		
336Pro									
337His									
338Leu									
339Met									
340Ser									
341Gln									
342Glu									
343Leu									
344Pro									
345Glu	0.88303	0.13196	0.046296	0.0016063	-0.52435	0.937	0.14	6.479	2.38
346Asp	1.0891	0.13074	0.052679	0.0074905	-0.2413	0.855	0.08		

347Trp	0.94337	0.076485	0.052811	0.0048633	-0.64544	0.877	0.071	4.784	2.087
348Asp	1.1531	0.090761	0.035723	0.0070816	-0.21356	0.717	0.056	16.416	5.624
349Lys	1.3013	0.41727	0.049722	0.010718	-0.37722	0.858	0.162		
350Gln	1.0741	0.084316	0.071932	0.012941					
351Pro									
352Val	1.2119	0.045782	0.048498	0.0096212	-0.33482	0.688	0.026		
353Lys	1.3368	0.29968	0.064931	0.01161	-0.42725	0.701	0.11		
354Val									
355Leu									
356Val	0.63887	0.1369			-0.37508				
357Gly	1.0961	0.34386							
358Lys	0.86562	0.065038	0.078749	0.002286	-1.1737	0.802	0.022		
359Asn	1.2333	0.16863	0.072629	0.010631	-0.56724	0.696	0.087		
360Phe									
361Glu	1.1264	0.29177			-0.42269	0.693	0.177		
362Asp	1.2069	0.14497	0.054832	0.0073382					
363Val	1.1765	0.28028	0.059852	0.01431	-0.55406	0.807	0.139		
364Ala			0.11498	0.0061982	-1.1892			8.698	0
365Phe	1.8219	0.17476			-0.58089	0.452	0.043		
366Asp					-1.0903				
367Glu	0.84002	0.079287			-0.1107	0.983	0.093		
368Lys									
369Lys	1.0577	0.14497	0.052911	0.018405	-0.13213				
370Asn									
371Val									
372Phe	1.4338	0.89937							
373Val	1.0232	0.47168	0.086308	0.017138					
374Glu	0.89247	0.075586	0.18312	0.039977					
375Phe	0.88853	0.25861	0.10452	0.013905	-0.10883	0.646	0.073		
376Tyr									
377Ala	1.4062	0.5119	0.12899	0.022429	-0.60216	0.623	0.137		
378Pro									
379Trp	1.1687	0.31805	0.070187	0.014855	-0.23123	0.672	0.157		
380Cys					-0.36345				
381Gly									
382His									
383Cys									
384Lys									
385Gln	1.437	0.27107	0.083444	0.017428	-0.21541	0.633	0.097		
386Leu	2.0341	0.27221			-0.37867				
387Ala	1.0685	0.11462	0.071216	0.011494	-0.46172	0.799	0.071		
388Pro									
389Ile	1.0263	0.081246	0.083771	0.010395	-0.29715	0.784	0.052		
390Trp									

391Asp	1.161	0.22369	0.10575	0.012172	-0.14315	0.604	0.083		
392Lys	1.2829	0.13001	0.060034	0.012929	-0.25636	0.675	0.063		
393Leu	1.0382	0.099279	0.073384	0.014522	-0.2212	0.805	0.069		
394Gly	0.87305	0.10197			-0.44404				
395Glu									
396Thr	1.1881	0.40067	0.058473	0.0032443	-0.30846				
397Tyr	1.1108	0.12498	0.045955	0.0030699	-0.77274				
398Lys	1.275	0.1049	0.016868	0.0084328	-0.21719	0.65	0.053		
399Asp									
400His	0.92277	0.054724	0.11993	0.009057	-0.73414	0.649	0.031		
401Glu	1.5058	0.31424			-0.059975	0.559	0.112		
402Asn	1.6236	0.6579			-0.37738				
403Ile	1.1693	0.18471			-0.29747	0.89	0.085		
404Val	0.66511	0.065521			-0.44768				
405Ile	1.0331	0.3382							
406Ala	1.0234	0.16807	0.091074	0.033657	-0.49521	0.78	0.117		
407Lys	1.0613	0.16278			-0.11269				
408Met									
409Asp									
410Ser	1.2029	0.21837	0.14945	0.037517	-0.56782	0.526	0.08		
411Thr	1.0914	0.18891	0.088316	0.028955	-0.1791	0.758	0.126		
412Ala					-1.0071				
413Asn	0.97406	0.041264	0.060735	0.0072325	-0.15391	0.863	0.034		
414Glu	0.7526	0.19295	0.053783	0.0052821				2.449	0
415Val	1.2189	0.090819			-0.17129				
416Glu	1.0398	0.070024	0.087065	0.013724	-0.23866	0.78	0.048		
417Ala	1.1498	0.17273	0.035429	0.0056713	-0.24829	0.719	0.108	16.615	4.843
418Val	1.268	0.059405							
419Lys									
420Val									
421His									
422Ser	1.3937	0.3335	0.10702	0.029127	-0.53481	0.587	0.105		
423Phe	1.1461	0.092273	0.088797	0.0093495	-0.30192	0.712	0.046		
424Pro									
425Thr	1.4921	0.2806			-0.10206				
426Leu	1.2603	0.229	0.083848	0.016249	-0.3166	0.69	0.092		
427Lys	1.2444	0.27012	0.18013	0.037522	-0.50811	0.407	0.064		
428Phe	1.1167	0.37634			-0.35003				
429Phe	1.1721	0.47637	0.06514	0.011594	-0.40525	0.888	0.146		
430Pro									
431Ala	1.1589	0.17923	0.12495	0.062812	-0.0042694	0.683	0.104		
432Ser	0.6286	0.094512							
433Ala									
434Asp	1.2414	0.12759	0.25163	0.054104	-0.0074083	0.403	0.042		

435Arg									
436Thr	1.0819	0.071161	0.063393	0.010407	-0.60882	0.783	0.048		
437Val									
438Ile	1.0124	0.10535			-0.073806				
439Asp	1.0495	0.088697	0.077382	0.0074361	-0.095012	0.775	0.056		
440Tyr									
441Asn									
442Gly									
443Glu									
444Arg									
445Thr	1.0362	0.11081	0.044931	0.0061527	-0.44135	0.798	0.085	9.373	3.345
446Leu	1.3364	0.17306	0.050475	0.0048935	-0.51247	0.809	0.066		
447Asp	0.86537	0.048147	0.098431	0.0030409	-0.70039	0.668	0.018		
448Gly	0.84797	0.15456							
449Phe	0.84279	0.07537	0.10121	0.013285	-0.56231	0.78	0.059		
450Lys									
451Lys	0.87742	0.084145	0.26823	0.082713	-0.12272	0.936	0.089		
452Phe	0.82392	0.10375	0.070744	0.0096195	-0.7065	0.936	0.087		
453Leu									
454Glu									
455Ser	1.0384	0.17997			-0.59076				
456Gly	0.96597	0.091511	0.067142	0.0091716	-0.57036	0.876	0.068		
457Gly	1.202	0.095062	0.042238	0.004319	-0.55051	0.688	0.054	12.57	2.575
458Gln	0.96649	0.13888	0.052968	0.0054436	-0.87738				
459Asp	0.82345	0.030751	0.31921	0.016357	-1.4891	0.247	0.01		
460Gly	0.9344	0.025398	0.1418	0.0082891	-1.1858	0.674	0.018		
461Ala	0.87186	0.018646	0.45757	0.017492	-1.6364	0.185	0.005		
462Gly	1.066	0.057161	0.16733	0.0046332	-1.734	0.393	0.01		
463Asp	0.90949	0.030061	0.39054	0.020392	-1.793	0.212	0.008		
464Asp									
465Asp									
466Asp									
467Leu									
468Glu									
469Asp									
470Leu									
471Glu									
472Glu									
473Ala	0.96367	0.031816	0.47231	0.027498	-1.9469	0.18	0.007		
474Glu	1.3887	0.18141			-3.402				
475Glu									
476Pro									
477Asp									
478Met									

479Glu									
480Glu									
481Asp									
482Asp									
483Asp									
484Gln									
485Lys									
486Ala	0.94141	0.046929	0.48199	0.041644	-2.3314	0.174	0.011		
487Val	0.95759	0.030645	0.54217	0.041609	-2.1895	0.183	0.008		
488Lys	1.0409	0.052068	0.51881	0.048477	-2.5121	0.169	0.011		
489Asp	1.1415	0.10401	0.51699	0.05565	-3.1051	0.142	0.013		
490Glu	1.3893	0.08645	0.62768	0.050732	-3.4986	0.121	0.008		
491Leu	1.6309	0.0099969	0.4129	0.0079486	-3.2753	0.315	0.002		

Appendix 5.6 I272A b'xa'c Relaxation and Model-free data

Residue	T ₁ /s	T ₁ error/s	T ₂ /s	T ₂ error/s	hetNOE	S ²	S ² error	R _{ex}	R _{ex} error
218Pro									
219Leu									
220Val									
221Ile									
222Glu	0.8657885	0.053061	0.0917765	0.001964	-1.2511985	0.708	0.014		
223Phe									
224Thr									
225Glu									
226Gln									
227Thr									
228Ala									
229Pro									
230Lys									
231Ile									
232Phe									
233Gly					-0.2588566				
234Gly	0.9949051	0.0681721	0.0471373	0.0017174	-0.4392863	0.811	0.056	8.463	1.167
235Glu	1.0100655	0.0853169			-0.3605421				
236Ile	0.7872179	0.0340933	0.0621799	0.0011108	-0.7523004			0.352	0
237Lys									
238Thr	1.1463649	0.1020008	0.038392	0.0049043	-0.1367723	0.704	0.063	14.981	3.47
239His	0.9057724	0.0448839			-0.3243161				
240Ile	0.8699022	0.0345871	0.035892	0.0030921	-0.1406389	0.927	0.037	13.278	2.469
241Leu	1.0207397	0.0851653	0.0529146	0.0060272	-0.2613754	0.868	0.059		
242Leu									
243Phe									
244Leu									
245Pro									
246Lys									
247Ser									
248Val									
249Ser									
250Asp									
251Tyr	0.9455891	0.0354998	0.0523416	0.0029719	-0.4676438	0.853	0.032	5.689	1.196
252Asp	0.9952644	0.0302875	0.0400069	0.0022811	-0.3794213	0.81	0.025	12.249	1.477
253Gly	0.9552576	0.0453635	0.0498527	0.0020749	-0.1734317	0.844	0.04	6.779	1.046
254Lys					-7.5418566				
255Leu					-0.6651447				
256Ser	0.9106205	0.0392055	0.0400099	0.0035206	-0.292589	0.886	0.038	11.062	2.28
257Asn									
258Phe	0.9576917	0.0513447	0.0844818	0.0078336	-0.6297284	0.816	0.038		

259Lys	0.9998992	0.0441079	0.0383011	0.0022971	-0.2009912	0.807	0.036	13.421	1.663
260Thr	0.9710654	0.0417225	0.0869663	0.0044628	-0.0759433	0.783	0.026		
261Ala	0.9357848	0.0503793	0.0548389	0.0041117	-0.4111466	0.928	0.041		
262Ala	0.9762813	0.0305991	0.0509315	0.0022359	-0.3211257	0.826	0.026	6.64	0.953
263Glu	1.0354593	0.0426376	0.0385205	0.0019808	-0.4364083	0.779	0.032	13.708	1.427
264Ser	1.12444	0.0494378			-0.7744348				
265Phe	0.9778937	0.0376882	0.0433346	0.0021642	-0.3410691	0.825	0.032	10.103	1.256
266Lys	1.0723055	0.0313918	0.035724	0.002604	-0.3225178	0.752	0.022	16.162	2.07
267Gly					-0.0426366				
268Lys									
269Ile									
270Leu	0.897747	0.0482918			-0.2820848				
271Phe									
272Ala									
273Phe									
274Ile									
275Asp									
276Ser									
277Asp									
278His									
279Thr									
280Asp	1.0255851	0.1285446	0.0392201	0.0047813	-0.3602967	0.786	0.099	13.127	3.474
281Asn	0.9531363	0.0450259	0.0658384	0.0098493	-0.2389923	0.855	0.039		
282Gln									
283Arg									
284Ile									
285Leu	0.9868001	0.038995	0.0647563	0.0040541	-0.4914541	0.853	0.029		
286Glu									
287Phe									
288Phe	0.766224	0.0269535	0.0483416	0.0040633	-0.2202869			4.956	0
289Gly	0.9992084	0.020387	0.0416873	0.0062295	-0.2220481	0.807	0.016	11.292	3.594
290Leu	1.04783	0.0319296	0.0571885	0.0052952	-0.6011904	0.787	0.023		
291Lys	0.8896019	0.0357137	0.1462808	0.0053385	-0.9486658	0.51	0.015		
292Lys	1.0895562	0.0492406	0.0395559	0.0022116	-0.2647635	0.74	0.033	13.637	1.508
293Glu	1.0156692	0.0370207	0.0393282	0.0018062	-0.3085213	0.794	0.029	12.937	1.253
294Glu	0.9433703	0.0302011	0.0449972	0.000726	-0.2857793	0.855	0.027	8.776	0.56
295Cys	1.0232564	0.019741	0.0448465	0.0012591	-0.9970792	0.788	0.015	9.901	0.67
296Pro									
297Ala									
298Val									
299Arg	1.0651301	0.0651218	0.1024153	0.0085612	-0.1198473	0.697	0.035		
300Leu	1.0406508	0.151602			-0.3477844				
301Ile	1.1928372	0.1969139							
302Thr	1.2572278	0.198674			-0.4293677				

303Leu	1.0266981	0.0775563	0.0419659	0.0046949	-0.3487878	0.786	0.059	11.473	2.825
304Glu	0.9520052	0.0486742	0.0623817	0.003273	-0.6268159	0.915	0.034		
305Glu	0.816866	0.0975797							
306Glu	1.1114446	0.0656397	0.0492332	0.0025922	-0.5115295	0.886	0.036		
307Met									
308Thr									
309Lys									
310Tyr									
311Lys									
312Pro									
313Glu	1.4974056	0.28972			-0.3516378				
314Ser	0.9979275	0.0394312	0.0825379	0.0067108	-1.2541972	0.8	0.028		
315Glu									
316Glu									
317Leu									
318Thr	1.1191403	0.0453063	0.0407427	0.0014197	-0.5530079	0.721	0.029	13.208	0.971
319Ala	0.969615	0.1814147	0.0728147	0.0139437	-0.2171858	0.851	0.114		
320Glu	0.9075377	0.0782389	0.0335049	0.0047274	-0.2509594	0.889	0.077	15.867	4.38
321Arg									
322Ile	0.8270934	0.0663408			-0.07002				
323Thr	0.9744621	0.0486349	0.0352891	0.0010343	-0.4699842	0.828	0.041	15.319	1.054
324Glu									
325Phe									
326Cys	1.1590752	0.0879978	0.0412052	0.0020994	-0.2957445	0.696	0.053	13.324	1.49
327His	1.0213101	0.0454901	0.03177	0.0015343	-0.3060627	0.79	0.035	19.055	1.618
328Arg									
329Phe									
330Leu									
331Glu	0.9878673	0.0462812	0.0423009	0.0018347	-0.2404001	0.816	0.038	10.798	1.189
332Gly	0.8673691	0.0229535	0.0991598	0.0041585	-0.6164224	0.798	0.018		
333Lys	0.9309263	0.0560109	0.075638	0.0065925	-0.0682911	0.858	0.042		
334Ile	0.8601859	0.0324615	0.3229947	0.0117143	-1.5204898	0.226	0.007		
335Lys	0.9440131	0.1136997	0.0371042	0.0012738	-0.1899617	0.854	0.103	13.513	1.864
336Pro									
337His									
338Leu									
339Met									
340Ser									
341Gln									
342Glu									
343Leu									
344Pro									
345Glu	1.1176688	0.0569891	0.0530541	0.0036796	-0.2269566	0.8	0.034		
346Asp	1.1058655	0.0655001			-0.4161433				

347Trp	0.9355902	0.0415801	0.1441733	0.018993	-0.6945023	0.734	0.032		
348Asp	1.1779332	0.0560017	0.0473924	0.004528	-0.2993048	0.724	0.032		
349Lys	1.2594959	0.0621607							
350Gln	1.1415756	0.0104831	0.0905245	0.0169717	-0.1818991	0.706	0.006		
351Pro									
352Val	1.1807569	0.0709493	0.0370979	0.0018546	-0.4018977	0.683	0.041	16.211	1.494
353Lys	1.109083	0.0990816	0.0458638	0.0062322	-0.2595854	0.727	0.065	10.365	3.134
354Val									
355Leu									
356Val					-0.4604829				
357Gly					-0.729881				
358Lys	0.8802423	0.0535182	0.1074493	0.0070845	-0.8537923	0.698	0.032		
359Asn	1.3839902	0.0673944	0.0454497	0.0050761	-0.1582148	0.609	0.028		
360Phe									
361Glu					-0.4392981				
362Asp	1.3081992	0.1119393	0.036021	0.0088786	-0.1962846	0.616	0.053	18.064	6.893
363Val	1.2347735	0.1548006	0.0785533	0.0138808	-0.2966348	0.692	0.071		
364Ala	1.0022453	0.0357448	0.1279008	0.0057974	-0.8355065	0.614	0.018		
365Phe	1.4238109	0.1440006			-0.322297				
366Asp					-5.3424005				
367Glu									
368Lys	1.2086081	0.1055867			-0.2795007				
369Lys	1.2594521	0.2008561	0.0584319	0.0031947	-0.4257869				
370Asn									
371Val									
372Phe	1.5274558	0.2984848	0.0840414	0.0191265	-0.244711	0.588	0.088		
373Val	0.7589645	0.0417981	0.1604366	0.0426651					
374Glu	0.8640315	0.0441949	0.0434953	0.00317	-0.1948411	0.933	0.048	8.308	1.836
375Phe			0.0734245	0.020673	-0.1513924				
376Tyr									
377Ala	1.5858706	0.1934743			-0.1136027				
378Pro									
379Trp	1.1763759	0.0785291	0.0322958	0.007847	-0.1910693	0.686	0.046	20	7.558
380Cys	1.0971953	0.1826264	0.0382239	0.0071329	-0.3533018	0.735	0.122	14.599	5.248
381Gly									
382His									
383Cys									
384Lys	1.0421347	0.110383	0.0636206	0.0112754					
385Gln	1.1878191	0.07818	0.0239387	0.0027094	-0.2119037	0.694	0.045	20	4.78
386Leu	1.1216094	0.0776292			-0.1011553				
387Ala	1.1006558	0.0369174	0.0412011	0.0080435	-0.6743822	0.734	0.025	10	4.754
388Pro									
389Ile	1.089089	0.1046289			-0.572229				
390Trp	1.0948016	0.177173	0.0685896	0.0101653					

391Asp	1.0898152	0.0469731	0.1387241	0.0201271	-0.1160787	0.687	0.029		
392Lys	1.2896342	0.0889561	0.0440251	0.0048366	-0.654172	0.682	0.042		
393Leu	1.160493	0.0800532	0.0427048	0.0046398	-0.2477334	0.695	0.048	12.485	2.654
394Gly									
395Glu					-0.4179524				
396Thr	0.7996652	0.0402174	0.066046	0.0042206	-0.4701433	0.99	0.039		
397Tyr	1.0509366	0.068629	0.0484902	0.0017754	-0.5560819	0.767	0.05	8.55	1.092
398Lys	0.8372012	0.1043281	0.0271631	0.0022129	-0.3002482	0.993	0.12	20	3.544
399Asp									
400His	0.8391009	0.058205	0.0823486	0.0028778	-0.9397504	0.799	0.025		
401Glu	1.1286534	0.1191573	0.0317254	0.0029258	-0.1828042	0.717	0.075	20	3.14
402Asn	0.9310689	0.1194161			-0.304284				
403Ile					-0.0440199				
404Val			0.0504469	0.0126135	-0.2494588				
405Ile			0.1142753	0.0209954	-0.2860353				
406Ala	1.0097308	0.110296	0.0375096	0.005782	-0.188379	0.799	0.087	14.097	4.333
407Lys	1.0881839	0.1208947							
408Met									
409Asp									
410Ser	0.9769521	0.0699195			-0.1636374				
411Thr					-0.2072613				
412Ala					-0.6099079				
413Asn	1.0556729	0.0362678	0.0372062	0.0020659	-0.2585701	0.764	0.026	14.86	1.548
414Glu									
415Val	1.1551434	0.0898636	0.0388238	0.0053736	-0.4173847	0.698	0.054	14.775	3.666
416Glu	1.1742396	0.0880371	0.0459651	0.0060743	-0.2613929	0.738	0.05		
417Ala	1.1420657	0.0774751	0.0327049	0.0021607	-0.3881908	0.706	0.048	19.468	2.156
418Val	1.3211388	0.1097839			-3.4610231				
419Lys									
420Val									
421His									
422Ser	1.3559811	0.1136434			-0.5710778				
423Phe	1.1451502	0.0498081	0.0624484	0.0081065					
424Pro									
425Thr	1.0039559	0.0986151			-0.1070298				
426Leu	1.0624079	0.0778756	0.0440732	0.007004	-0.4279466	0.759	0.056	10.75	3.71
427Lys	1.1065098	0.0723999			-0.2079964				
428Phe					-0.0670697				
429Phe	1.2394729	0.1469693	0.0485571	0.0091064	-0.3295475	0.71	0.074		
430Pro									
431Ala	1.3155424	0.0522057			-0.4429593				
432Ser	0.8347549	0.0587638	0.0526891	0.0073835	-0.3325891	0.966	0.068	3.783	2.867
433Ala									
434Asp	1.1769274	0.0907767			-0.2610787				

435Arg									
436Thr	1.1588669	0.1111516	0.0346933	0.0018448	-0.5034227	0.696	0.067	17.877	1.858
437Val	1.04783	0.0684382			-0.398349				
438Ile	0.9530225	0.1548119	0.1013483	0.0337006	-0.4011353	0.78	0.115		
439Asp	1.1503889	0.0825541	0.0419693	0.004773	-0.2805491	0.701	0.05	12.799	2.823
440Tyr									
441Asn									
442Gly									
443Glu									
444Arg									
445Thr	1.112639	0.0477695	0.0453182	0.0019392	-0.4767054	0.725	0.031	10.664	1.064
446Leu	1.1717494	0.0557598	0.0506786	0.0043654	-0.0271312	0.736	0.031		
447Asp	0.8819286	0.034978	0.1359215	0.0074258	-0.7139761	0.616	0.021		
448Gly	1.1667969	0.1836047							
449Phe	0.9219761	0.0765809			-0.2464833				
450Lys									
451Lys	1.1328831	0.0801807	0.1606832	0.0307013					
452Phe	0.7362345	0.0313989	0.0625575	0.0032631	-0.7868927	1	0	0.255	0
453Leu									
454Glu									
455Ser	1.276383	0.1206294			-1.6170946				
456Gly	0.978422	0.0660027	0.0482304	0.0039139	-0.4405891	0.824	0.056	7.768	1.896
457Gly	0.9987044	0.0552533	0.0433219	0.0042172	-0.4987121	0.808	0.045	10.38	2.354
458Gln	1.0263165	0.0559632	0.0466907	0.0002936	-0.735535	0.786	0.043	9.055	0.687
459Asp	0.8045711	0.0278771	0.3292223	0.0155212	-1.4900463	0.245	0.009		
460Gly	0.9416667	0.0427083	0.1527926	0.0068247	-1.1792803	0.498	0.017		
461Ala	0.8697442	0.0192825	0.447579	0.0166892	-1.5591128	0.191	0.005		
462Gly	1.0350428	0.0531961			-1.8399625				
463Asp	0.8754269	0.0294961	0.3727984	0.0195186	-1.6592665	0.228	0.009		
464Asp									
465Asp									
466Asp									
467Leu									
468Glu									
469Asp									
470Leu									
471Glu									
472Glu									
473Ala	0.94223	0.032721	0.4771922	0.0263685	-1.8692891	0.175	0.007		
474Glu	1.3006152	0.1012529			-0.3782554				
475Glu									
476Pro									
477Asp									
478Met									

479Glu									
480Glu									
481Asp									
482Asp									
483Asp									
484Gln									
485Lys									
486Ala	0.9236309	0.0459367			-1.8470411				
487Val	0.9452844	0.0350752	0.5518533	0.0279023	-2.1150401	0.139	0.006		
488Lys	1.0277005	0.0578372	0.5210996	0.0444134	-2.3875716	0.157	0.01		
489Asp	1.0682527	0.053977	0.5416986	0.0479071	-2.6645145	0.161	0.01		
490Glu	1.2689928	0.0757987	0.629564	0.0550278	-2.8945189	0.128	0.009		
491Leu	1.6293568	0.0103265	0.3519658	0.0051126	-3.2296583	0.31	0.002		