

https://dspace.org/

Topologically conserved hydrophobic residues of the thioredoxin C-subdomain stabilise GSTs

Parbhoo, Nishal

2014-06-12
http://hdl.handle.net10539/14751

Downloaded from a DSpace repository. https://dspace.org/

Topologically conserved hydrophobic residues of the thioredoxin C-subdomain stabilise GSTs

Nishal Parbhoo

A thesis submitted to the Faculty of Science, University of the Witwatersrand, Johannesburg, in fulfillment of the requirements for the degree of Doctor of Philosophy.

Johannesburg, 2013

Declaration

I declare that this thesis is my own, unaided work. It is being submitted for the degree of Doctor of Philosophy in the University of the Witwatersrand, Johannesburg. It has not been submitted for any other degree or examination at any other University.



Nishal Parbhoo

6th day of December, 2013

ABSTRACT

The thioredoxin-like fold is a well conserved fold that is present in many families of proteins. One such superfamily of proteins include the GSTs which are involved in phase II detoxification. GSTs primarily catalyse the metabolism of xenobiotics but are also involved in transporting non–substrate ligands and reactive compounds. The GST fold comprises an N-terminal thioredoxin domain and an all alpha helical C-terminal domain and is present in at least 18 classes of proteins. The N-terminal thioredoxin domain is characterised by the βαβαββα topology and can be further divided into two structural motifs, an N-terminal (βαβ) and a C-terminal (αββα) motif. A well conserved hydrophobic network exists between these two motifs and the role of the C-terminal motif is elucidated in this study using class Alpha GST as a model protein. A topologically conserved valine (Val58) and an isoleucine (Ile75) located on β3 and α3, respectively, were mutated to alanine. Secondary and tertiary structural characterisation as well as ligandin function of the mutant enzymes displayed no major structural alteration with respect to the wild-type enzyme. This was confirmed with high resolution crystal structures obtained. Enzymatic activity was maintained indicating that no structural alterations have occured that affects the active site dynamics and the domain interface as a result of the induced mutations. Thermal denaturation studies, however, indicated a slight destabilisation in the enzyme in the case of the valine mutation, but a large destabilisation was witnessed as a result of the isoleucine mutation. This is further observed in denaturant-induced equilibrium studies where the thermodynamic stability of proteins can be determined. Furthermore, as a result of the isoleucine mutation, the enzyme unfolds via a populated intermediate in contrast to the wild-type which globally unfolds via a two-state mechanism with no stable intermediates being populated. Pulse-proteolysis was employed as an additional probe for thermodynamic stability where the enzyme was digested by thermolysin at varying denaturant concentrations. Pulse-proteolysis results were in agreement with the thermal and denaturant-induced stability studies further confirming that the isoleucince substitution causes a large destabilisation. Thus these conserved hydrophobic

residues of the thioredoxin C-subdomain play a crucial stabilising role in the GST fold.

"Alles ist einfacher, als man denken kann,
zugleich verschränkter, als zu begreifen ist."

- Maximen Goethe

Everything is simpler than can be imagined,
yet more intricate than can be comprehended.

ACKNOWLEDGMENTS

My supervisor, Professor H.W. Dirr for his support, guidance, Patience, criticism and for allowing me to develop my postgraduate career in his research unit.

Dr. I. Achilonu and Dr. S. Fanucchi of the Protein Structure Function Research Unit and Dr M.A. Fernandes, School of Chemistry, University of the Witwatersrand for their assistance on the crystallography aspect of this work.

All of my colleagues, both past and present of the Protein-Structure Function Research Unit for their support, informal discussions and the great laughter they provided.

My former co-supervisor, Dr. Samantha Gildenhuys for her support and advice throughout my studies.

The University of Witwatersrand, National Research Foundation and the Andrew Mellon foundation for financial assistance.

Finally, my parents and friends who have put up with me and continued to support me during this entire process.

TABLE OF CONTENTS

Declaration ii
ABSTRACT iii
ACKNOWLEDGMENTS vi
TABLE OF CONTENTS vii
List of Figures x
List of tables xii
Abbreviations xiii
Research output xv
CHAPTER 1. INTRODUCTION 1
1.1 Protein folding and stability 1
1.2 Folding of multidomain proteins 3
1.3 Redox control 4
1.3.1 Thioredoxin 5
1.3.2 The thioredoxin fold 5
1.4 GST family of proteins 7
1.4.1 Structural characteristics of GSTs 7
1.4.2 Trx fold in the GSTs 9
1.4.3 Stability and folding of GSTs 11
1.4.3.1 Equilibrium unfolding of GSTs 11
1.4.3.2 Folding pathway of GSTs 14
1.5 Probes for (un)folding reactions 15
1.6 Aims and objectives 16
CHAPTER 2. EXPERIMENTAL PROCEDURES 18
2.1 Materials 18
2.2 Experimental 18
2.2.1 Generation of hGSTA1-1 mutants 18
2.2.2 Transformation, over-expression and purification of hGSTA1-1 19
2.2.3 SDS-PAGE 21
2.2.4 Absorbance spectroscopy 21
2.2.5 Circular dichroism spectroscopy 22
2.2.6 Fluorescence spectroscopy 24

2.2.6.1 Intrinsic fluorescence-tryptophan fluorescence 24
2.2.6.2 Extrinsic fluorescence - ANS binding..... 24
2.2.7 Glutathione (GSH) quantification..... 25
2.2.8 Standard GSH-CDNB conjugation assay 25
2.2.9 Thermal denaturation studies..... 26
2.2.10 Urea - induced equilibrium unfolding studies of the hGSTA1-1 proteins..... 27
2.2.10.1 Recovery of unfolding 27
2.2.10.2 Urea - induced equilibrium unfolding/refolding..... 27
2.2.11 Data fitting 28
2.2.11.1 Two-state fitting model..... 28
2.2.11.2 Three-state fitting model..... 31
2.2.12 Pulse proteolysis 33
2.2.13 Crystallography..... 33
2.2.13.1 Crystal growth..... 33
2.2.13.2 Diffraction..... 34
2.2.13.3 Refinement..... 34
CHAPTER 3. RESULTS 35
3.1 Sequence identity..... 35
3.2 Over-expression and purification..... 35
3.3 Size and purity determination 35
3.4 Absorbance spectroscopy..... 35
3.5 Structural Characterisation..... 39
3.5.1 Secondary structure characterisation..... 39
3.5.2 Tertiary structure characterisation..... 39
3.5.2.1 Intrinsic fluorescence 39
3.5.2.2 Near-UV circular dichroism..... 42
3.6 Functional characterisation 42
3.6.1 Ligandin (H-site) binding..... 42
3.6.2 Enzymatic activity..... 44
3.7. Conformational stability..... 48
3.7.1 Thermal-induced unfolding..... 48

3.7.2 Recovery of urea-induced unfolding..... 50
3.7.3 Urea-induced equilibrium unfolding..... 50
3.7.3.1 Equilibrium unfolding of wild-type and V58A hGSTA1-1..... 50
3.7.3.2 Equilibrium unfolding of I75A hGSTA1-1..... 55
3.7.4 Properties of the unfolding intermediate formed by the I75A mutation 57
3.7.5 Data fitting and thermodynamic parameters..... 59
3.8 Pulse Proteolysis 62
3.10 Protein crystallography 65
CHAPTER 4: DISCUSSION..... 72
4.1 Role of the C-subdomain in structural integrity..... 74
4.2 Role in catalysis and ligandin function 77
4.3 Impact of the Trx C-subdomain on stability of GSTs..... 78
4.4 Conclusion 88
CHAPTER 5. REFERENCES 89

List of Figures

Figure 1-1. Protein folding energy landscape. 2

Figure 1-2. The architecture of the Trx-fold 6

Figure 1-3. Ribbon representation of a dimeric GST 8

Figure 1-4: Structure based sequence alignment of the C-subdomain in GSTs. .. 10

Figure 1-5. Trx domain in hGSTA1-1 12

Figure 2-1. Urea denaturation curve 30

Figure 3-1. Plasmid sequencing results of the V58A and I75A encoding mutations. 36

Figure 3-2. SDS-PAGE analysis of hGSTA1-1 eluted proteins from cation-exchange chromatography column and size determination. 37

Figure 3-3. Absorbance spectra of the hGSTA1-1 proteins. 38

Figure 3-4. Far-UV circular dichroism spectra of hGSTA1-1 proteins. 40

Figure 3-5. Fluorescence spectra of hGSTA1-1. 41

Figure 3-6. Near UV CD spectra of the hGSTA1-1 proteins. 43

Figure 3-7. ANS binding to apo-hGSTA1-1 45

Figure 3-8. Specific activity of the hGSTA1-1 proteins 46

Figure 3-9. Thermal unfolding of hGSTA1-1. 49

Figure 3-10. Urea-induced equilibrium unfolding of wt hGSTA1-1 52

Figure 3-11. Urea-induced equilibrium unfolding of V58A hGSTA1-1 53

Figure 3-12. Urea-induced equilibrium unfolding of the I75A hGSTA1-1 56

Figure 3-13. Spectra of the native, intermediate and unfolded species of I75A hGSTA1-1 58

Figure 3-14. SDS-PAGE analysis of pulse proteolysis..... **Error! Bookmark not defined.**

Figure 3-15. Densitometric analysis of pulse proteolysis 64

Figure 3-16. Crystal images of the mutant hGSTA1-1 proteins 66

Figure 3-17. Fit of crystal models into electron density map 68

Figure 3-18. Ramachandran plots of the mutant hGSTA1-1 proteins 70

Figure 3-19. Structural alignment of wild-type and mutant hGSTA1-1 structures 71

Figure 4-1. Trx domain in GSTs 73

Figure 4-2. Domain interface of intersubunits of GSTs..... 76

Figure 4-3. Contribution of Val58 to the hydrophobic core in hGSTA1-1 80

Figure 4-4. Contribution of Ile75 to the hydrophobic core in hGSTA1-1 83

Figure 4-5. Thioredoxin subdomain interface interaction network 87

List of tables

Table 1-1: List of GSTs used for sequence alignment 10

Table 3-1: Specific activity values 47

Table 3-2. Extent of recovery of the native forms of the hGSTA1-1 proteins. 51

Table 3-3. Conformational stability parameters 60

Table 3-4. Conformational stability parameters for I75A hGSTA1-1 61

Table 3-5: Crystallographic X-ray data collection and refinement statistics of V58A and I75A hGSTA1-1. 67

Abbreviations

A_{280}	absorbance at 280 nm
A_{340}	absorbance at 340 nm
Å	Ångström
ANS	8-anilino-1-naphthalene sulphonate
ATP	adenosine triphosphate
CD	circular dichroism
CDNB	1-chloro-2,4-dinitrobenzene
C_m	the denaturant concentration at the midpoint of the unfolding curve
Da	Dalton
ΔG	change in Gibbs free energy
$\Delta G(H_2O)$	change in Gibbs free energy of unfolding in the absence of denaturant
dH_2O	distilled water
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
DTNB	5,5'-dithiobis(2-nitrobenzoic acid)
DTT	dithiothreitol
EDTA	ethylenediaminetetra-acetic acid
ε	molar extinction coefficient
E222	ellipticity at 222 nm
F325	fluorescence intensity at 325 nm
F340	fluorescence intensity at 340 nm
F490	fluorescence intensity at 490 nm
Far-UV CD	far-ultraviolet circular dichroism
G	Gibbs free energy
EcGrx2	*Escherichia coli* glutaredoxin-2
GSH	reduced glutathione
GST	glutathione transferase

hClic	human chloride intracellular channel
hGST A1-1	human class Alpha glutathione transferase
HT	the voltage applied to the circular dichroism photomultiplier tube
I	intermediate state
I75A	Ile75 substituted with alanine
IPTG	isopropyl-β-D-thiogalactopyranoside
K_{eq}	equilibrium constant
mdeg	millidegrees
m-value	the dependence of the free energy of unfolding on denaturant concentration
N	native state
OD_{600}	optical density at 600 nm
ORF	open reading frame
PDB	Protein Data Bank
SCOP	Structural Classification of Proteins
SDS-PAGE	sodium dodecyl sulphate-polyacrylamide gel electrophoresis
SE-HPLC	size exclusion-high performance liquid chromatography
*Sj*GST	*Schistosoma japonicum* glutathione transferase
T	temperature
TEMED	*N*, *N*, *N*', *N*'-tetramethylethylenediamine
Trx	thioredoxin
[θ]	mean residual ellipticity
$[\theta]_{222}$	mean residual ellipticity at 222 nm
U	unfolded state
UV	ultraviolet
V58A	Val58 substituted with alanine
YT	yeast tryptone

The IUPAC-IUBMB one and three letter codes for the amino acids were used throughout.

Research output

Conferences:

- *Poster*: 'A conserved structural motif of a hydrophobic network impacts on stability of glutathione transferases'. **N. Parbhoo**, S. Gildenhuys and H.W. Dirr. Molecular Biosciences Research Thrust Reasearch Day, University of Witwatersrand PDH, Dec 2011.

- *Poster*: 'A conserved structural motif of a hydrophobic network impacts on stability of glutathione transferases'. **N. Parbhoo**, S. Gildenhuys and H.W. Dirr. 23rd SASBMB Congress, Drakensburg, Jan 2012.

- *Oral*: 'Topologically conserved hydrophobic residues of the thioredoxin C-subdomain stabilises GSTs'.4th Cross-Faculty Postgraduate Symposium, University of Witwatersrand PDH, Oct 2012.

Protein Data Bank depositions:

- [3ZFB]: Crystal structure of the I75A mutant of human class alpha glutathione transferase in the apo form. Parbhoo, N., Fanucchi, S., Achilonu, I.A., Fernandes, M.A., Gildenhuys, S., Dirr, H.W. 19-Dec-2012.

- [3ZFL]: Crystal structure of the V58A mutant of human class alpha glutathione transferase in the apo form. Parbhoo, N., Fanucchi, S., Achilonu, I.A., Fernandes, M.A., Gildenhuys, S., Dirr, H.W. 19-Dec-2012.

Original Publication:

Proposed manuscript title: Contribition of topologically conserved residues of the Trx-like domain to the stability of GSTs.

Authors: Parbhoo, N., Khoza, T., Achilonu, I., Fanucchi, S., Fernandes, M.A. and Dirr, H.W.

Target Journal: PLoS ONE

CHAPTER 1. INTRODUCTION

1.1 Protein folding and stability

Proteins play a pivotal role in intricate reactions in most, if not all biological processes, as catalysts, structural, storage and carrier molecules. Depending on the function, the nascent amino-acid chain, or part thereof, is required to correctly fold into the biologically active, three-dimensional structure of the native state.

Various models have been proposed over the years to increase the understanding of the highly cooperative process of the folding of proteins to their native states. The current accepted model is the concept of energy landscapes and folding funnels (Wolynes *et al*., 1995; Dill and Chan, 1997; Dobson and Karplus, 1999). The width of the funnel represents the conformational entropy of the polypeptide chain, while the depth represents the free energy of stabilisation (Figure 1-1). Folding proceeds (proceeding down the funnel), through a progressive organisation of partially folded structures in which it may encounter intermediate states and/or kinetic traps, driven by the accumulation of favourable enthalpic interactions. An ensemble of the native structure exists at the bottom of the funnel, at a global energy minimum, having the most stable conformation. The unfolded protein can also form partially folded or misfolded states which may lead to the formation of aggregates if a high thermodynamic or kinetic barrier is experienced in the pathway (Hartl and Hayer-Hartl, 2009). In many instances, aggregate formation is the result of exposed hydrophobic residues coalescing from the unstructured backbone and thus forming non-native interactions between other protein molecules (Figure 1-1). The formation of aggregates can also result in highly ordered, fibrillar amyloids (Figure 1-1) which often lead to the toxic state of the cells giving rise to neurodegenerative diseases. No correlation has been found between sequence and amyloid formation and is thus believed to be an inherent property of the protein (Dobson, 2004).

The folding process involves a limited number of intermediate species. Point mutations in proteins sometimes block the folding pathway at the level of stable intermediate states thus disabling the protein to adopt its native conformation, thereby causing misfolded proteins or the formation of aggregates. The consequences are altered or lost biological activity and ultimately genetic diseases (Santucci e*t al*., 2008)

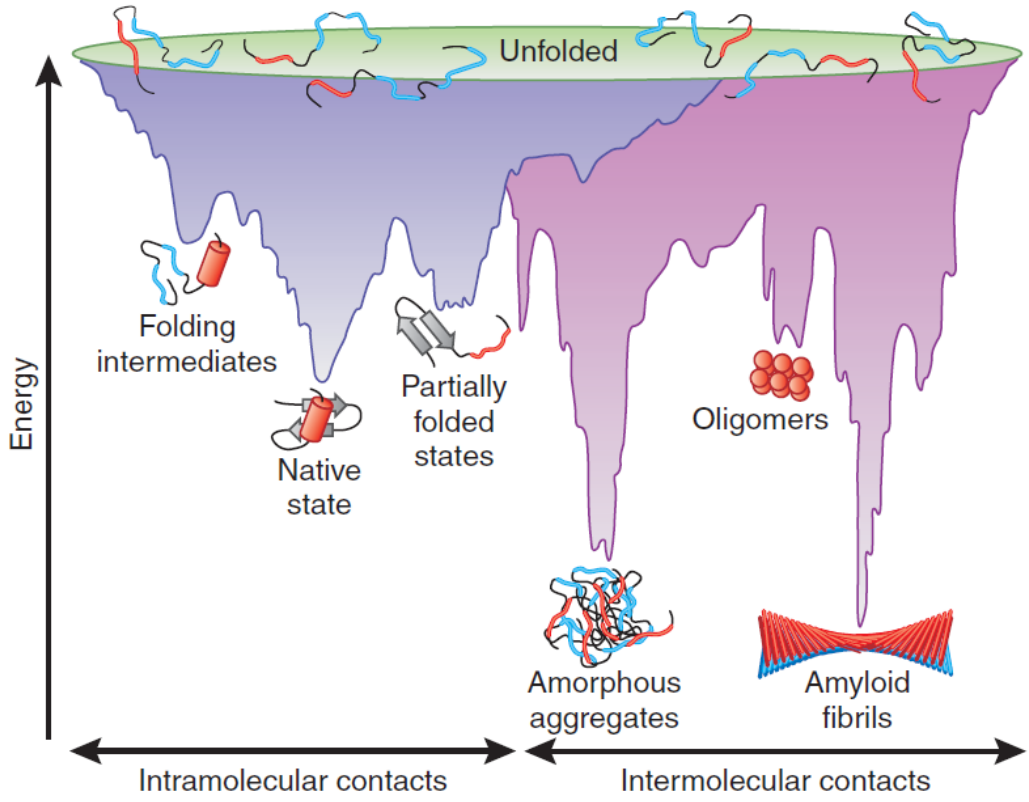


Figure 1-1. Protein folding energy landscape.
A representation schematic of a typical folding energy funnel illustrating a rugged landscape that contains kinetic and thermodynamic traps. (Image from Hartl and Hayer-Hartl, 2009).

Thermodynamic and kinetic studies have shown the presence of stable intermediate states in a number of proteins (Privalov, 1996; Horwick 2002; Calamai *et al*., 2005). One such intermediate, known as the 'molten globule state' (Ohgushi and Wada, 1983; Bhattacharyya and Varadarajan, 2013), is characterised by compact secondary structure, but fluctuating tertiary conformation allowing these regions to be probed by binding of a hydrophobic dye, e.g. 8-anilino-1-naphthalene sulphonate (ANS) (Semisotnov *et al*., 1991; Chaffotte *et al*., 1992). In multidomain proteins, molten-globule states occur when one domain or subdomain of the partially unfolded protein remains folded, while the other is unfolded (Freire *et al*., 1992; Freire, 1995; Privalov, 1996; Vreuls *et al*., 2004; Bhattacharyya and Varadarajan, 2013). Molten-globule states are also involved in biological processes such as membrane insertion and transmembrane trafficking that require the protein to become partially unfolded (Hartl *et al*., 1994). Intermediates are however generally unstable and poorly populated at equilibrium (Yon, 2001).

It is crucial to understand the forces that give rise to the folded state of a protein and that govern its conformational stability. The relative stability of a protein is determined by the energy difference between the folded and unfolded state (ΔG). The forces contributing to the stability include van der Waals forces, hydrogen bonds, hydrophobic interactions and electrostatic interactions. While van der Waals forces may be relatively weak, they are ubiquitous in proteins and can accumulate to a significant amount (Pace, 2001). The resulting clusters of the hydrophobic residues, aided by van der Waals interactions, minimises the surface area exposed to the solvent (Pace *et al*., 1996). The hydrophobic free energy contribution to protein stability is estimated to be about 60 kcal.mol^{-1} (Dill *et al*., 1989). In the interior, domain and/or subunit interface of proteins, steric complementarity is crucial to maximise the stabilising effects of these forces (Ratnaparkhi and Varadarajan, 2000).

1.2 Folding of multidomain proteins

The majority of studies on protein folding have focused on monomeric proteins, which have been extensively characterised using numerous experimental and theoretical methods (Gianni *et al*., 2003; Lindorff-Larsen *et al*., 2005; Wolynes, 2005; Dill *et al*., 2007; Han *et al*., 2007). Experimental methods include spectroscopic techniques such as tryptophan fluorescence and circular dichroism (CD) which provide low resolution information but

together with protein engineering, the folding mechanisms can be mapped in more detail. Techniques such as nuclear magnetic resonance (NMR) and hydrogen-deuterium exchange mass spectrometry (H/DXMS) are able to provide atomic detail of the folding mechanisms. Theoretical techniques mainly involve the use of computational tools such as molecular dynamic simulations. Computational tools allow the prediction of folding processes at very small time scales (picoseconds to milliseconds) and are often useful for proteins that are difficult to isolate in their native form such as membrane proteins. Due to their relatively small size and simplicity, monomeric proteins are particularly good candidates for in-depth analyses. However, monomeric proteins represent a very small fraction of the proteome, 19% of proteins in *Escherichia coli,* and the percentage is even lower in higher organisms (Goodsell and Olson, 2000). In contrast, the proteome mainly consists of oligomeric proteins, the most common assemblies being homodimers, which constitute about 38% of proteins in *Escherichia coli* (Goodsell and Olson, 2000). Nature has opted to construct quaternary complexes as it is simpler to make smaller individual subunits rather than a single chain large protein (Marianayagam *et al*., 2004). A shortcoming in using monomeric proteins as models for folding is that it overlooks the contribution to stability from quaternary interactions in multisubunit / protein-protein complexes. A dimeric protein can be considered as a stable complex composed of two identical or nonidentical subunits. In the past decade, a few in-depth folding studies have been conducted on oligomeric proteins, in particular on homodimers, shedding more light on the complexes formed *in vivo* (reviewed in Rumfeldt *et al*., 2008). Given the biological importance of these quaternary structural complexes (Rumfeldt *et al*., 2008), it is essential to extend our understanding of the folding of higher order complexes.

1.3 Redox control

The energy production (synthesis of ATP) in cells occurs via tightly controlled metabolic cycles in which many reactions involve the transfer of electrons with oxygen being the ultimate acceptor. As a result, many oxygen derived radicals or reactive oxygen species (ROS), such as superoxide (O_2^-), hydrogen peroxide (H_2O_2) hydroxide radical (OH^-), can accumulate causing oxidative stress and may have deleterious consequences to the cell (Alfadda and Sallam, 2012). Therefore it is crucial that the redox homeostasis of a cell is maintained.

Cells possess many antioxidant enzymatic systems to control the excessive ROS formation. The tripeptide glutathione (GSH) is the most abundant free thiol antioxidant intracellularly (Meister and Tate, 1976) and serves as a redox buffer through the process of glutathionylation where a mixed disulfide is formed between the cysteine moieties of the protein and GSH (Dalle-Donne *et al*., 2007; Gallogly and Mieyal, 2007). Furthermore it can detoxify foreign compounds upon conjugation with GSH dependant enzymes such as glutaredoxins (Fernandes and Holmgren, 2004), glutathione peroxidase (Epp *et al*., 1983), glutathione reductases (Meister and Anderson, 1983) and glutathione transferases (Armstrong, 1997). Many of the enzymes involved in redox control possess a common structural feature, a thioredoxin-like fold (Pan and Bardwell, 2006).

1.3.1 Thioredoxin

Thioredoxin (Trx) was first discovered as an electron donor for ribonucleotide reductase in *Escherichia coli* (Laurent *et al*, 1964), an enzyme required for DNA synthesis and repair (Nordlund and Reichard, 2006). In addition to redox control, Trx has a wide range of functions ranging from facilitating protein folding (Berndt *et al*., 2008) to regulating apoptosis (Ravi *et al*., 2005), modulating inflammatory responses (Nakamura *et al*., 2005) and playing a role as a growth factor (Powis *et al*., 2000). Its vast role in many metabolic processes has also made it a target for chemotherapeutic agents (Gromer *et al*., 2004; Arnér and Holmgren, 2006). It is a 12 kDa protein ubiquitous in prokaryotes and eukaryotes and contains the CxxC active site motif which is well conserved among proteins consisting of the thioredoxin-like fold (Martin, 1995). Although the Trx proteins bear a low sequence similarity, the Trx fold is highly homologous in the Trx superfamily (Pan and Bardwell, 2006).

1.3.2 The thioredoxin fold

The thioredoxin-like superfamily encompasses at least 23 families (SCOP database; Murzin *et al*., 1995) with some containing subfamilies. The thioredoxin fold, central to the superfamily, consists of a βαβ motif and a ββα motif connected by an α-helix (Figure 1-2B). The two β-strands in the N-terminal motif are parallel to one another, and those of the C-terminal motif are antiparallel. The resulting mixed β-sheet forms the core of the thioredoxin fold and is sandwiched between three α-helices (Figure 1-2). In addition to the conserved active site motif, a *cis*-Pro residue located adjacent to the β-strands on the N-

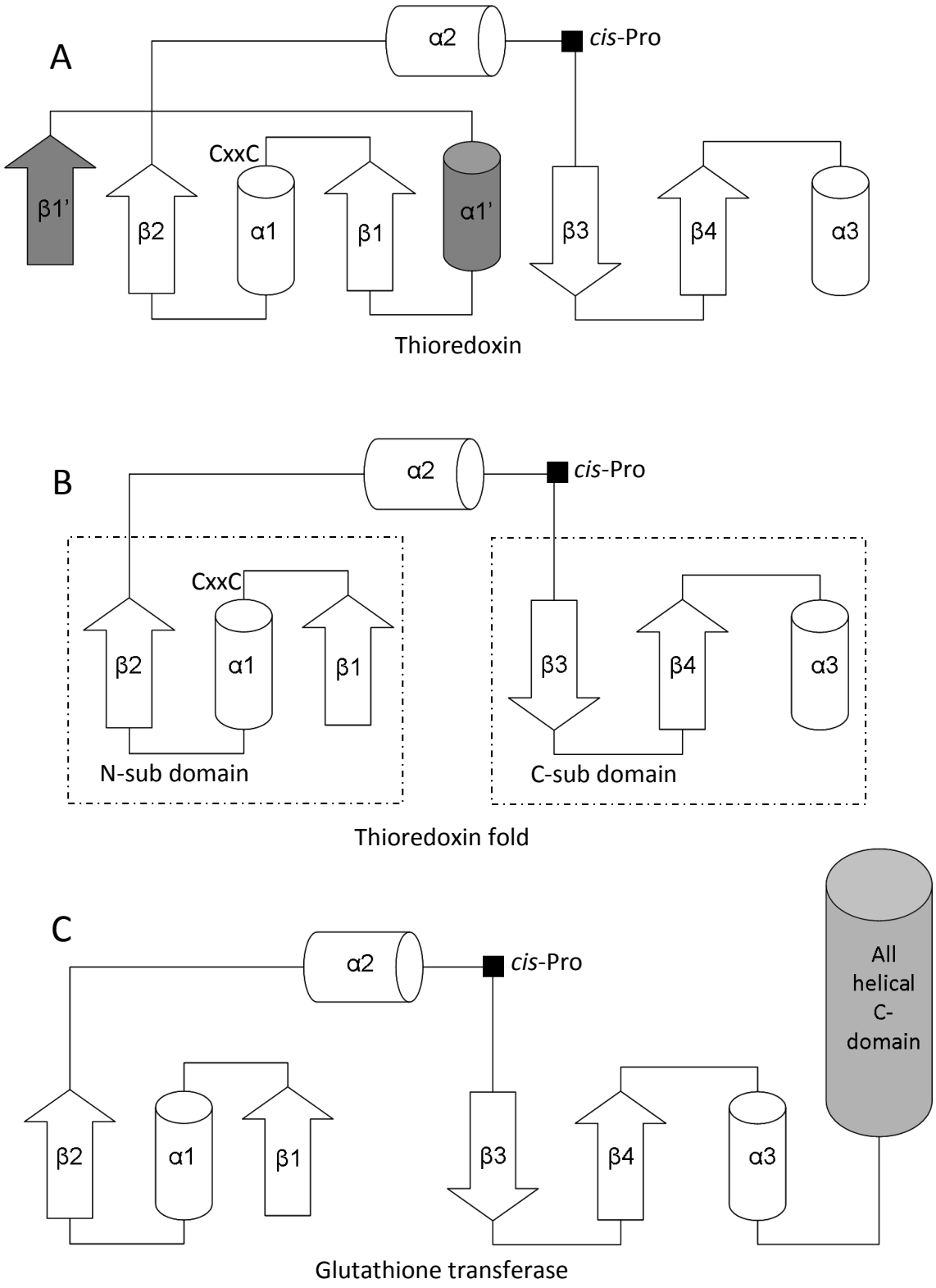


Figure 1-2. The architecture of the Trx-fold
The conserved Trx-fold topology in (A) Trx and (C) GSTs. The conserved active site motif (CxxC) and *cis*-Pro is located in the loop preceding α1 and between α2 and β3, respectively. Image adapted from Pedone *et al*. (2010).

terminal motif is also well conserved (Martin, 1995). The largest subfamily containing the Trx fold is the glutathione transferases (GSTs).

1.4 GST family of proteins

Cytosolic glutathione transferases (GST; EC 2.5.1.18) comprise a family of multifunctional homo- or heterodimeric proteins, which catalyse the conjugation of glutathione to a large variety of endogenous and exogenous electrophilic compounds (Mannervik and Danielson, 1988; Hayes and Pulford, 1995; Armstrong, 1997). Besides their detoxification roles, GSTs are also involved in the modulation of signal transduction, ion channels and the storage of nitric oxide (Adler *et al*., 1999; Cho *et al*., 2001; Dulhunty *et al*., 2001; Wang *et al*., 2001).

1.4.1 Structural characteristics of GSTs

GSTs consist of subunits of ~ 200 amino acids each. Visual inspection of the structures of the GSTs reveal an overall conserved architecture, as illustrated in Figure 1-3 using class Alpha as a model. Each subunit is composed of two structurally distinct domains: an N-terminal thioredoxin-like domain (domain 1), characterised by the Trx-fold topology as previously described (section 1.3.2) and an all α-helical C-terminal domain (domain 2). A structural feature that distinguishes class Alpha GSTs from those of other classes is the presence of an extended C-terminal region that forms an amphipathic helix (α9) at the active site which provides a "lid" over the active site in the presence of ligands (Cameron *et al*., 1995) and is otherwise conformationally flexible in the ligand-free form (Dirr and Wallace, 1999; Kuhnert *et al*., 2005).

The active site of the GSTs can be divided into two adjacent subsites: the G-site for binding GSH and the H-site for binding hydrophobic electrophilic substrates (Dirr *et al*., 1994b). The G-site is formed primarily by the thioredoxin-like N-terminal domain, whereas the C-terminal domain forms the H-site. The peptidyl moiety of the product lies at the end of the β-sheet of domain 1 and interacts with the protein through a number of electrostatic and hydrogen-bonding interactions. The γGlu moiety of GSH interacts with a hydrophilic complementary pocket near the subunit interface formed by residues in domain 1 of one subunit and domain 2 of the adjacent subunit. The dimeric structure of the glutathione transferases is required for the formation of the active site on each subunit (Figure 1-3)

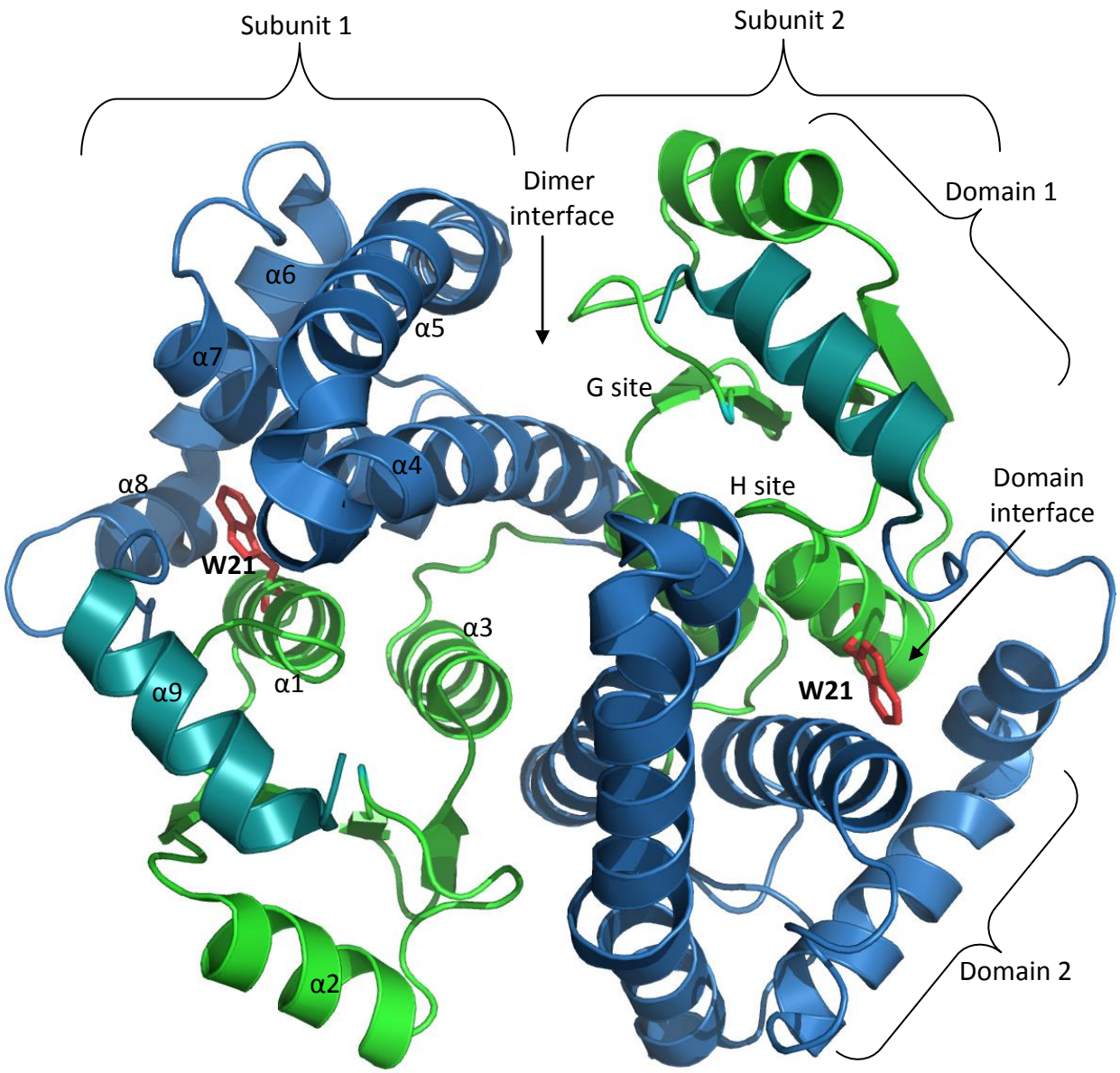


Figure 1-3. Ribbon representation of a dimeric GST
The N-terminal domain (domain 1) shown in green, and the all α helical C-terminal domain shown in blue. The helix (α9) that 'caps' the active site of domain 1 shown in teal. The tryptophan residues are involved in the interdomain contacts (highlighted in red as stick representations). Human class Alpha glutathione transferase (hGSTA1-1) is used here as a model illustration of GSTs. [Image rendered using PyMOL™, DeLano Scientific, 2006; PDB code: 1K3L (Le Trong *et al*., 2002)].

(Dirr *et al*., 2005; Kinsley *et al*., 2008) and thus the dimeric form of GSTs is essential for enzyme activity.

The GSTs share little sequence similarity but their tertiary structure is conserved (Wilce and Parker, 1994). The exception is class Kappa, a mitochondrial GST, whose C-terminal domain is inserted into the N-terminal domain (Ladner *et al*., 2004). The dimeric structure also has been shown to be involved in stabilisation of tertiary structures of individual subunits (Erhardt and Dirr, 1995) as well as to provide a non-substrate ligand-binding site at the subunit interface (Sluis-Cremer *et al*., 1996; Sayed *et al*., 2000; Lyon and Atkins, 2002; Yassin *et al*., 2004). Two types of interactions are found at the dimer interface, a conserved hydrophobic interaction found in mammalian classes of Alpha (Sayed *et al*., 2000), Mu (Hornby *et al.*, 2002; Thompson *et al*., 2006; Alves *et al*., 2006) and Pi (Stenberg *et al.*, 2000) and a hydrophilic interaction, where electrostatic forces predominate as in the classes of Sigma (Stevens *et al.*, 1998; Stevens *et al.*, 2000) and Theta (Rossjohn *et al.*, 1998).

1.4.2 Trx fold in the GSTs

All members of the GST family contain the thioredoxin fold in domain 1 (Sheehan, 2001), which is comprised of approximately 80 amino acid residues. The fold can be subdivided into an N and C-subdomain connected by a loop of residues that incorporates a third helix [Figure 1-2B, (Martin, 1995)]. The packing of secondary structures in the thioredoxin fold form a compact hydrophobic core (Figure 1-5). The α1 and α3 helices as well as the β1, β2 and β4 strands make up the hydrophobic core of the fold due to the packing of the α1 and α 3 against the one side of the β-sheet (Martin, 1995).

The superimposed structures of domain 1 of the GSTs reveal a high degree of structural similarity, however the different gene classes are characterised by specific structural features (Dirr *et al*, 1994a). As mentioned earlier, class Alpha has the terminal helix, α9, over domain 1 that is crucial for catalysis. An extended mobile loop in class Mu is present connecting the sequence between β2 strand and α2 helix (Ji *et al*., 1992). The class Kappa has domain 2 inserted into domain 1 containing the thioredoxin fold (Ladner *et al*., 2004).

A structure-based sequence alignment of domain 1 (Trx domain) of the different GST classes shows that topologically conserved residues are predominantly hydrophobic in

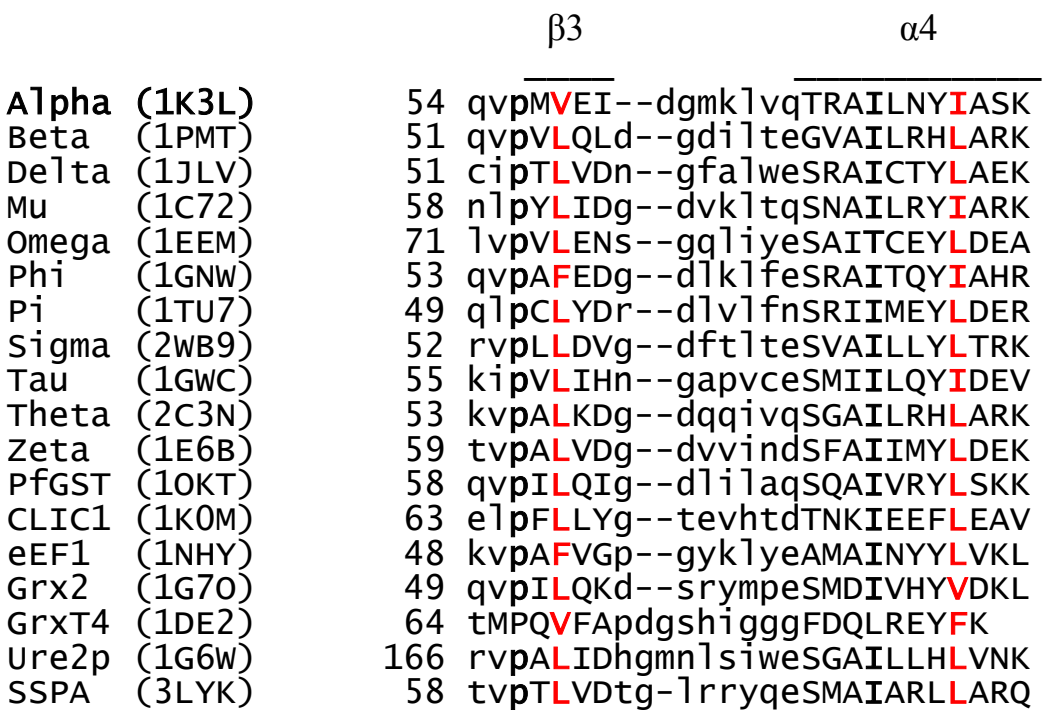


Figure 1-4: Structure based sequence alignment of the C-subdomain in GSTs.
The alignment was performed using sequences of one member of domain 1 of the 18 classes of GSTs. The PDB file for each class was obtained and the sequence alignment was performed using UCSF Chimera, v1.5.3 (Pettersen *et al*., 2004). The displayed alignment only shows residues of C-subdomains containing β3 and α4 for clarity purposes. The topologically conserved residues are boxed and those involved in this study are highlighted in red. The PDB accession codes and their reference are shown in Table 1-1. Each class of the GST fold was identified in the SCOP database (http://scop.mrc-lmb.cam.ac.uk/scop/data/scop.b.d.fh.b.html; Murzin *et al.*, 1995).

Table 1-1: List of GSTs used for sequence alignment

PDB Code	GST Class	Reference
1k3l	Alpha	Le Trong *et al.*, 2002
1c72	Mu	Chern *et al.*, 2000
1e6b	Zeta	Thom *et al.*, 2001
1eem	Omega	Board *et al.*, 2000
1g6w	yeast prion protein ure2p	Bousset *et al.*, 2001
1g7o	Grx2	Xia *et al.*, 2001
1gnw	Phi	Reinemer *et al.*, 1996
1gwc	Tau	Thom *et al.*, 2002
1jlv	Delta	Oakley *et al.*, 2001
1k0m	Clic1	Harrop *et al.*, 2001
1nhy	GST-like domain of EF-1	Jeppesen *et al.*, 2003
1okt	*Plasmodium falciparum* GST	Fritz-Wolf *et al.*, 2003
1pmt	Beta	Rossjohn *et al.*, 1998
2c3n	Theta	Tars *et al.*, 2006
2wb9	Sigma	Line *et al.*, 2010
3lyk	Stringent starvation protein A	Ramagopal *et al.*, 2010
1tu7	Pi	Perbandt *et al.*, 2005

in nature with a minimal difference in bulk of the side chains (Figure 1-4). The present study focuses on the C-subdomain of the Trx domain in GSTs and the conserved residues of interest are Val58, and Ile75 as they form a hydrophobic core between the subdomains of the Trx-fold in GSTs (Figure 1-5). Other well conserved residues of the C-subdomain include Pro56 and Ile71. Mutation of Pro56 highlighted its role in maintaining conformational stability and enzymatic activity (Nathaniel *et al*., 2003) and that of Ile71, which also contributes to the hydrophobic core, has been implicated in maintaining the structural stability of GSTs (Achilonu *et al*., 2010).

1.4.3 Stability and folding of GSTs

1.4.3.1 Equilibrium unfolding of GSTs

A three-state mechanism is observed for the equilibrium unfolding of most of the dimeric GSTs. In the case of the class Mu (Hornby, *et al*. 2000), Sigma (Stevens *et al*., 1998), Beta (Sacchetta *et al*., 1993) GSTs stable monomeric intermediates are present while a dimeric intermediate is present for yeast prion protein 2 (Ure2p) (Lian *et al*., 2006). In the case of class Alpha (Wallace *et al*., 1998) and class Pi (Dirr and Reinemer, 1991; Erhardt and Dirr, 1995, Gildenhuys *et al*., 2010b) however, no stable dimeric or monomeric intermediate is observed. A three-state model of unfolding is proposed as both these classes of GSTs posses a dynamic helix which is mobile at room temperature and unfolded and low concentrations of denaturant (Wallace *et al*., 1998; Gildenhuys *et al*., 2010b) and thus the following model is proposed:

$$N_2 \rightleftharpoons N_2^* \rightleftharpoons 2U$$

where N_2^* represents the dimeric state where the dynamic helix (α9) is unfolded.

The SjGST (Kaplan *et al*., 1997) and a monomeric counterpart of the GST family, EcGrx2 (Gildenhuys *et al*., 2008) follow a two-state equilibrium unfolding mechanism. Another monomeric homologue of GST, hClic1 (Fanucchi *et al*., 2008), also follows a two-state equilibrium mechanism at pH 7, however, it is changed to a three-state mechanism when encountering an acidic environment (pH 5.5) where it is required to undergo a structural change for membrane binding (Littler *et al*., 2004)

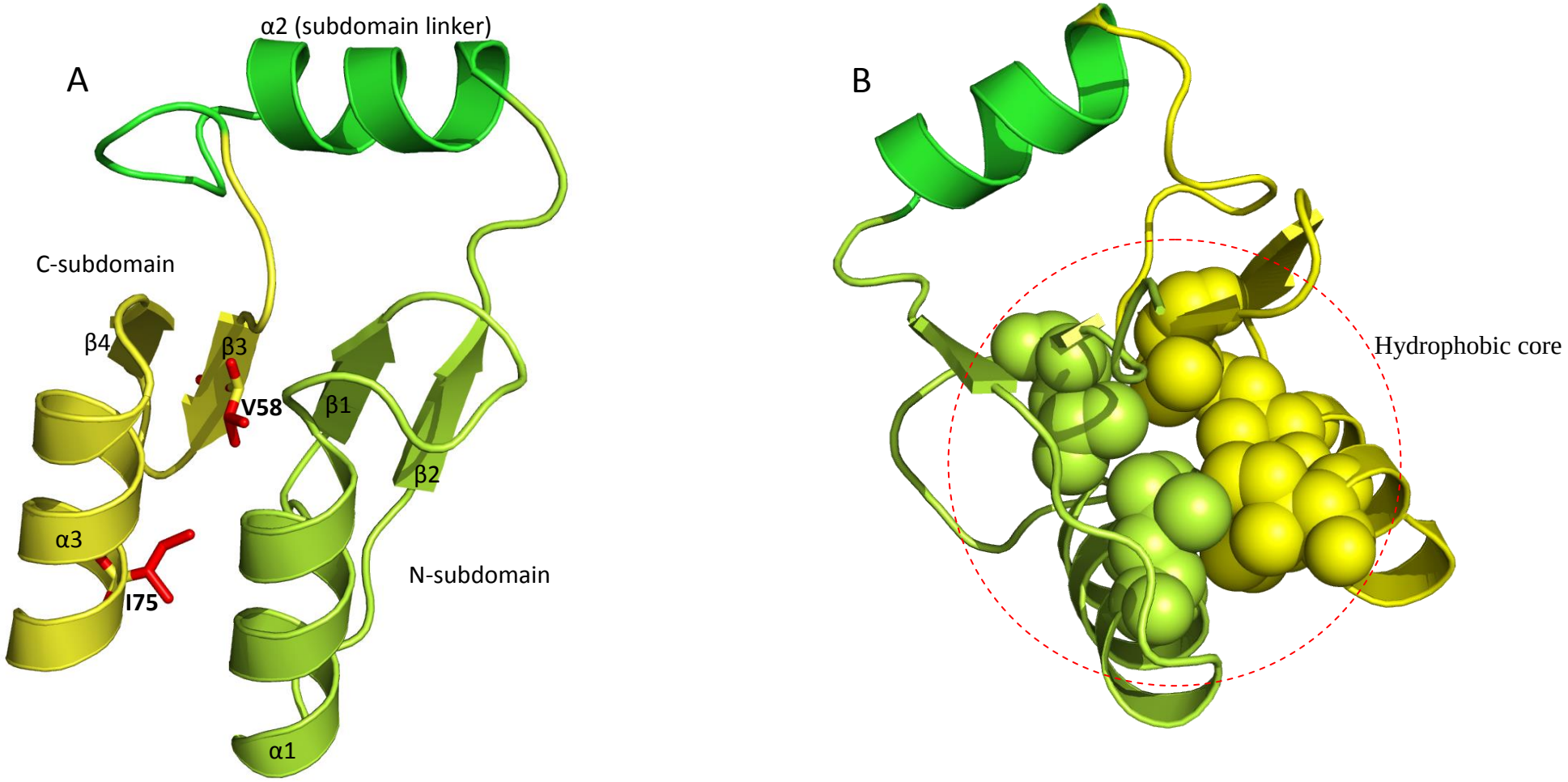


Figure 1-5. Trx domain in hGSTA1-1
The Trx domain (domain 1) of GSTs is made up of two subdomains; the N-subdomain ($\beta_1\alpha_1\beta_2$; lime green) and the C-subdomain ($\beta_3\beta_4\alpha_3$; yellow) connected by a linker, α2 (green). (A) Topologically conserved residues of the C-subdomain are shown in red. (B) Hydrophobic core (encircled in red) constituted by residues shown as spheres. (Images rendered using PyMOL™, DeLano Scientific, 2006; PDB code: 1K3O (Le Trong *et al*., 2002)).

The N- and C-terminal domains of GST proteins have been shown to unfold cooperatively in the case of class Alpha (Wallace *et al.*, 1998), Sigma (Stevens *et al.*, 1998), Pi (Erhardt and Dirr, 1995) and *Schistoma japonicum* GST (Kaplan *et al.*, 1997). In addition, there is no evidence to suggest that the individual domains of GST proteins unfold independently as indicated by the equilibrium unfolding profiles which display monophasic, coincident transitions. Therefore, the burial of a significant amount of hydrophobic surface area upon domain association indicates that the domain interfaces of GSTs play an important role in the stability and folding of these proteins. The N-terminal domain of Ure2p has no contribution to the stability and dimerisation of the protein (Lian *et al*., 2006).

The details of the domain-domain interfaces of GST proteins are not as well studied as their subunit-subunit interfaces. Luo *et al.* (2002) investigated the conformational stability and equilibrium unfolding of two domain-exchanged chimeric isoenzymes. The domains of the class Mu isoenzymes M1-1 and M2-2 were exchanged resulting in chimeras (M12 and M21) with one domain from M1 and one domain from M2 (Luo *et al.*, 2002). It was shown that the M12 and M21 monomers are less stable then the wild-type monomers (Luo *et al.*, 2002). A hydrophobic 'lock-and-key' motif is located at the domain interface of GSTs (Parbhoo *et al*., 2011). The effect of domain packing on stability and function was investigated in hGSTA1-1 (Wallace *et al.*, 2000; Balchin *et al*., 2010), hClic1 and EcGrx2 (Parbhoo *et al*., 2011). In all cases, a cavity forming mutation was created to assess the role of the 'lock-and-key' motif and it was found that the mutation was both disruptive and destabilising with the equilibrium unfolding results pointing to the accumulation of one or more intermediate species (Wallace *et al*., 2000; Parbhoo *et al*., 2011), hence the domain interface complementarity is critical for correct domain-domain packing which in turn plays a role in protein stability.

In addition to the domain interface, a 'lock-and-key' motif is located at the subunit interface of GSTs and has been found to be involved in the conformational stability of GSTs (Sayed *et al*., 2000; Hornby *et al*., 2002; Stenberg *et al*., 2000; Abdalla *et al*., 2002; Hegazy *et al*., 2004; Vargo *et al*., 2004; Alves *et al*., 2006). The GSTs display two types of interactions at the subunit interface (Armstrong, 1997). The first is a 'lock-and-key' hydrophobic interaction, involving an aromatic 'key' residue from

domain 1 of one subunit that inserts into several hydrophobic 'lock' residues of domain 2 in the other subunit (Figure 1-3) found in mammalian classes of Alpha [Phe52, (Sayed *et al.*, 2000)], Mu [Phe56, (Hornby *et al.*, 2002)] and Pi [Tyr50, (Stenberg *et al.*, 2000)]. The inter-subunit 'lock-and-key' is also found in the Zeta class but the 'key' is not an aromatic side chain but the side chain of methionine (Wilce and Parker, 1994). The second type of intersubunit interaction is more hydrophilic and lacks the hydrophobic 'lock-and-key' motif and electrostatic forces predominate as in the classes of Sigma (Stevens *et al.*, 1998; Stevens *et al.*, 2000) and Theta (Rossjohn *et al.*, 1998). An interesting variant of the 'lock-and-key' motif exists in the Delta class of GSTs, known as the 'clasp lock-and-key' (Wongsantichon and Ketterman, 2006). The aromatic 'key' residue not only inserts into a hydrophobic 'lock' of the neighbouring subunit, but also acts as part of the 'lock' for the other subunit 'key'. The 'key' residues from both subunits show aromatic ring stacking with each other in a pi– pi interaction, generating a 'clasp' in the middle of the subunit interface (Wongsantichon and Ketterman, 2006). Thus dimeric GST structure is required to maintain catalytically functional conformations of the individual subunits as well as the non-substrate ligand-binding site at the dimer interface (Dirr, 2001; Abdalla *et al.*, 2002; Vargo *et al.*, 2004). The extent of subunit stabilisation by quaternary interactions is less for classes Sigma and Mu but essential for the stability of the subunits for the classes Alpha, Pi and Mu GSTs (Dirr, 2001).

1.4.3.2 Folding pathway of GSTs

Protein families show low sequence similarity, however the fold remain highly conserved and the folding pathways are conserved in the different classes of the family (Gunesakaran *et al.*, 2001). Cytosolic GSTs share a canonical fold (Sheehan *et al.*, 2001), therefore are most likely to share a conserved folding mechanism.

Kinetics of (un)folding of GSTs have been characterised in class Alpha GST (Wallace *et al.*, 1998; Wallace and Dirr, 1999) and in Ure2p (Lian *et al.*, 2006). The kinetics of unfolding are complex for class Alpha GST as two phases of unfolding are witnessed, a fast (millisecond time range) and slow (second to minute time range). The *cis*-Pro (Pro56) has an influence on the unfolding of class Alpha GST, where the rates of unfolding on both phases are dramatically increased when the Pro56 was replaced with Gly (Nathaniel, *et al.*, 2003). Formation of a structured intermediate is ruled out

as there is an absence of a 'rollover' (non-linearity in an arm of a chevron plot) in the urea dependence of the rates of unfolding (Wallace *et al*., 1998). In contrast, an equilibrium unfolding intermediate of Ure2p is witnessed in its unfolding kinetics (Lian *et al*., 2006).

1.5 Probes for (un)folding reactions

Understanding the mechanism of protein folding and unfolding remains one of the most challenging questions in life sciences (Dill *et al*., 2007). It requires not only detailed knowledge of different species involved in folding/unfolding reactions but also detailed kinetic information that connects these species (Lindberg and Oliveberg, 2007). During the folding/unfolding reactions, a large number of noncovalent interactions form or break on various time scales, leading to distinct processes/phases in protein (un)folding reactions, such as hydrophobic collapse and secondary structure formation.

The chromophores in proteins, i.e. the peptide backbone and the aromatic amino acids (Phe, Tyr and Trp) have allowed us to exploit optical spectroscopic techniques to probe (un)folding processes in protein. While circular dichroism and fluorescence spectroscopy provide low resolution images of protein structures, they are able to give information on the specific events such as subunit and domain (dis)association in solution which cannot be obtained from high resolution imaging such X-ray crystallography. Much of the information on folding mechanism of GSTs has been characterised using spectroscopic techniques such as circular dichroism and fluorescence spectroscopy which report on changes occurring on the polypeptide backbone and tertiary environment of the aromatic amino acids, respectively. The information provided by any structural changes is highly dependent on the location of the probe. Tryptophan is the dominant intrinsic fluorophore in proteins and highly sensitive to the local environment surrounding its indole ring (Lacowicz, 2006). Tryptophan fluorescence together with circular dichroism was used to probe unfolding mechanism of the GSTs such as class Alpha (Wallace *et al*., 1998), Mu (Hornby *et al*., 2000) Pi (Dirr and Reinemer, 1991; Erhardt and Dirr, 1995; Gildenhuys *et al*., 2010b), Sigma (Stevens *et al*., 1998) and GST-like proteins such as EcGrx2 (Gildenhuys *et al*., 2008) and hClic1 (Fanucchi *et al*., 2008). These

spectroscopic techniques were also used to probe events occurring at the domain interface in GSTs (Wallace *et al*., 2000; Balchin *et al*., 2010; Parbhoo *et al*., 2011) as well as at the subunit interface (Stevens *et al*., 1998; Alves *et al*., 2006, Thompson *et al*., 2006).

A more accurate depiction of events upon (un)folding can be provided by the use of extrinsic probes, e.g. ANS, an amphipathic dye that can bind exposed hydrophobic patches of proteins (Slavik, 1982; Semisotnov *et al*., 1991; Engelhard and Evans, 1995; reviewed in Hawe *et al*., 2008). The use of ANS has allowed probing for intermediates in GST-like proteins (Fanucchi *et al*., 2008; Parbhoo *et al*., 2011). The use of spectroscopic independent probes e.g. proteolysis (Park and Marqusee, 2005) can be also used to either provide additional or to verify the information provided by the spectroscopic probes.

1.6 Aims and objectives

This present study examines the role of conserved residues in the C-subdomain of the Trx fold in GSTs. A well studied GST is the Alpha class, which possesses all the features found in the GST family of proteins and thus represents a suitable model. Topologically conserved residues in the C-subdomain have been identified and are represented by residues Val58 and Ile75 in hGSTA1-1 and are found in the hydrophobic core of the Trx domain. These residues are thus postulated to have a role in the stability of the protein as core residues are believed to drive the hydrophobic collapse and form stabilising van der Waals interactions through tight packing in the native state (Kellis *et al*., 1988; Dill, 1990). The role of these residues was determined by mutating them to Ala, which removes the bulk and any interactions they have in the hydrophobic core while not introducing a large degree of flexibility as to completely destabilise the protein with the risk of misfolded protein being formed.

The impact of the mutations on the native structure in solution was assessed by spectroscopic techniques using far-UV CD and intrinsic fluorescence. The integrity of the native structure was further confirmed by X-ray crystallography. The impact on the G and H-sites of the enzymes was assessed by determining the specific activity and the binding of a dye, ANS, to the enzymes. The conformational stability of the

proteins was determined by denaturant-induced equilibrium unfolding experiments using intrinsic and extrinsic spectroscopic probes. In all the experiments, the data were compared to that of the wild-type protein.

CHAPTER 2. EXPERIMENTAL PROCEDURES

2.1 Materials

The cDNA encoding wild-type hGSTA1-1 cloned into the pKHA1 plasmid was a generous gift from Prof. B. Mannervik (Uppsala, Sweden; Stenberg *et al*., 1992). The QuikChange® Lightning kit was purchased from Stratagene (La Jolla, CA, USA). Ampicillin, DTT and IPTG were purchased from Melford (Ipswich, United Kingdom). 8-Anilino-1-naphthalenesulfonic acid (ANS) and reduced GSH was purchased from Sigma-Aldrich (St. Louis, MO USA). SDS-PAGE molecular weight marker (SM0431) was purchased from Fermentas Life Sciences (St. Leon-Rot, Germany). CM-Sepharose was purchased from GE Healthcare Life Sciences (Uppsala, Sweden). Ultrapure (99.5%) urea was purchased from Merck chemicals (Darmstadt, Germany). All other reagents were of analytical grade. Inqaba Biotech (Pretoria, South Africa) conducted synthesis of oligonucleotide primers and performed all sequencing to confirm the identity of the plasmids.

2.2 Experimental

2.2.1 Generation of hGSTA1-1 mutants

The plasmid DNA encoding the V58A and I75A hGSTA1-1 mutants were created using site-directed mutagenesis. Oligonucleotide primers were designed in accordance with the published wild-type nucleotide sequence of hGSTA1-1 (Stenberg *et al*., 1992) with the aid of the Primer-X software (http://bioinformatics.org/primerx) and Gene Runner software v3.04 (Hastings Software Inc., NY, USA). The sequences of the primers synthesised used to construct the V58A and I75A hGSTA1-1 mutants are listed below with the respective codons coding for Ala underlined:

V58A

Forward: 5′ CAAGTGCCAATG**GCA**GAGATTGATGGGATGAAGCTG 3′

Reverse: 5′ CAGCTTCATCCCATCAATCTC**TGC**CATTGGCACTTG 3′

I75A:

Forward: 5′ GCCATTCTCAACTAC**GCC**GCCAGCAAATACAAC 3′

Reverse: 5′ GTTGTATTTGCT**GGC**GGCGTAGTTGAGAATGGC 3′

The DNA encoding for the mutant hGSTA1-1 proteins was generated by following the protocol described in the QuikChange® Lightning mutagenesis kit from Stratagene (La Jolla, CA USA) (Braman *et al.*, 1996). The sample reaction had a final volume of 51 µl which comprised of 5 µl (10x) reaction buffer, 1 µl (50 ng) double stranded DNA template, 1 µl of each (125 ng) primer (forward and reverse), 1 µl dNTP mix, 1.5 µl QuikSolution reagent, 39.5 µl Milli-Q water and 1 µl (2.5 U/µl) QuikChange Lightning enzyme. The product was generated by 18 amplification cycles of 20 seconds at 95°C to denature the wild-type dsDNA, 10 seconds at 60°C to anneal the mutant primers and 120 seconds at 68°C for DNA extension. Parental DNA template was digested with 2 µl (10 U/µl) Dpn I for 5 minutes at 37°C. The reaction products were then used to transform (section 2.2.2) *Escherichia coli* XL10-Gold ultracompetent cells supplied with the mutagenesis kit. The cells were plated onto LB agar plates [1% (w/v) tryptone, 0.5% (w/v) yeast extract, 1.0% (w/v) NaCl, 1.5% (w/v) agar] supplemented with ampicillin (100 µg/ml) and were incubated for 12-16 hours at 37°C. Colonies were chosen at random and overnight cultures were made in LB medium (1% (w/v) tryptone, 0.5% (w/v) yeast extract, 1.0% (w/v) NaCl). Plasmid DNA was then extracted from the overnight culture of cells, using the Genejet plasmid miniprep kit from Fermentas Life Sciences (St. Leon-Rot, Germany). The identity of the desired mutation, and that no other mutations were generated during the mutagenesis amplification reaction, was confirmed by DNA sequencing (Inqaba Biotech; Pretoria, South Africa) of the plasmid DNA, using the pKHA1 forward and reverse primers.

2.2.2 Transformation, over-expression and purification of hGSTA1-1

The expression system was constructed by transforming *Escherichia coli* BL21(DE3)/pLysS cells (Lucigen, Middleton, WI, USA) with the plasmids encoding for the wild-type, V58A and I75A hGSTA1-1 proteins. The bacterial cells were transformed using a one-step method as described by (Chung *et al.*, 1989). Competent cells (50 µl) were thawed on ice for 10 minutes to which 1 µl of mutant plasmid DNA (100 ng/µl) was added and the reaction mixture incubated on ice for 30 minutes. The cells were then heat-shocked at 42°C for 45 seconds on a heating block, followed by a rapid transfer to ice for 2 minutes. SOC medium [2% (w/v) tryptone, 0.5% (w/v) yeast extract, 250 mM KCl, 1 M glucose, 2 M $MgCl_2$] was added to the reaction mixture

followed by incubation at 37°C for 90 minutes. The cells were then plated on LB-agar plates [1% (w/v) tryptone, 0.5% (w/v) yeast extract, 1.0% (w/v) NaCl, 1.5% (w/v) agar] supplemented with the antibiotics ampicillin (100 μg/ml) and chloramphenicol (100 μg/ml). The plates were then incubated at 37°C for 12 - 16 hours. The transformed *Escherichia coli* expression cells transformed with the respective plasmid containing the cDNA sequence coding for the hGSTA1-1 protein were added to fresh, sterile 2xYT medium [1.6% (w/v) tryptone, 1.0% (w/v) yeast extract, 0.5% (w/v) NaCl] supplemented with ampicillin (100 μg/ml) and chloramphenicol (100 μg/ml). The cells were grown at 37°C with shaking at 250 rpm for 12 - 16 hours of which a 50-fold dilution was then used to inoculate into fresh, sterile 2xYT medium supplemented ampicillin (100 μg/ml) and chloramphenicol (100 μg/ml). Cells were grown at 37°C with shaking at 250 rpm till an OD_{600} of ~ 0.6 was reached after which over-expression of the hGSTA1-1 proteins were induced by the addition of 1 mM IPTG. The cells were grown for a further 12 - 16 hours at 37°C, with shaking at 250 rpm, in order to achieve optimum protein expression. Cells were harvested via centrifugation (4200 × *g*, 20 min) and re-suspended in $1/20^{th}$ of the culture volume of re-suspension buffer [20 mM sodium phosphate buffer, pH 7.45, 1 mM EDTA, 0.02 % (w/v) NaN_3]. DNAse I (50 μg/ml) and lysozyme (100 μg/ml) were added to the cell suspension and was rotated at 4°C for 30 minutes. The cells were then sonicated on ice for 4 cycles of 30 seconds with 40% pulse intensity using an ultrasonic liquid processor (Misonix Inc. (model: XL-2020), Farmingdale, NY, USA). The lysed cells were centrifuged (16 100 × *g*, 30 min) at 4°C. Samples of whole cell extract, soluble and insoluble fractions were electrophoresed on 12% acrylamide SDS-PAGE gels (section 2.2.3).

The hGSTA1-1 proteins were purified using cation-exchange chromatography as described by Wallace and Dirr (1999). The soluble fraction after centrifugation was applied onto a CM-Sepharose column (GE Healthcare Life Sciences, Uppsala, Sweden) pre-equilibrated with buffer A [10 mM sodium phosphate buffer, pH 7.45, 0.02% (w/v) NaN_3]. The column was then washed with 10 column volumes of buffer A. Bound hGSTA1-1 was eluted off the column using a linear NaCl gradient (10 column volumes) from 0 M to 0.3 M NaCl. The gradient was produced by mixing buffer A with 10 mM sodium phosphate buffer, pH 7.45, 1 mM EDTA, 300 mM NaCl, 0.02% (NaN_3). Cation-exchange chromatography was conducted using an

ÄKTA Prime system attached to a computer with PrimeView 1.0 software (GE Healthcare Life Sciences, Uppsala, Sweden). The purified hGSTA1-1 proteins was then dialysed against three changes of storage buffer [20 mM sodium phosphate buffer, pH 7.4, 0.02% (w/v) NaN_3], lasting 4 hours each, after which it was snap frozen and stored at -80ºC. Unless otherwise stated, all experiments for the hGSTA1-1 proteins were conducted in the storage buffer.

2.2.3 SDS-PAGE

The solubility, homogeneity and purity of the expressed mutant proteins were assessed by separation on a 12% SDS-PAGE (Laemmli, 1970) gel. The discontinuous gel system consisted of a 4% (w/v) acrylamide/bis-acrylamide stacking gel [0.1% (w/v) SDS, 0.05% (w/v) ammonium persulphate, 0.1% (w/v) TEMED and 125 mM Tris-HCl buffer, pH 6.8] and a 12% (w/v) acrylamide/bis-acrylamide separating gel [0.1% (w/v) SDS, 0.05% (w/v) ammonium persulphate, 0.1% (w/v) TEMED and 375 mM Tris/HCl, pH 8.8]. Protein samples were diluted five-fold with sample buffer [10% (w/v) glycerol, 2% (w/v) SDS, 5% (w/v) β-mercaptoethanol, 0.05% (w/v) bromophenol blue and 62.5 mM Tris-HCl buffer, pH 6.8]. Samples were then boiled for 5 minutes to ensure that the proteins were denatured. The electrode buffer used contained 1% (w/v) SDS, 192 mM glycine and 25 mM Tris, pH 8.5. The protein samples were applied to the SDS-PAGE wells (15μl) and electrophoresed at 140 V for 2 hours using a Hoefer MiniVE electrophoresis system (Holliston, MA, USA). The molecular weight marker (Fermentas Life Sciences, (cat. no.: SM0431), Ontario, Canada) used contained a mixture of seven proteins: β-galactosidase (116 kDa), bovine serum albumin (66.2 kDa), ovalbumin (45 kDa), lactate dehydrogenase (35 kDa), restriction endonuclease Bsp98I (25 kDa), β-lactoglobulin (18.4 kDa) and lysozyme (14.4 kDa). The gels were stained in 2% (w/v) Coomassie Blue R250 staining solution containing 13.5% (v/v) glacial acetic acid and 18.75% (v/v) ethanol and destained with 7% (v/v) ethanol and 5% (v/v) glacial acetic acid until the background was clear.

2.2.4 Absorbance spectroscopy

Most experiments conducted in this study involved a spectroscopic technique which allows the experiment to be done under solution conditions and relatively low

amounts of protein are consumed. It is therefore imperative that the concentrations of protein and its ligands are accurately determined.

The concentrations of the hGSTA1-1 proteins, ANS and GSH were determined spectrophotometrically using a Jasco V-630 UV-VIS spectrophotometer (Jasco Inc., Tokyo, Japan) and by applying the Beer-Lambert law:

$$A = \varepsilon_\lambda c \ell \quad (3)$$

where A is the absorbance at the respective wavelength, ε_λ is the molar extinction of the absorber at wavelength λ, c is the concentration of the absorbing solution and ℓ is the path length of light through the solution (cuvette).

A molar extinction coefficient (ε_λ) of 38 200 $M^{-1}.cm^{-1}$ (Stenberg *et al.*, 1992) was used for determining the concentrations of wild-type and V58A and I75A hGSTA1-1 proteins at 280 nm. The same extinction coefficient is applicable for both mutant proteins as the amino acids involved in the mutation are not chromophores. The extinction coefficients used for determining ANS concentration at 350 nm was 5000 $M^{-1}cm^{-1}$ (Weber and Young, 1964). The quantification of GSH, which employed the DTNB assay (section 2.2.7), at 412 nm used an extinction co-efficient of 13 600 $M^{-1}cm^{-1}$ (Habeeb *et al.*, 1972) for the 2-nitro-5-thiobenzoate anion.

The concentrations were determined by fitting a linear regression to 10 or more points from a serial dilution. All readings were buffer corrected with the appropriate buffer used for the concentration determination.

2.2.5 Circular dichroism spectroscopy

The secondary structure of proteins is chiral and polarised light interacts in a specific manner with chiral molecules and are thus optically active. Circular dichroism (CD) is a technique that measures the differential absorption of left- and right-handed circularly polarised light by such optically active molecules. Any conformational change in a protein can therefore be detected by CD. Optical activity in proteins arises from disulfide groups, aromatic side chains, and the peptide backbone (Woody, 1995). Disulfide groups and aromatic amino acids have characteristic absorption bands in the

near-UV range (250 - 300 nm). In the far-UV region (180 - 250 nm), the predominant signal arises from the peptide backbone. The adoption of different secondary structures by the peptide backbone results in distinctive CD spectra (Woody, 1995). As a result, this wavelength range gives a good indication of the secondary structural content of proteins. Proteins with a high α-helical content display characteristic minima at 208 and 222 nm, as well as a stronger positive band near 190 nm (Woody, 1995).

Far-UV CD spectra (180 - 250 nm) were recorded using 2 μM hGSTA1-1 protein in storage buffer. In some cases, a stock protein solution (20 μM) and the storage buffer was diluted 10 fold in order to improve the signal to noise ratio at the lower wavelengths. All far-UV CD spectra were recorded at 20°C and represent an average of 10 accumulations, at a scan speed of 200 nm/min. The bandwidth used was 1 nm and the data pitch 0.2 nm. Measurements were obtained using on a Jasco J-810 spectropolarimeter with Spectra Manager software v1.5.00 (Jasco Inc., Tokyo, Japan) using a pathlength of 2 mm. All spectra were buffer corrected and were normalised by calculating the mean residue ellipticity [θ] (deg.cm^2/dmol/residue) using the following equation (Woody, 1995):

$$[\theta] = (100 \times \theta) / c \times n \times \ell \qquad (4)$$

where (θ) is the ellipticity signal in mdeg, c is the protein concentration in mM, n is the number of residues in the protein chain and ℓ is the path length in cm.

Near-UV CD spectra (250 – 350 nm) were recorded using 20 μM hGSTA1-1 protein at 5 °C in storage buffer. The temperature was maintained by a Jasco PTC-423S Peltier-type temperature control system. All near-UV CD spectra were recorded at a scanning speed of 100 nm/min, using a 10 mm path length cuvette. The sensitivity was set to high (10 mdeg), data pitch was 0.05 nm, response: 1 sec, bandwidth: 0.5 nm, and each spectrum was the result of 10 accumulations. Low temperature CD measurements on proteins sharpen the CD bands because of lowered motility of the side-chains, and at very low temperatures increase the intensity of the signal (Strickland, 1974). Near-UV CD signal data were not normalised to mean residue

ellipticity since only four types of residues contribute to the signal, making averaging over all residues unjustified.

2.2.6 Fluorescence spectroscopy

Fluorescence is the emission that results from the return of an unpaired electron from the excited to the ground state (Lakowicz, 2006). The energy lost between excitation and emission, known as Stokes shift, results in the bathochromic (red) shift of emission spectra. In proteins, the naturally occurring fluorophores are tryptophan, tyrosine and phenylalanine. Due to the small quantum yield of phenylalanine in proteins, its emission is rarely observed (Lakowicz, 2006). The fluorescence of most proteins is dominated by tryptophan, with its quantum yield being more than double that of tyrosine. In the native state, tyrosine emission is quenched by energy transfer to tryptophan, and to quenching due to nearby charged carboxyl and uncharged amino groups (Lakowicz, 2006). The indole ring of tryptophan is highly sensitive to solvent polarity (Lakowicz, 2006). Emission spectra of this residue reflect the polarity of its surrounding environment. Therefore, tryptophan fluorescence is used to monitor tertiary structural changes in proteins. All fluorescence measurements were recorded in a quartz cuvette with a 10 mm path-length using a Jasco FP 6300 fluorimeter with SpectraManager software, v1.2 (Waltham, MA, USA).

2.2.6.1 Intrinsic fluorescence-tryptophan fluorescence

The hGSTA1-1 protein has a lone tryptophan residue (W21) in the N-terminal domain whose indole ring protrudes into the C-terminal domain and 10 tyrosine (Tyr) residues (3 in domain 1 and 7 in domain 2). For optimal signal the hGSTA1-1 proteins were excited at 280 nm. Fluorescence emission spectra were recorded using 2 μM hGSTA1-1 from 280 - 500 nm. The excitation and emission slit widths were 5 nm each. The spectra were recorded at 20°C, buffer corrected, and reflect averages of three accumulations at a scan speed of 500 nm/min.

2.2.6.2 Extrinsic fluorescence - ANS binding

8-anilino-1-naphthalene sulphonate (ANS) is a hydrophobic dye used as an extrinsic fluorescence probe (Engelhard and Evans, 1995). It binds to hydrophobic patches in proteins. In an aqueous environment ANS fluorescence is quenched, but upon binding

to a hydrophobic surface its fluorescence quantum yield increases and a hypsochromic shift of its maximum emission wavelength is witnessed (Engelhard and Evans, 1995).

A stock solution (10 mM) of ANS was prepared in storage buffer. The concentration of ANS was checked by recording the absorbance at 350 nm and using an extinction coefficient of ε_{350} = 5000 $M^{-1}.cm^{-1}$ (section 2.2.4). Native protein (2 μM) was incubated for 30 minutes with ANS to a final concentration of 200 μM. In the case of urea induced equilibrium unfolding studies, the proteins were incubated for 60 min at different concentrations of urea (0 - 8 M) with ANS added to the protein/urea mixture to a final concentration of 200 μM. The solution was incubated for at 30 min to achieve equilibrium. A series of blanks was generated, each containing 200 μM ANS with the appropriate urea concentration (0 – 8 M). The samples were excited at 390 nm and emission spectra were recorded from 390 to 600 nm. Spectra were produced from an average of three accumulations at 500 nm/min scan speed at 20°C in storage buffer, using excitation and emission slit widths of 5 nm. The fluorescence emission intensities at 465 nm were extracted and plotted as a function of urea concentration.

2.2.7 Glutathione (GSH) quantification

Glutathione (GSH) is a substrate used in the standard GSH/CDNB assay (Habig and Jakoby, 1981) and can be accurately quantified by using a DTNB assay (Thannhauser *et al.*, 1984). The DTNB assay was conducted by titration of a dilution series GSH in 20 mM sodium phosphate buffer, pH 7.4, 1 mM EDTA, 0.02 % (w/v) NaN_3, with 0.2 mM DTNB. The 2-nitro-5-thiobenzoate anion (absorbs maximally at a wavelength of 412 nm) is released when DTNB reacts with free thiol groups (Habeeb *et al.*, 1972). The concentration of the released 2-nitro-5-thiobenzoate anion was determined spectrophotometrically at 20ºC as described previously (section 2.2.4). The concentration of GSH was determined immediately after preparation and stored at -20 ºC for further use.

2.2.8 Standard GSH-CDNB conjugation assay

The specific activities of the hGSTA1-1 proteins were determined using the standard GSH-CDNB conjugation assay (Habig and Jakoby, 1981). The enzyme catalysed conjugation of 1-chloro-2,4-dinitrobenzene (CDNB) to GSH was monitored spectrophotometrically by measuring the absorbance at 340 nm upon formation of the

chromophoric product 1-(*S*-glutathionyl)-2,4-dinitrobenzene. Triplicate sample solutions of 1–10 nM hGSTA1-1 protein were prepared in 0.1 M sodium phosphate buffer, pH 6.5, containing 1 mM EDTA and 0.02% sodium azide, in the presence of 1 mM GSH (added fresh to reduce the risk of oxidation). The conjugation reaction was initiated by the addition of 100 µl of a 30 mM stock solution of CDNB (solubilised in ethanol) to create a final assay CDNB concentration of 1 mM and a final assay volume of 3.0 ml. The reaction was followed as linear progress curves by measuring the absorbance at 340 nm for 1 min using a Jasco V–630 UV/Vis spectrophotometer at 20 °C. The change in concentration of the 1-(*S*-glutathionyl)-2,4-dinitrobenzene complex formed at the active sites was calculated using the extinction coefficient of the chromophore at 340 nm (ε_{340} = 9600 M^{-1} cm^{-1}) (Habig and Jakoby, 1981) after correcting for the absorbance contribution from the 1-(*S*-glutathionyl)-2,4-dinitrobenzene complex formed in the absence of enzyme. The specific activity (µmol/min/mg) was determined by linear regression as the slope of a plot between the initial velocity of complex formation (µmol/min) versus the amount of protein (mg).

2.2.9 Thermal denaturation studies

Thermal or chemical unfolding is generally used to determine the thermodynamic stability of a protein, which is the Gibbs free energy difference (ΔG) between the folded (G_F) and the unfolded (G_U) states. In order to obtain thermodynamic parameters from thermal unfolding experiments, e.g. melting temperature (T_m), enthalpy [$\Delta H(T_m)$], and heat capacity increment (ΔC_p), the key requirement is that protein unfolding is thermodynamically reversible which is usually only valid for chemical denaturation while thermal denaturation is often irreversible. The general root cause of this irreversibility is aggregation of the heat-induced unfolded polypeptide (Benjwal *et al.*, 2006).

Temperature induced unfolding studies were conducted to assess the relative stabilities of wild-type and the mutant hGSTA1-1 proteins. Temperature-induced denaturation of hGSTA1-1 was determined by monitoring the ellipticity at 222 nm (section 2.2.5) over the temperature range 20°C to 80°C. The temperature unfolding profiles were recorded using 2 µM hGSTA1-1. The temperature was controlled by a Jasco PTC-423S Peltier-type temperature control system and the rate at which the

temperature was increased was 1°C.min^{-1}. The bandwidth was 1 nm and data pitch 0.2°C. Measurements were obtained using a Jasco J-810 spectropolarimeter with Spectra Manager software v1.5.00 (Jasco Inc., Tokyo, Japan) and a pathlength of 2 mm. The temperature unfolding profiles were normalised by calculating the mean residue ellipticity [θ] (deg.cm^2.dmol^{-1}.residue^{-1}).

2.2.10 Urea - induced equilibrium unfolding studies of the hGSTA1-1 proteins

The impact of a mutation on the conformational stability of a protein can be determined by unfolding the protein under equilibrium conditions using a chemical denaturant (Pace, 1986). In order to do so, its recovery to the native state from the unfolded state and reversibility of the unfolding event need to be established.

2.2.10.1 Recovery of unfolding

Determining recovery (resembling native – like structure) establishes that the protein is able to refold after being denatured thus displaying equilibrium between the native and unfolded states. The recovery of unfolding was determined for wild-type and mutant hGSTA1-1 proteins.

The refolding of unfolded protein samples, incubated in 8 M urea for one hour at 20°C, was achieved by a 10-fold dilution of each sample reaction with storage buffer for 1 hour at 20°C. The refolded state of the protein was assessed using far-UV CD (section 2.2.5) and intrinsic tryptophan fluorescence (section 2.2.6).

2.2.10.2 Urea - induced equilibrium unfolding/refolding

A denaturant is used to shift the equilibrium from the native to the unfolded state. The equilibrium constant (K_{eq}) can be calculated and hence the thermodynamic conformational stability parameters: $\Delta G(H_2O)$ and *m*-value can be determined (section 2.2.11) provided reversibility is established.

Determining recovery alone does not imply that the unfolding of a protein is reversible along the same pathway as the equilibrium between the two states cannot be observed under such conditions. Therefore the reversibility of unfolding must be determined under conditions where the equilibrium between native and unfolded

states can be witnessed i.e. in transition region of unfolding/refolding. Establishing reversibility also determines hysteresis (if any), i.e. unfolding pathway different to that of refolding pathway.

Urea unfolding and refolding (reversibility) of the hGSTA1-1 proteins were conducted over a range of urea concentrations from 0 M to 8 M urea and vice versa (refolding) in the absence and presence of ANS (section 2.2.6.2). Unfolding was conducted at 20ºC for 1 hour to allow equilibrium to be reached. Thereafter, the samples for the range of urea concentrations were monitored using far-UV CD (section 2.2.5) and fluorescence spectroscopy (section 2.2.6).

All of the urea used for experiments was prepared according to methods previously described (Pace, 1986) using storage buffer as the solvent. Following preparation, the pH of the stock urea solution was adjusted to pH 7.4, filtered and the concentration of 10 M confirmed using an Atago R5000 refractometer (Tokyo, Japan) and the method of (Pace, 1986).

2.2.11 Data fitting

Unfolding/refolding data obtained for all the proteins were analysed according to either a two-state or three-state unfolding process for a dimeric protein.

2.2.11.1 Two-state fitting model

During two-state reversible unfolding there is an equilibrium reached between the native [N_2 (dimer)] and the unfolded [2U (monomer)] species:

$$N_2 \overset{K_{eq}}{\rightleftharpoons} 2U \qquad (5)$$

During a two-state unfolding transition, only the unfolded and native states are present at significant concentrations (Pace, 1986).

Therefore, for a two-state mechanism:

$$f_N + f_U = 1 \qquad (6)$$

where f_N is the fraction folded or native protein and f_U is the fraction unfolded protein. At any point during unfolding, there is a contribution to the signal from the concentration of both species:

$$y = y_N f_N + y_U f_U \quad (7)$$

where y is the signal obtained for the respective spectroscopic probe, f_N represents the fraction of folded protein, f_U represents fraction unfolded protein. In addition y_N represents the y value for the folded state and can be extrapolated from linear pre-transition region of the unfolding data and y_U represents the y value for the unfolded state and can be extrapolated from the linear post-transition region of the unfolding data (Figure 2-1). By combining equations 6 and 7, the fraction of unfolded protein can be obtained:

$$f_U = (y_N - y) / (y_N - y_U) \quad (8)$$

similarly the fraction of folded or native protein can be obtained:

$$f_N = (y - y_U) / (y_N - y_U) \quad (9)$$

The equilibrium constant for the unfolding reaction (K_{eq}) is:

$$K_{eq} = f_U / f_N = [U]^2 / [N_2] = 2 P_t [f_U^2/(1-f_U)] \quad (10)$$

So, therefore, if equations 8 and 9 are substituted into 10:

$$K_{eq} = (y_N - y) / (y - y_U) \quad (11)$$

And

$$\Delta G° = - RT \ln K_{eq} \quad (12)$$

where $\Delta G°$ is the free energy of unfolding, R is the gas constant, T is temperature in Kelvin and K_{eq} is the equilibrium constant for a reaction. In order to determine $\Delta G(H_2O)$ it is assumed that $\Delta G°$ has a linear dependence on denaturant concentration [D] for all urea concentrations (Tanford, 1968, 1970).

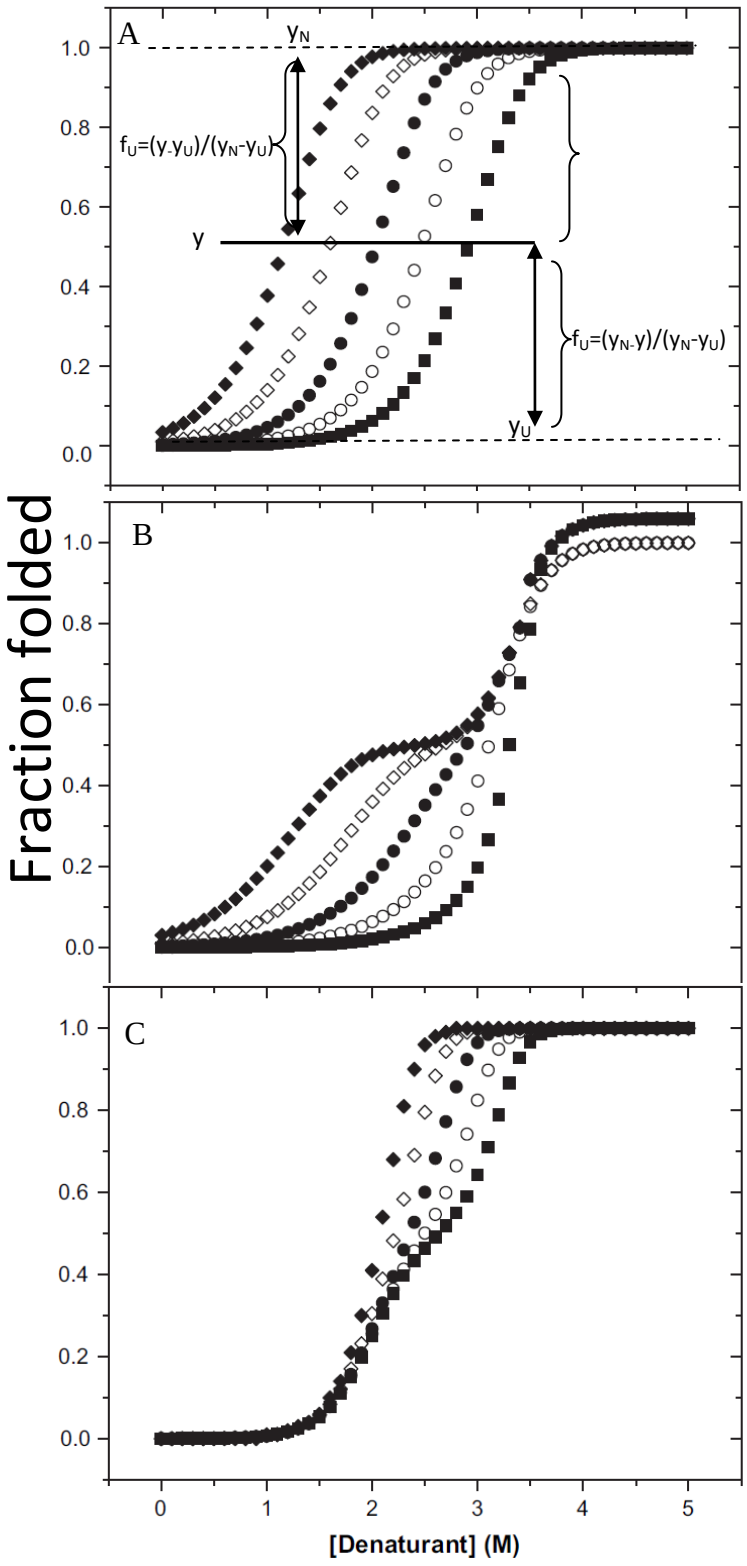


Figure 2-1. Urea denaturation curve
Equilibrium unfolding analysis of a (A) two state mechanism (B) three-state with monomeric intermediate and (C) three-state with dimeric intermediate. Dimeric protein concentrations of 0.1 µM (♦), 1 µM (◊), 10 µM (●), 100 µM (○) and 1000 µM (■). Image adapted from Rumfeldt *et al*. (2008).

Therefore:

$$\Delta G° = \Delta G(H_2O) - m\,[D] \quad (13)$$

where $\Delta G(H_2O)$ represents the free energy difference between the folded and unfolded states in the absence of denaturant, *m* is the *m*-value for the dependence of free energy on denaturant concentration which is also an indicator of co-operativity and can be related to the change in solvent-accessible surface area ($\Delta SASA$) , and [D] is the denaturant concentration.

Combining equations 11, 12 and 13 thus gives:

$$y = [y_N + y_U * e^{-(\Delta G\,(H2O) - m\,[D])/RT}] / [1 + e^{-(\Delta G\,(H2O) - m\,[D])/RT}] \quad (14)$$

The equilibrium unfolding data obtained were fitted to equation 14 using SigmaPlot version 11.0 (Systat Software Inc; Chicago, IL, USA) and the parameters $\Delta G(H_2O)$ and *m* were obtained.

2.2.11.2 Three-state fitting model

For a three-state dimer unfolding transition, a protein may exist in the native (N_2), intermediate (I_2) and unfolded (2U) states:

$$N_2 \overset{K_1}{\rightleftharpoons} I_2 \overset{K_2}{\rightleftharpoons} 2U \quad (15)$$

where K_1 represents the equilibrium constant for the N $\leftrightarrow$ I transition and K_2 represents the equilibrium constant for the I $\rightleftharpoons$ U transition.

As the native, unfolded and intermediate conformations in the three-state unfolding transitions are the only species present,

$$f_N + f_I + f_U = 1 \quad (16)$$

$$K_1 = f_I / f_N \; K_2 = f_U / f_I \; K_U = f_U / f_N = K_1K_2 \quad (17)$$

where f_I is the fraction of intermediate present. The equilibrium constant for N $\rightleftharpoons$ U is represented by K_1K_2. y (signal obtained for the respective spectroscopic probe) can be represented as:

$$y = y_N f_N + y_I f_I + y_U f_U \quad (18)$$

where y_I depicts the measured signal of the intermediate. y_N and y_U are estimated from the linear extrapolation of the pre- and post-transition baselines, respectively.

By solving f_U in terms of K_1 and K_2 and by combining equations (16) and (17):

$$f_U = (K_1K_2)/(K_1K_2 + 1 + K_1) \quad (19)$$

Rearranging and substituting equation (17) into equation (18) and solving for f_N and f_I in terms of K_1 and K_2:

$$f_N = f_U / K_1K_2$$

therefore,

$$f_N = 1/[(K_1K_2 + 1 + K_1)/K_1K_2] \quad (20)$$

Similarly

$$f_I = K_1/[(K_1K_2 + 1 + K_1)/K_1K_2] \quad (21)$$

Substituting equations (19), (20) and (21) into equation (18) gives:

$$y = (y_N + y_I K_1 + y_U (K_1K_2))/(1 + K_1 + K_1K_2) \quad (22)$$

Given that:

$$\Delta G_1 = -RT\ln K_1 \text{ hence } K_1 = e^{-\Delta G1/RT} \quad (23)$$

And

$$\Delta G_1 = \Delta G_2(H_2O) - m_1[\text{denaturant}] \quad (24)$$

By combining equations (23) and (24):

$$K_1 = e^{[-\Delta G1(H_2O)-m1[\text{denaturant}]]/RT} \quad (25)$$

Similarly,

$$K_2 = e^{[-\Delta G2(H_2O)-m2[\text{denaturant}]]/RT} \quad (26)$$

Substituting equations (25) and (26) into (22) the final equation to fit the data is achieved:

$$y = [y + y(e^{[-\Delta G1(H2O)-m1[\text{denaturant}]]/RT}) + y(e^{[-\Delta G1(H2O)m1[\text{denaturant}]]/RT})(e^{[-\Delta G2(H2O)-m2[\text{denaturant}]]/RT})]/[1 + (e^{[-\Delta G1(H2O)-m1[\text{denaturant}]]/RT}) + (e^{[-\Delta G1(H2O)-m1[\text{denaturant}]]/RT})(e^{[-\Delta G2(H2O)-m2[\text{denaturant}]]/RT})] \quad (27)$$

2.2.12 Pulse proteolysis

The native and unfolded forms of protein can be distinguished by the proteolytic digestion coupled with electrophoresis. Unlike the spectroscopic probes, the enzymatic digestion of proteins is independent of secondary or tertiary structural content. Under equilibrium conditions ($N \rightleftharpoons U$), only the unfolded population of protein will be digested in a short incubation time (pulse proteolysis) whereas the folded/native protein would remain intact unless the protein contains an intrinsically disordered region (Park and Marqusee, 2005). Pulse proteolysis is therefore an alternate method to estimating relative protein stability.

The hGSTA1-1 proteins (10 µM) were prepared in 20 mM Tris-HCl buffer, pH 7.5, 50 mM NaCl, 10 mM $CaCl_2$ and incubated in different concentrations of urea as for the urea-induced equilibrium unfolding experiments (section 2.2.10.2). The protein samples were then developed as by the method of Park and Marqusee (2005). Two µl of thermolysin (10 mg/ml) was added to the hGSTA1-1 protein solution pre-equilibrated with urea at 20°C and proteolytic activity was allowed to proceed for 60 seconds. This duration is sufficient for 'pulse proteolysis' as proteins are in equilibrium between the folded and unfolded states and both populations can be detected. Proteolytic activity was then quenched by adding 15 µl of 50 mM EDTA. Samples were prepared for SDS-PAGE (Laemmli, 1970; section 2.2.3).

2.2.13 Crystallography

An accurate 3-D representation of biomolecular structures such as proteins is best provided by X-ray crystallography. This technique provides atomic resolution of the molecule of interest. Briefly, a crystal (repeating unit of the molecule) is diffracted by X-rays yielding an electron density map from which the relative positions of the atoms are solved. The fundamental step in protein crystallography is obtaining large crystals of pure protein.

2.2.13.1 Crystal growth

Crystals of the V58A and I75A hGSTA1-1 protein were grown as described previously (Achilonu *et al*., 2010; Gildenhuys *et al*., 2010a). Crystals of both mutant proteins were grown by the hanging-drop vapour-diffusion method at 293 K using a

24-well microplate. Each hanging drop (4, 6 or 8 µl) was comprised of equal volumes of protein stock solution and reservoir buffer. The stock protein concentration was 7 or 10 mg/ml in 0.1 M Tris–HCl pH 7.5 containing 10 mM DTT and 0.02% sodium azide. The reservoir buffer contained in 0.1 M Tris–HCl pH 7.5, 10 mM DTT, 0.02% sodium azide and PEG 3350 [18% (w/v)] in the case of the V58A crystal and PEG 4000 [17% (w/v)] in the case of the I75A crystal.

2.2.13.2 Diffraction

X-ray diffraction data for V58A and I75A hGSTA1-1 were collected using an in-house Bruker X8 Proteum system with a Microstar copper rotating-anode generator with Montel 200 optics, a PLATINUM 135 CCD detector and an Oxford Cryostream Plus system. Crystals were cooled to 113 K in a stream of nitrogen during data collection and images were collected covering an oscillation angle of 0.5° per image. The data sets were processed using APEX and SAINT software (Bruker AXS Inc., Madison, Wisconsin, USA).

2.2.13.3 Refinement

The structures of both mutants were solved by molecular replacement using MOLREP (Vagin & Teplyakov, 2000) as implemented in the CCP4 suite of programs (Collaborative Computational Project, Number 4, 1994), using wild-type hGSTA1-1 (PDB code 1K3O; Le Trong *et al*., 2002) as the search model. The refinement was perfomed with REFMAC5 (Murshudov *et al*., 1997) and model building was performed with Coot (Emsley & Cowtan, 2004). Stereochemical validation of the model was performed using and MolProbity (Chen *et al*., 2010).

CHAPTER 3. RESULTS

3.1 Sequence identity

The pKHA1 plasmid containing the open reading frames (ORF) encoding the wild-type, V58A and I75A hGSTA1-1 proteins were sequenced. The regions where the DNA was mutated are highlighted in Figure 3-1. The presence of the engineered mutation was confirmed and no further mutations were found to be introduced during the thermal cycling reactions.

3.2 Over-expression and purification

The over-expression of the recombinant hGSTA1-1 proteins was conducted as described in section 2.2.2. All proteins were found to be soluble as determined by SDS-PAGE of the cytosol and pellet fractions (Figure 3-2A). The soluble fractions, after sonication and centrifugation, containing wild-type and mutant hGSTA1-1 proteins were purified by means of cation-exchange chromatography and eluted with a NaCl gradient (section 2.2.2). The fractions containing protein were analysed on SDS-PAGE (section 3.3).

3.3 Size and purity determination

SDS-PAGE was used to determine the molecular mass and purity of wild-type and mutant hGSTA1-1 proteins. The purity of the hGSTA1-1 protein was judged to be electrophoretically pure as seen by the single bands (Figure 3-2A). The hGSTA1-1 proteins were found to have a subunit molecular mass of ~ 26 kDa (Figure 3-2B). This is in agreement with the published data for wild-type hGSTA1-1 (Stenberg *et al*., 1992).

3.4 Absorbance spectroscopy

The concentrations of the proteins were determined using absorbance spectroscopy (section 2.2.4). A yield between 250 - 300 mg of pure protein was obtained from 1 ℓ of culture. Absorbance spectra of the hGSTA1-1 proteins (Figure 3-3) indicated that the protein stock solutions were free from aggregates (negligible absorbance 340 nm) and nucleic acids (A_{280}/A_{260} = 1.8)

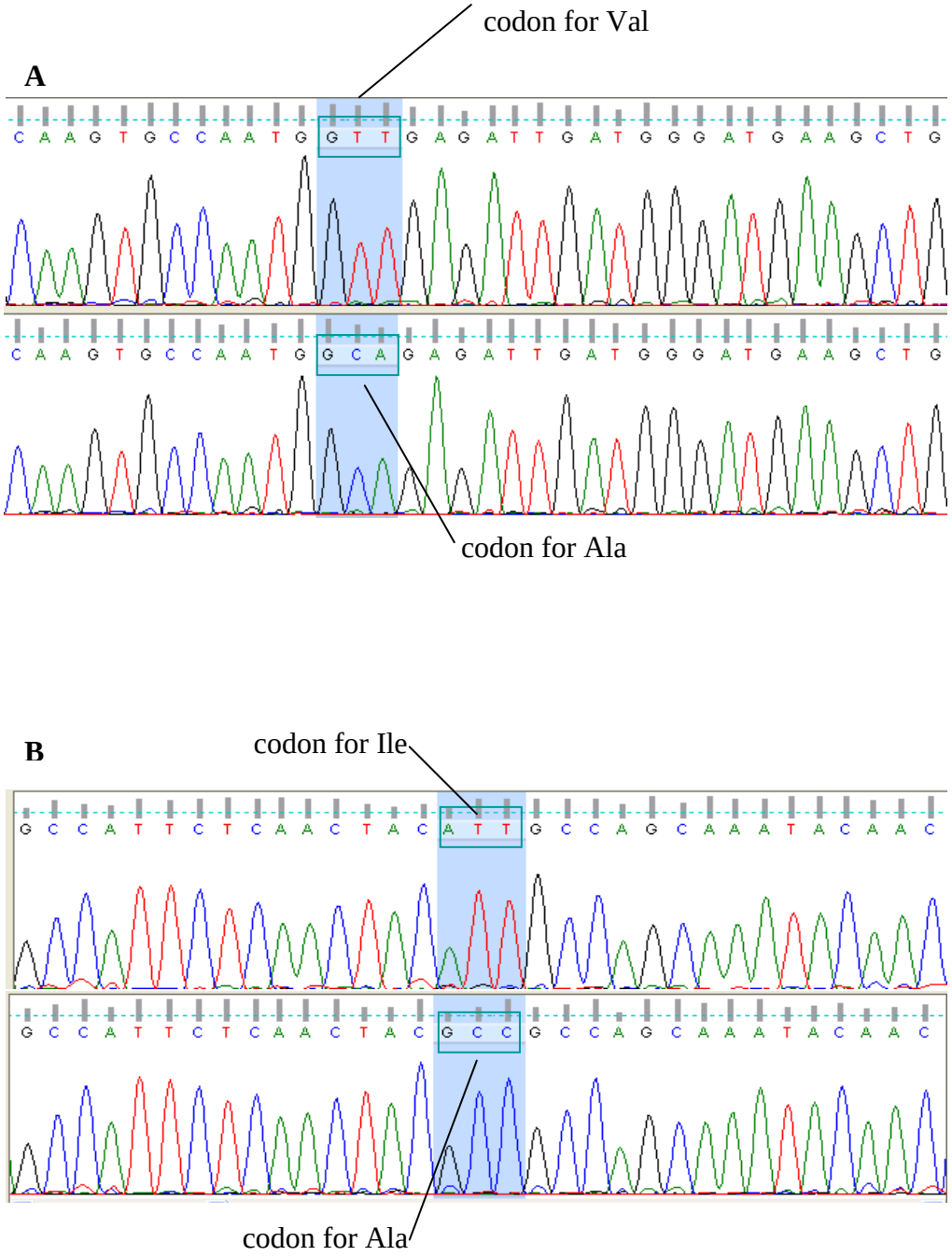


Figure 3-1. Plasmid sequencing results of the V58A and I75A encoding mutations.
A selected segment of the pKHA1 plasmid sequence encoding wild-type (upper panels) and mutants (lower panel). The mutated codons for (A) V58A and (B) I75A hGSTA1-1 proteins are boxed and highlighted. The sequencing results were viewed using the program Finch TV version 1.4.0 (http://www.geospiza.com/FinchTV: Geospiza Inc.).

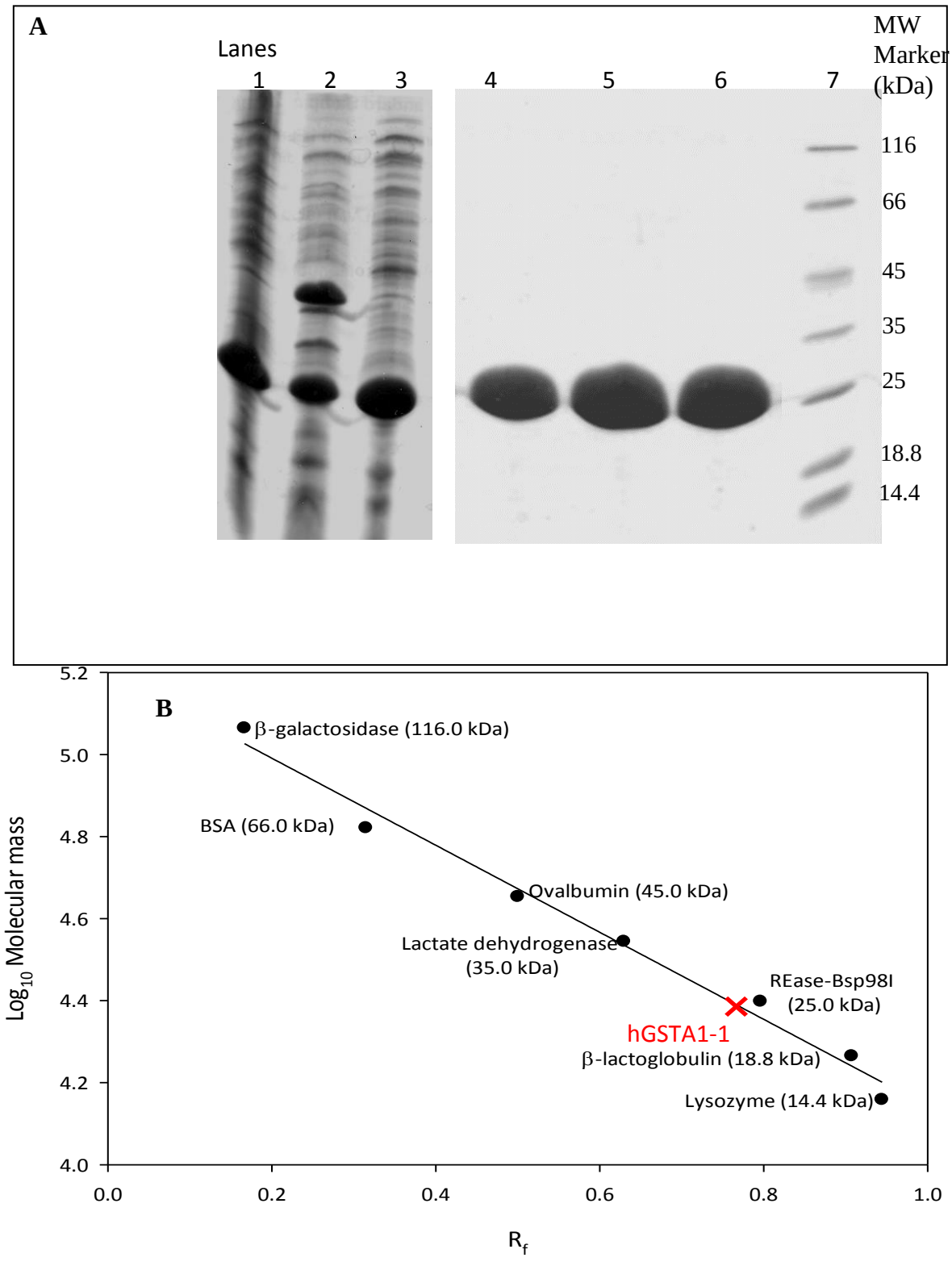


Figure 3-2. SDS-PAGE analysis of hGSTA1-1 eluted proteins from cation-exchange chromatography column and size determination.
(A) SDS-PAGE gel of lysate (lane 1), pellet (lane 2) and cytosolic (lane 3) and the pooled hGSTA1-1 proteins displaying its molecular mass: I75A hGSTA1-1 (lane 4), wt hGSTA1-1 (lane 5), V58A hGSTA1-1 (lane 6) and MW marker (lane 7) (B) Calibration curve (R^2 = 0.96) of MW marker proteins and the relative R_f of the hGSTA1-1 proteins (as indicated on the figure) with respect to the MW marker.

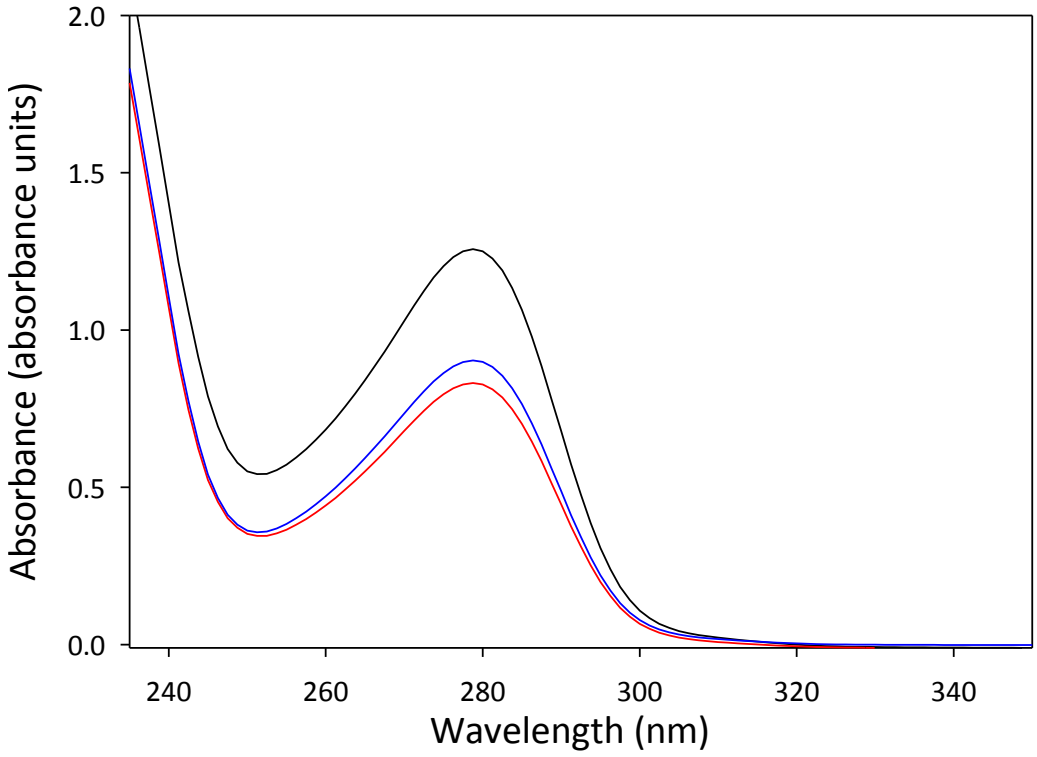


Figure 3-3. Absorbance spectra of the hGSTA1-1 proteins.
Spectra of pure wild-type hGSTA1-1 (black), V58A hGSTA1-1 (blue) and I75A hGSTA1-1 (red). Spectra were collected on samples containing the hGSTA1-1 proteins prepared in 20 mM sodium phosphate buffer, pH 7.4, containing 1 mM EDTA, 0.02% NaN_3, using a Jasco V-630 spectrophotometer.

3.5 Structural Characterisation

3.5.1 Secondary structure characterisation

Studies of the far-UV region (typically 250 – 190 nm) can be used to assess the overall secondary structural content of a protein quantitatively and have characteristic features based on the secondary structure adopted by the protein (Woody, 1995). The CD spectra of native forms of hGSTA1-1 proteins in the folded and unfolded conformational states are illustrated in Figure 3-4. It is evident that the mutations did not induce gross conformational changes in the secondary structure. Each protein exhibited minima at 222 nm and 208 nm and a peak at 190 nm, which is typical of proteins predominated by α-helices. This is consistent with the crystal structure of hGSTA1-1 [Figure 1-3, (Stenberg *et al*., 1992)].

3.5.2 Tertiary structure characterisation

3.5.2.1 Intrinsic fluorescence

Intrinsic tryptophan fluorescence is a useful spectroscopic technique used to measure the local environment of the indole side chain of the Trp residues in proteins. The number as well as the location of this residue within a protein, provides a useful means of probing for local or global conformational changes.

hGSTA1-1 has a lone tryptophan residue (Trp21) located in domain 1 that protrudes into domain 2, forming a 'lock-and-key' motif (Wallace *et al*., 2000) (Figure 1-3). Structurally related GST proteins with buried tryptophan residues display emission maximum for the native proteins at a wavelength of 335 nm (Kaplan *et al*., 1997; Hornby *et al*., 2000; Hornby *et al*., 2002; Luo *et al*., 2002). In the native state of wild-type and mutant GSTA1-1 proteins in this study, the indole ring of the Trp residues are slightly more buried than its dimeric GST counterparts, rendering an emission maximum at 325 nm (Figure 3-5). The absence of any shift in wavelength suggests that the mutations did not impact upon the environment of the tryptophan residues. The emission maximum for several proteins that contain tryptophan residues displays a red shift to 355 nm when they become denatured (Teale, 1960). This characteristic red shift in the maximum wavelength as well as a decrease in the intensity occured for wild-type and mutant hGSTA1-1 proteins in the presence of high concentrations of urea (Figure 3-5).

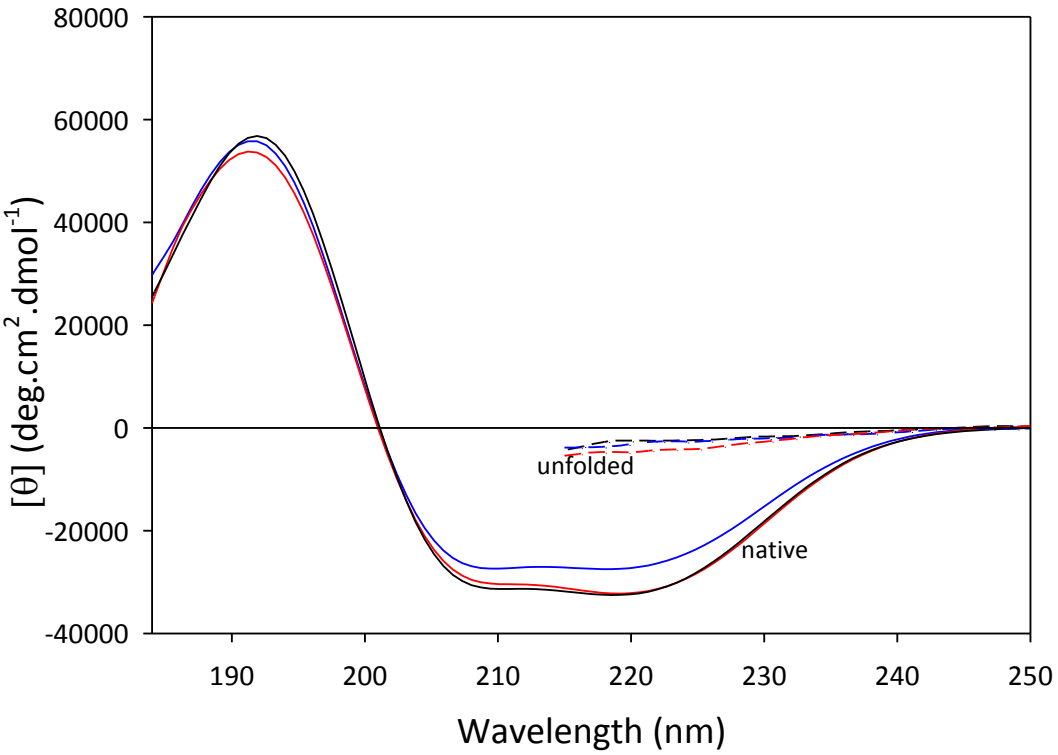


Figure 3-4. Far-UV circular dichroism spectra of hGSTA1-1 proteins.
Spectra shown for the native (solid lines) and unfolded forms (dashed lines) of wild-type (−), V58A (−) and I75A (−) hGSTA1-1 proteins. Spectra were collected using 2 μM protein in 2 mM sodium phosphate buffer, pH 7.4, containing 1 mM EDTA, to minimise the contribution of noise from the salt in the buffer for the native proteins. Spectra for unfolded protein were collected in the presence of high concentrations of denaturant (8.0 M urea).

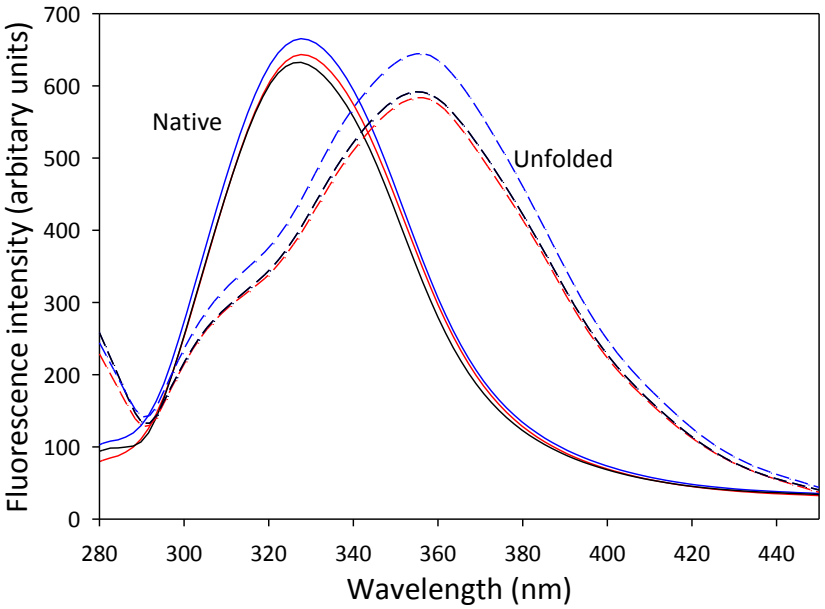


Figure 3-5. Fluorescence spectra of hGSTA1-1.
Fluorescence emission spectra shown for the native (solid lines) and unfolded (dashed lines) forms of wild-type (−) V58A (−) and I75A (−) hGSTA1-1. Spectra were collected using 2 μM protein in 20 mM sodium phosphate buffer, pH 7.4, 1 mM EDTA, 0.02% NaN_3 for native protein and in high concentration of denaturant (8.0 M urea). Samples were excited at 280 nm.

The indole ring of tryptophan (Lakowicz, 2006) is sensitive to changes in the environment as has been show by the changes in emission maxima for native and denatured proteins, hence making intrinsic fluorescence a good probe to monitor local tertiary structural changes for hGSTA1-1.

3.5.2.2 Near-UV circular dichroism

The near-UV CD spectra of proteins confer information about the packing and solvent exposure of the aromatic functional groups as well as possible contributions from disulfide bonds, or non-protein cofactors which might absorb in this spectral region (Kelly *et al*., 2005). The shape and magnitude of near UV CD spectra are dependant on the mobility and environment of the aromatic amino acid involved (Kelly *et al*., 2005). Near-UV CD was the second probe used to investigate the tertiary structure of hGSTA1-1 (Figure 3-6), however in contrast to fluorescence emission which predominates from tryptophan residues, the contributions made by each aromatic to the near UV CD signal are distinguishable. The hGSTA1-1 has 10 Phe, 10 Tyr and 1 Trp residues per subunit. The near UV CD spectra displayed bands at 255 nm, 265 nm, and 270 nm which are typical of those contributed by Phe residues (Kelly *et al*., 2005), and less prominent bands at 275 nm and 282 nm, characteristic of Tyr residues and large bands at 288 nm and 294 nm which are contributed by Trp residues. The strong overlap of near UV CD spectra suggests that there are no packing differences in the tertiary structure of hGSTA1-1 as a result of the mutations.

3.6 Functional characterisation

The hGSTA1-1 protein has both catalytic and ligandin function and these were tested on the hGSTA1-1 proteins to assess the impact of the mutations.

3.6.1 Ligandin (H-site) binding

The amphipathic ligand, ANS, binds the H-site of hGSTA1-1 with the hydrophobic anilino and naphthyl rings occupying the H-site and the polar sulfonate moiety located at the interface of the G- and H-sites (Dirr *et al*., 2005). The fluorescence emission spectrum ANS are characterised by the polarity of its environment. A change in the environment would result in a change of the spectral band in both wavelength and intensity (Slavik, 1982).

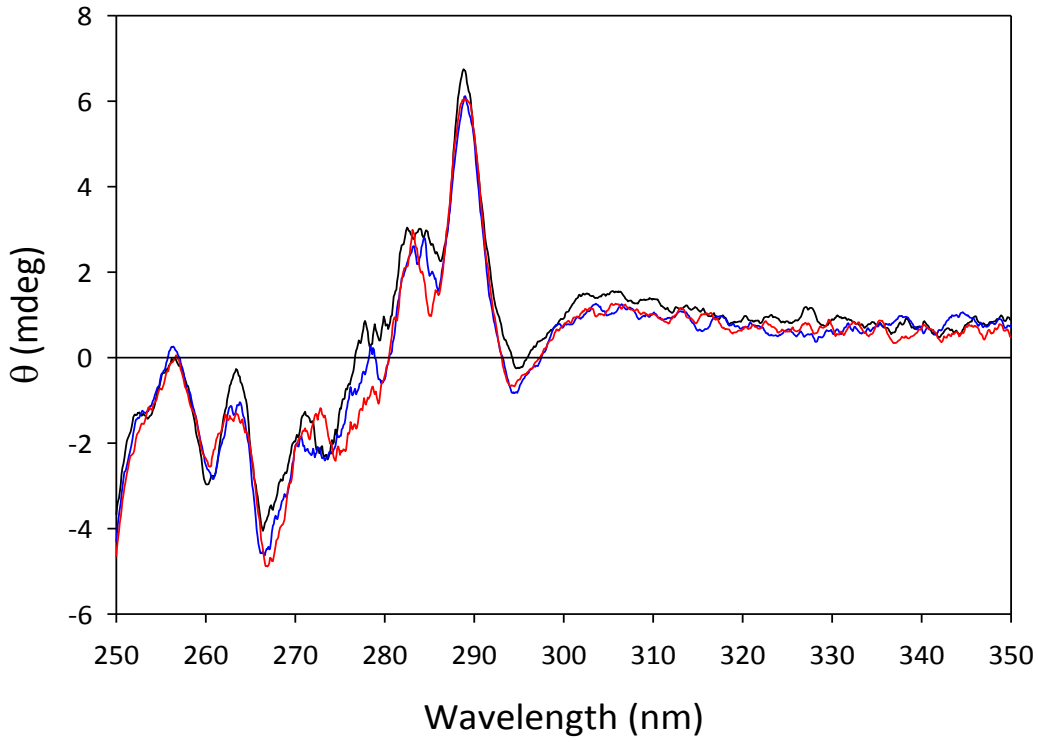


Figure 3-6. Near UV CD spectra of the hGSTA1-1 proteins.
Native CD spectra of wild-type (−), V58A (−) and I75A (−) hGSTA1-1 proteins. Spectra were recorded using 20 µM protein at 5ºC in 20 mM sodium phosphate buffer, pH 7.4, 1 mM EDTA, 0.02% NaN_3.

Thus the fluorescence properties of ANS were used to probe for any changes that may have resulted at the H-site of hGSTA1-1 as a result of the mutations.

The binding of ANS to the hGSTA1-1 proteins have emission spectra that peaked maximally at 480 nm with that of free ANS at 500 nm (Figure 3-7). This is in good agreement with previous studies [480 nm (Sluis-Cremer *et al*., 1996), 475 nm (Wallace *et al*., 1999) and 480 nm (Sayed *et al*., 2000)]. The absence of a shift in wavelength where the peak in observed and the absence of a significant intensity change, together suggest that the H-site remains unaffected as a result of the induced mutation.

3.6.2 Enzymatic activity

The specific activity of an enzyme is a rapid test that reveals whether any significant structural change has occurred at the enzyme's active site as a result of a mutation. The specific activity of hGSTA1-1 proteins was determined using the standard GSH-CDNB assay (Habig and Jakoby, 1974) (Figure 3-8). The activity obtained for the wild-type enzyme (Table 3-1) is comparable with those reported previously [39.4 $\mu mol.min^{-1}.mg^{-1}$ (Mosebi *et al*., 2003); 44.6 $\mu mol.min^{-1}.mg^{-1}$ (Nathaniel *et al*., 2003); 34.9 $\mu mol.min^{-1}.mg^{-1}$ (Gildenhuys *et al*., 2010a)].

The induced mutations did not extinguish enzymatic activity although the specific activity values were lower (Table 3-1) than the wild-type but this was by less than an order of magnitude. The mutations in the hydrophobic core of hGSTA1-1 therefore do not induce any structural conformation changes of the residues at the G-site (active site).

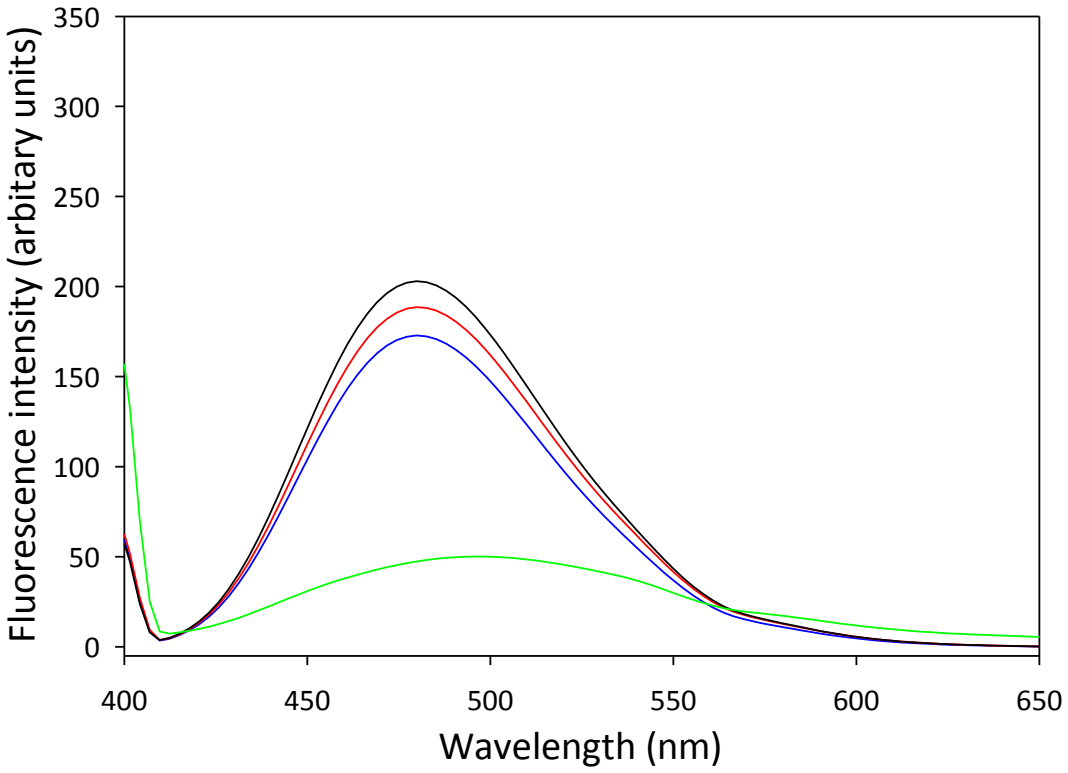


Figure 3-7. ANS binding to apo-hGSTA1-1
Emission spectra of ANS bound to the hGSTA1-1 proteins. Saturating amounts of ANS (200 µM) was added to (2 µM) wild-type (−) V58A (−) and I75A (−) protein. ANS was selectively excited at 390 nm and the emission spectrum read from 400 nm. The emission of free ANS (absence of protein; green) has been subtracted from the spectra showing ANS bound to protein. Spectra were recorded using 2 µM protein at 20 ºC in 20 mM sodium phosphate buffer, pH 7.4, 1 mM EDTA, 0.02% NaN_3.

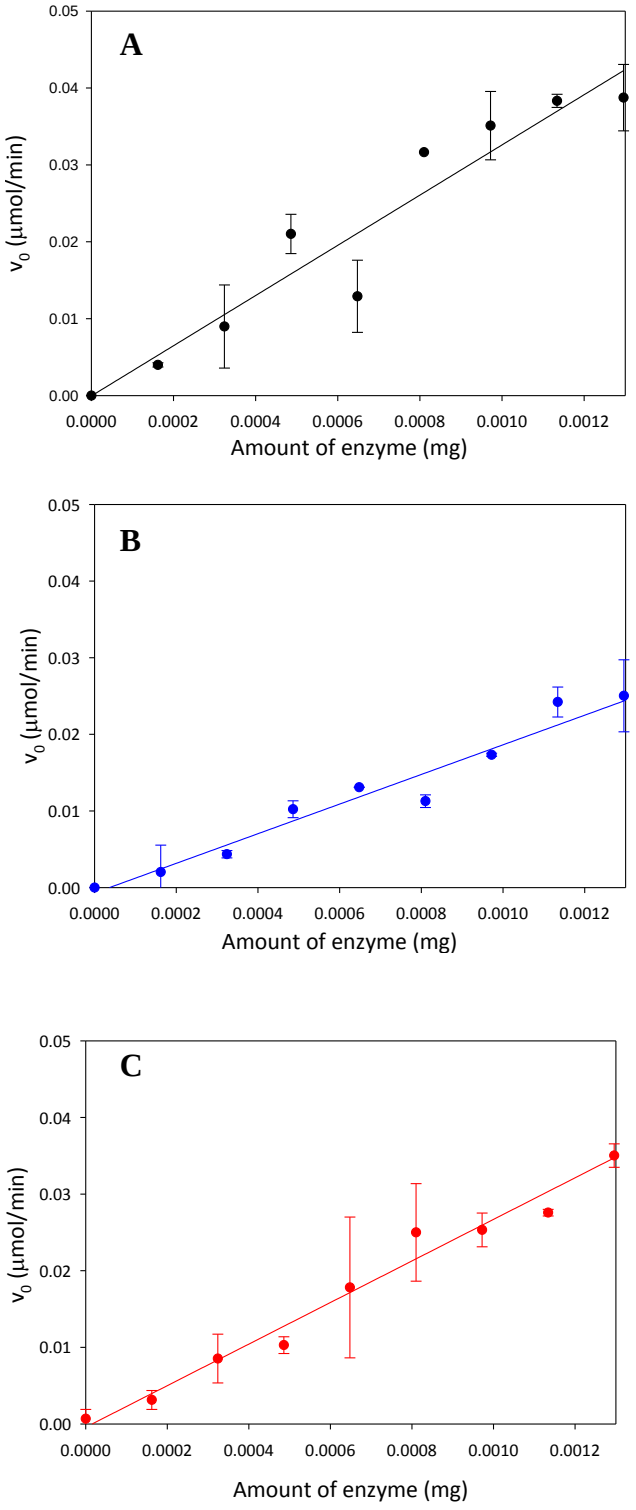


Figure 3-8. Specific activity of the hGSTA1-1 proteins
Specific activities of (A) wild-type (B) V58A and (C) I75A hGSTA1-1 proteins. The conjugation of CDNB by enzyme-bound GSH was monitored by measuring the absorbance at 340 nm of the 1-(*S*-glutathionyl)-2,4-dinitrobenzene product (ε_{340} = 9600 M^{-1} cm^{-1}). The absorbance contribution from the formation of the 1-(*S*-glutathionyl)-2,4-dinitrobenzene product in the absence of enzyme has been subtracted. The reaction was monitored in 0.1 M sodium phosphate buffer, pH 6.5, containing 1 mM EDTA and 0.02% NaN_3 in the presence of 1 mM GSH and CDNB.

Table 3-1: Specific activity values

The specific activity of each enzyme is the slope of the linear regression plots (Figure 3-8). The values are a mean from three replicate samples (n = 3) and the values in parentheses represent the standard deviation.

	wt	V58A	I75A
Specific activity (µmol/min/mg)	32.68 (±2.22)	19.38 (±1.09)	27.12 (±1.57)

3.7. Conformational stability

3.7.1 Thermal-induced unfolding

Thermal unfolding is widely used to determine protein stability. In thermal unfolding experiments, protein solution is heated at a constant rate, and changes in the protein conformation can be monitored by spectroscopy. Thermal unfolding was conducted on the hGSTA1-1 proteins using far-UV CD at 222 nm as a probe to follow the change in the α-helical content of the protein.

The proteins were found to unfold in a co-operative manner (Figure 3-9A). The $T_{1/2}$ values, the temperature at which 50% of the protein is folded/unfolded, of wild-type, V58A and the I75A hGSTA1-1 proteins were 59ºC, 57ºC and 52ºC respectively. The induced mutations therefore caused a distinct destabilisation. The turbidity recorded in CD experiments can be used as an indicator of aggregation as a protein is heat unfolded (Benjwal *et al.*, 2006). Thermal unfolding of the hGSTA1-1 proteins were shown to be irreversible as a result of aggregation shown by the increase in turbidity (Figure 3-9B) at the same temperature at which the respective proteins began to unfold. Heat-induced unfolding of the hGSTA1-1 proteins was associated with aggregation and hence irreversibility of unfolding. Therefore no thermodynamic analysis could be inferred by heat induced unfolding.

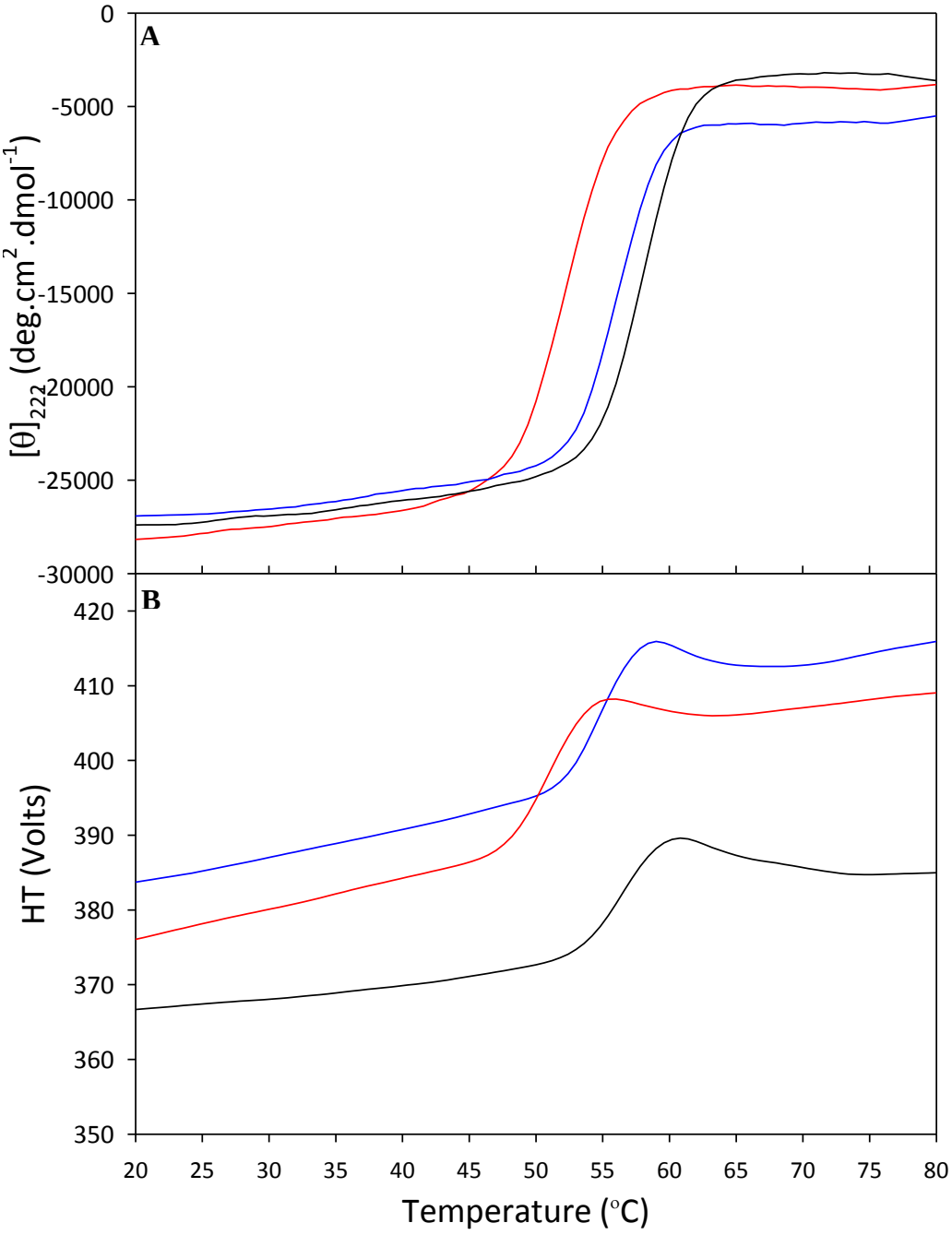


Figure 3-9. Thermal unfolding of hGSTA1-1.
(A) Heat unfolding curves of wild-type (−), V58A (−) and I75A (−) hGSTA1-1. (B) Turbidity (dynode voltage) plot indicating aggregation followed by precipitation of the aggregates indicated by the reduction in voltage at the temperature where the proteins are largely unfolded.

3.7.2 Recovery of urea-induced unfolding

Equilibrium unfolding transitions can only be analysed in terms of their thermodynamic parameters if it has been established that the unfolding reaction is reversible and that the native fold can be recovered (Pace, 1986). The recovery of unfolded hGSTA1-1 was established using far-UV CD and intrinsic tryptophan fluorescence as secondary and tertiary probes, respectively. The proteins were unfolded in the presence of 8.0 M urea and allowed to refold in storage buffer. The ellipticity at 222 nm was chosen as the indicator of recovery of the secondary as this is a signature of predominantly helical proteins. The maximum fluorescence emission at 325 nm (where native hGSTA1-1 maximally fluoresces) was chosen as the indicator of tertiary structure recovery. Refolded hGSTA1-1 proteins displayed at least 95% recovery (Table 3-2). Thus the equilibrium unfolding transitions of could be analysed to determine the thermodynamic parameters which would indicate the conformational stability of the hGSTA1-1 proteins.

3.7.3 Urea-induced equilibrium unfolding

Denaturation curves using chemical denaturants are a convenient method for estimating the conformational stability of a protein. Urea was used as the chemical denaturant and the conformational stability was determined by setting up a range (0 M - 8M urea) of unfolding reactions which were allowed to reach equilibrium (1 hour). The resulting urea-induced equilibrium-unfolding reactions were monitored using the structural probes (far-UV CD and intrinsic tryptophan fluorescence). The probes detect the structure of the predominating species present at equilibrium at the respective urea concentrations (Pace, 1986). The ellipticities at 222 nm and fluorescence emission at 325 nm for the secondary and tertiary structural probes, respectively, were plotted as a function of urea.

3.7.3.1 Equilibrium unfolding of wild-type and V58A hGSTA1-1

It is clear that the unfolding curves of the wild-type (Figure 3-10A) and V58A hGSTA1-1 (Figure 3-11A) proteins gave monophasic transition curves for both unrelated spectroscopic probes (far-UV CD and intrinsic tryptophan fluorescence). The unfolding curves for the wild-type and V58A hGSTA1-1 protein also superimposed well. Thus the concerted loss of the secondary and tertiary structures at

Table 3-2. Extent of recovery of the native forms of the hGSTA1-1 proteins.

Recoveries expressed as a percentage and determined by far-UV CD (E_{222}) and fluorescence emission spectroscopy (F_{325}).

	% recovery (E_{222})	% recovery (F_{325})
Wild-type hGSTA1-1	96	98
V58A hGSTA1-1	98	99
I75A hGSTA1-1	95	96

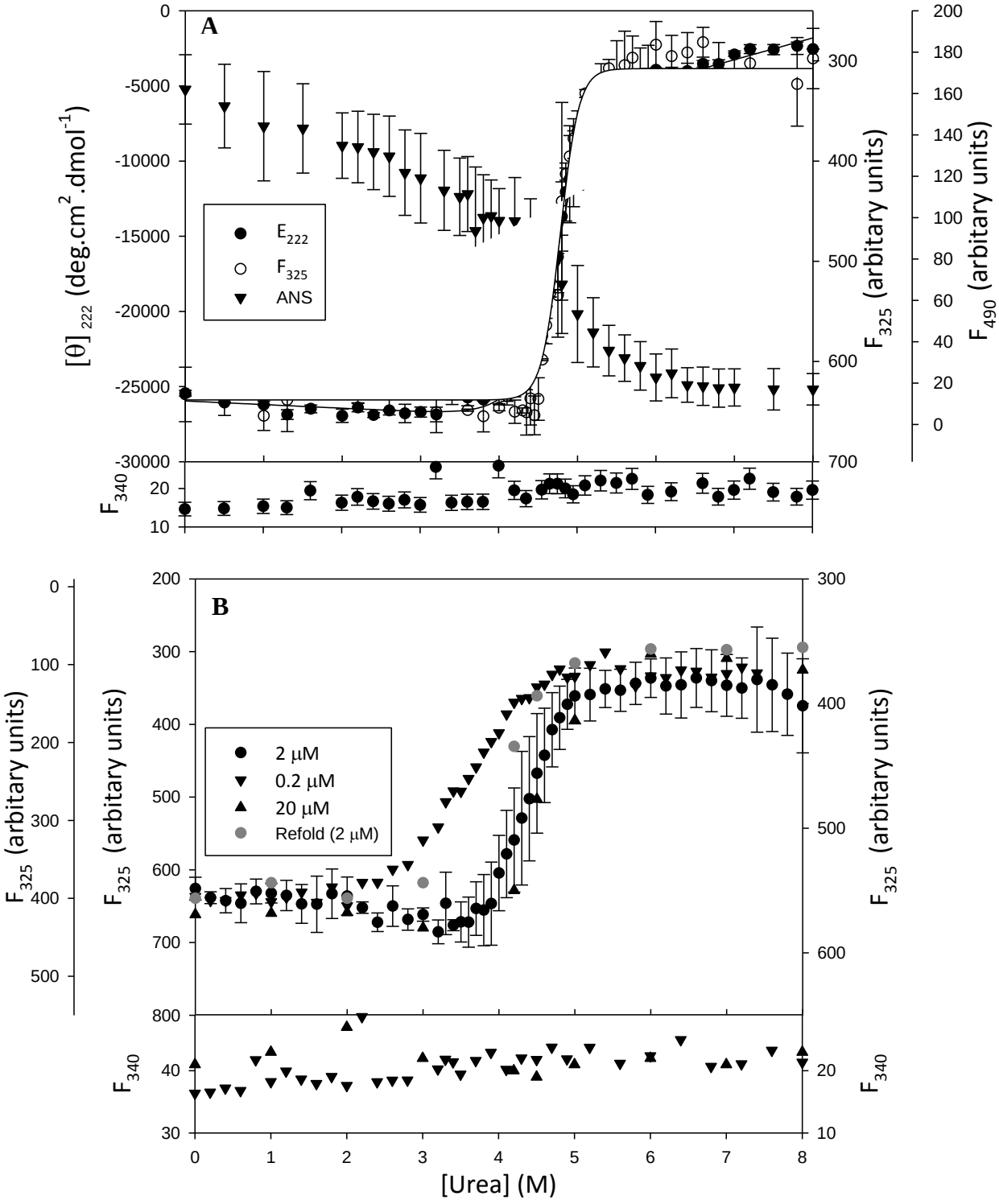


Figure 3-10. Urea-induced equilibrium unfolding of wt hGSTA1-1
(A) Unfolding curves for 2 μM wt hGSTA1-1 using far UV CD, tryptophan fluorescence and ANS fluorescence as probes. The solid lines through the datum points represent the fit of the two-state model. (B) Unfolding curves as a function of protein concentration using tryptophan fluorescence as a probe. The test for unfolding reversibility is shown by symbols coloured in gray. Light scattering at 340 nm was used to detect for aggregates (lower panels).

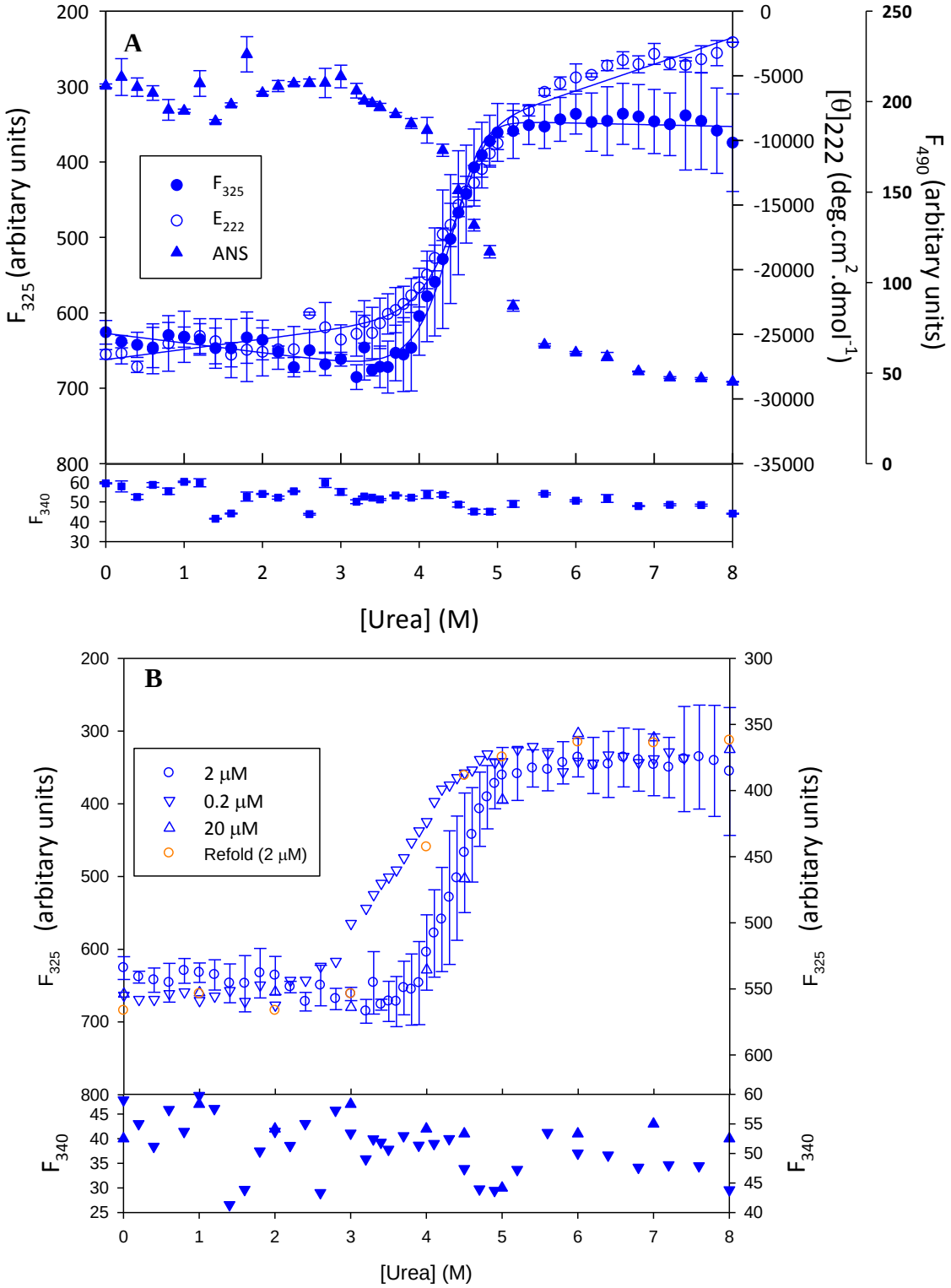


Figure 3-11. Urea-induced equilibrium unfolding of V58A hGSTA1-1
(A) Unfolding curves for 2 μM V58A hGSTA1-1 using far UV CD, tryptophan fluorescence and ANS fluorescence as probes. The solid lines through the datum points represent the fit of the two-state model. (B) Unfolding curves as a function of protein concentration using tryptophan fluorescence as a probe. The test for unfolding reversibility is shown by symbols coloured in orange. Light scattering at 340 nm was used to detect for aggregates (lower panels).

the domain interface (Trp21), as seen by the coincident unfolding curves, is characteristic of a highly co-operative two-state folding protein (Sancho 2013). Two additional probes were used to confirm the co-operative behaviour of unfolding by performing the aforementioned equilibrium unfolding studies in the presence of ANS as well as carrying out the experiments at varying protein concentrations (Figure 3-10A and Figure 3-11A).

The binding of the hydrophobic ligand, ANS, is widely used to monitor the conformational changes induced in proteins during the unfolding process (Semistonov *et al*., 1991). ANS is used to probe for the presence of intermediates during protein unfolding events (Semistonov *et al*., 1991). It has been reported that ANS, when bound to the molten globule state, shows higher emission intensities than when bound to protein in either the folded or partially unfolded state (Ptitsyn, 1995). In the case of hGSTA1-1, the native state bound ANS (Figure 3-7) due to its ligandin function (Sluis-Cremer *et al*., 1996). However, in the unfolding experiments presented here, there was no enhancement in the ANS fluorescence intensity with increasing denaturant concentration (Figure 3-10A and Figure 3-11A). A gradual decrease in the fluorescence was witnessed between 0 M and 4.2 M urea, followed by a sharper decrease between 4.3 M and 6.0 M urea, coinciding with the transition region of the far-UV CD and fluorescence data. No further change in the intensity occurred in the unfolded region (6.0 M to 8.0 M urea). The light scattering data (Figure 3-10) indicated no significant aggregation, hence the binding of ANS to protein aggregates could be ruled out.

The hGSTA1-1 protein is a dimer in the native state (Figure 1-3) and a dimeric protein can unfold via two pathways if an intermediate (I) is present:

$$N_2 \rightleftharpoons I_2 \rightleftharpoons 2U \text{ (dimeric intermediate)}$$

or

$$N_2 \rightleftharpoons 2I \rightleftharpoons 2U \text{ (monomeric intermediate)}$$

The existence of a dimeric (I_2) or monomeric (I) intermediate can be probed by performing protein concentration dependence studies during equilibrium unfolding.

The choice of protein concentrations should differ by 10 fold to witness a significant change (Rumfeldt *et al*., 2008), therefore in this study, concentrations of 0.2 µM, 2 µM and 20 µM were used. The unfolding curves of the wild-type and V58A hGSTA1-1 proteins at all protein concentrations were sigmoidal with monophasic transitions ((Figure 3-10B and Figure 3-11B). An expected difference of the C_m (concentration where 50% of the protein population is unfolded) was witnessed where it is shifted to slightly lower denaturant concentration (3.5 M) in the case of 0.2 µM protein concentration and higher (4.6 M) for the 20 µM protein concentration.

3.7.3.2 Equilibrium unfolding of I75A hGSTA1-1

The unfolding curves of the I75A hGSTA1-1 were markedly different to that of the wild-type protein. The unfolding curves (Figure 3-12) displayed both biphasic and non-superimposable transitions suggesting the formation of an intermediate. The probes for the tertiary structure at the domain interface (tryptophan fluorescence) and the secondary structure (far-UV CD) deviate between denaturant concentrations of 2.6 M and 6.5 M urea (Figure 3-12).

The non co-incident biphasic transitions differed to that of the wild-type in that they are broader and less steep for both spectroscopic probes suggesting that the unfolding co-operativity was lost (Sancho, 2013). The lack of change in signal for both probes at denaturant concentrations between 3.2 M and 4.2 M suggests the formation of a stable intermediate.

As with the wild-type and the V58A hGSTA1-1 proteins, the use of additional probes, ANS and protein concentration dependence during unfolding would determine if the intermediate formed is dimeric or monomeric. The unfolding curves for all the protein concentrations were biphasic and with no divergence at the pre- or post-transition regions. The absence of protein concentration dependence suggests that there is no change in the oligomeric state of the hGSTA1-1 protein and dissociation into monomers must occur after the second transition region.

The ANS fluorescence intensity decreased gradually in the pre-transition region with a steeper decrease in the transition region similar to that of the wild-type protein.

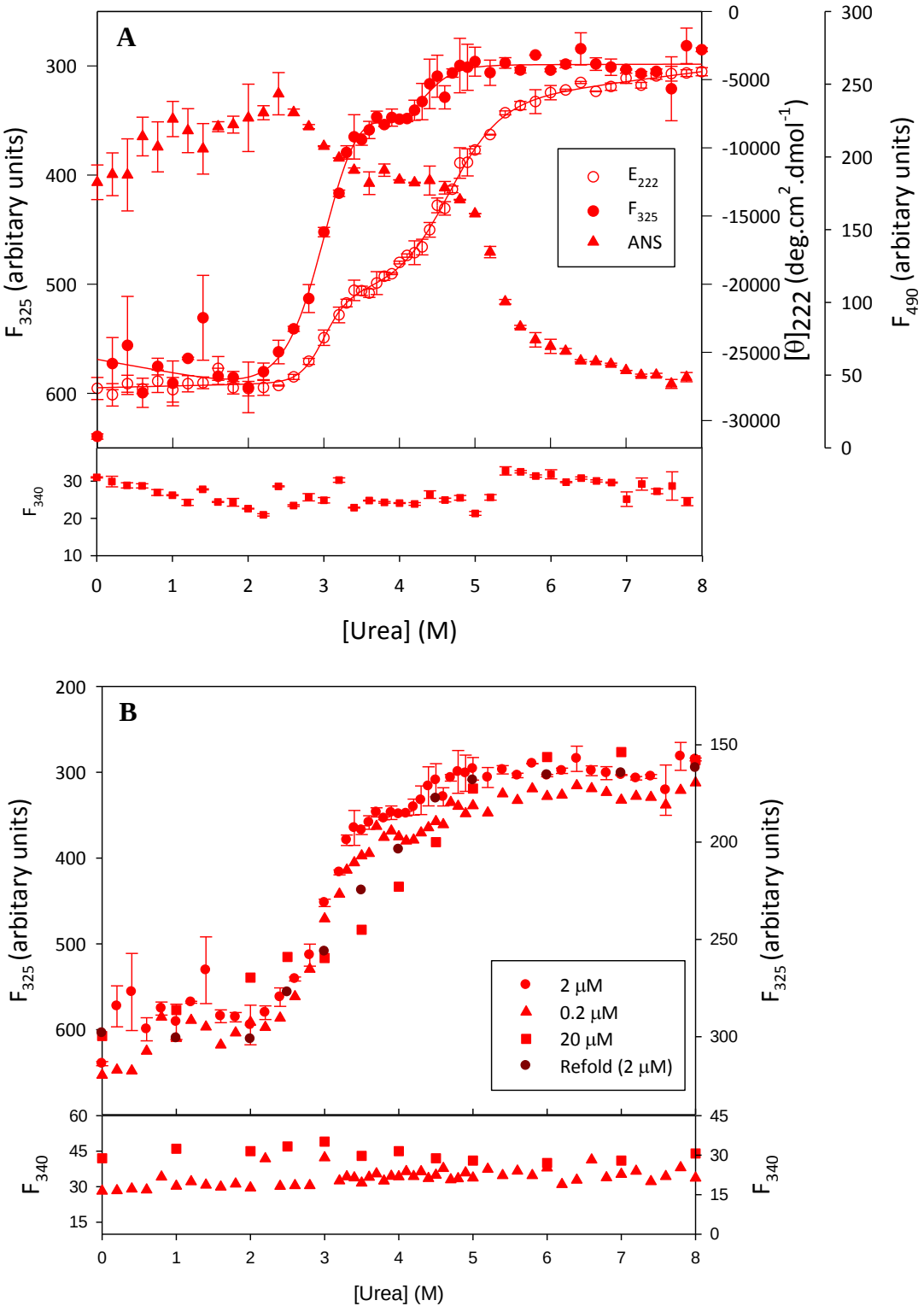


Figure 3-12. Urea-induced equilibrium unfolding of the I75A hGSTA1-1
(A) Unfolding curves for 2 µM I75A hGSTA1-1 using far UV CD, tryptophan fluorescence and ANS fluorescence as probes. The solid lines through the datum points represent the fit of the three-state model. (B) Unfolding curves as a function of protein concentration using tryptophan fluorescence as a probe. The test for unfolding reversibility is shown by symbols coloured in gray. Light scattering at 340 nm was used to detect aggregates (lower panels).

However, no change in fluorescence intensity was seen between denaturation concentrations of 3.2 M and 4.2 M, suggesting the formation of an intermediate.

3.7.4 Properties of the unfolding intermediate formed by the I75A mutation

The I75A mutation causes a formation of an intermediate that populates at denaturant concentrations between 3.2 M and 4.2 M urea. Its secondary and tertiary structural properties as determined by far-UV CD and intrinsic fluorescence are shown in Figure 3-13. The fractional helical content of the intermediate can be determined by calculating the proportion to the determined amount of the native state (Marqusee and Baldwin, 1987). Therefore, the fraction of helical content is 65 % relative to the native state (Figure 3-13A). There is a hypsochromic shift in ellipticity minima, 222 nm to 218 nm, which is indicative of a structural shift from an α-helical structure toward a random coil (Woody, 1995, Corrêa and Ramos 2009).

The fluorescence emission at 325 nm (λ_{max} of native state) of the intermediate state is close to that of the unfolded form of the protein (Figure 3-13B). The unfolded form has a λ_{max} of 346nm which is similar to that of exposed tryptophan residues (Lacowicz, 2006). The fluorescent signal predominates as a result of the emission of Trp21, which is located at the domain interface, hence the large quenching of the fluorescence emission can be attributed to solvent exposure of the domain interface.

A plot displaying maximum emission wavelength at each denaturant concentration is useful in displaying fluorescence data independent of emission intensity of the fluorophore and any shift in wavelength of the maximum emission can provide more information on the nature of the surrounding environment of the fluorophore (Trp21) used. All the hGSTA1-1 proteins used in this study displayed a bathochromic shift which is consistent for proteins undergoing a transition from a native state to unfolded state (Figure 3-13C). As this fluorescence probe reports around the local environment of the domain interface, the bathochromic shift suggests that the domain interface of the protein in the intermediate state has become solvent exposed.

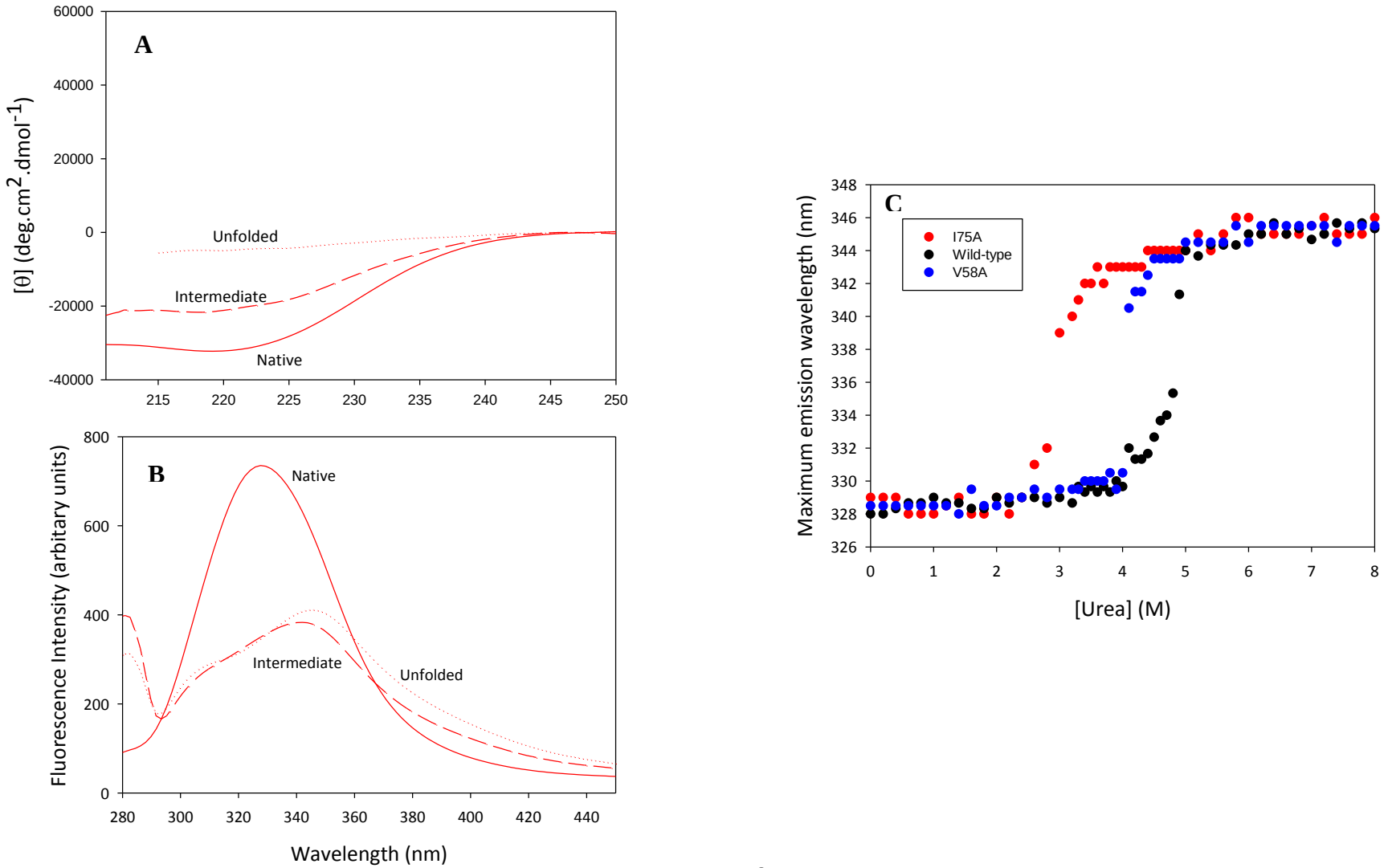


Figure 3-13. Spectra of the native, intermediate and unfolded species of I75A hGSTA1-1

(A) Far-UV CD and (B) fluorescence emission spectra of the native (0 M urea, solid line), intermediate (3.7 M urea, dashed line) and unfolded (8.0 M urea, dotted line) forms of I75A hGSTA1-1. (C) Maximum emission wavelength of the hGSTA1-1 proteins at each urea concentration.

3.7.5 Data fitting and thermodynamic parameters

The data was fitted to the linear extrapolation method (Pace, 1986) as described in section 2.2.11 from which the thermodynamic properties of unfolding were determined Although the equilibrium unfolding mechanism is described as three-state (section 1.4.3.1) due to the unfolding of α9, a two-state model was used to fit the wild-type and V58A hGSTA1-1 unfolding data as there is no evidence to suggest that a stable intermediate is formed and α9 does not significantly contribute toward the stability of hGSTA1-1 (Mosebi *et al*., 2003; Kuhnert *et al*., 2005).

The conformational stability parameters are tabulated in Table 3-3. The $\Delta G(H_2O)$ and *m*-values determined for the wild-type hGSTA1-1 protein correlate well with that reported previously [25.9 kcal.mol^{-1}, (Wallace *et al*., 1998)]. A slight decrease (~7%) in the $\Delta G(H_2O)$ was measured for the V58A hGSTA1-1 protein indicating a marginal destabilisation as a result of the V58A mutation. It is the slope in the transition region in the unfolding curve. The *m*-value, was also decreased marginally (~5%). The *m*-value is defined by dependence of the free energy of unfolding on denaturant concentration (Pace, 1986), i.e the slope of the transition region in the unfolding curve. The *m*-value therefore indicates the degree of surface area exposed upon unfolding (Myers *et al*., 1995) A lower *m*-value is usually indicative of an intermediate forming (Soulages, 1998), however, none of the data obtained with the other probes seem to suggest that an intermediate formed.

A two-state model was precluded for the unfolding data for I75A hGSTA1-1 as the two spectroscopic probes do not coincide with one another and was not sigmoidal. These observations suggest the formation of an intermediate (Sancho 2013, Tsytlonok and Itzhaki, 2013), thus a three-state model was used to fit the unfolding data for I75A hGSTA1-1 and the thermodynamic parameters of which are tabulated in Table 3-4. The $\Delta G(H_2O)$ values for the N – I transition are higher than the I – U transition indicating that the intermediate species is energetically less stable than the native state as less energy is required to unfold this species. Although the total $\Delta G(H_2O)$ is the sum of $\Delta G_{N\text{-}I}(H_2O)$ and $\Delta G_{I\text{-}U}(H_2O)$, the only relevant form of comparison is the $\Delta G_{N\text{-}I}$ (H_2O) (Privalov *et al*.,1981; Sancho *et al*., 2002). The lower *m*-values indicate a loss

Table 3-3. Conformational stability parameters
The values were obtained from fitting the equilibrium unfolding data to a two-state model ($N_2 \leftrightarrow 2U$).

Parameter	Far UV CD probe		Intrinsic fluorescence probe	
	wild-type	***V58A***	***wild-type***	***V58A***
$\Delta G_{N\text{-}U}(H_2O)$ (kcal/mol)	25.1 (±0.18)	23.2 (±0.22)	24.8 (±0.18)	22.8 (±0.26)
$m_{N\text{-}U}$ (kcal/mol/M urea)	4.3 (±0.21)	4.1 (±0.19)	4.3 (±0.18)	4.1 (±0.22)
C_m (M)	4.8	4.4	4.7	4.3

Table 3-4. Conformational stability parameters for I75A hGSTA1-1
The values were obtained from fitting the equilibrium unfolding data to a three-state model ($N_2 \leftrightarrow I_2 \leftrightarrow 2U$).

Three state model (I75A hGSTA1-1)

Probe	$\Delta G_{N\text{-}I}(H_2O)$ (kcal/mol)	$m_{N\text{-}I}$ (kcal/mol/M urea)	C_m (M)	$\Delta G_{I\text{-}U}$ (H_2O) (kcal/mol)	$m_{I\text{-}U}$ (kcal/mol/M urea)	C_m (M)
Far-UV CD	14.8 (±0.52)	2.1 (±0.11)	2.8	12.4 (±0.07)	2.9 (±0.05)	4.6
Intrinsic fluorescence	10.7 (±2.2)	2.6 (±1.2)	2.8	12.1 (±0.18)	2.6 (±0.2)	4.6

of co-operativity in the unfolding and thus the protein no longer follows a single step ($N \rightleftharpoons U$) of unfolding.

3.8 Pulse Proteolysis

A spectroscopic-independent probe was also employed to determine the relative stabilities of the proteins based on their susceptibility to proteolysis. Samples were prepared as for conformational stability studies involving urea followed by proteolytic digestion by thermolysin (section 2.2.12). In pulse proteolysis, as the name suggests, the reaction is allowed to take place for a brief fixed period of time. This is useful in that the unfolded species of protein is cleaved at a higher rate than the folded species thus allow the relative stabilities to be determined based on the fractions of folded versus unfolded protein at a given denaturant concentration (Park and Marqusee, 2005).

Proteolysis by thermolysin of the hGSTA1-1 proteins were analysed by SDS-PAGE (Figure 3-14). Densitometric analyses of the bands corresponding to the GST protein on the SDS-PAGE gels of the proteolytic digests (Figure 3-15) revealed a similar profile of unfolding as obtained with the spectral probes. Proteolytic digestion of the wild-type hGSTA1-1 occurs between urea concentrations of 3.2 M and 4 M with the midpoint of unfolding (C_m) occurring at ~ 4.1 M (Figure 3-15). Digestion of the V58A and the I75A hGSTA1-1 proteins occurs at urea concentrations of 2.4 M and 1.6 M, respectively, and midpoints of unfolding (C_m) occurring at ~ 3 M and 2 M, respectively (Figure 3-15). The digested products are visible on the gel at urea concentrations of 5 M and higher (Figure 3-14) and increases with increasing urea concentration as the population of unfolded protein is the predominating species at these concentrations. The V58A and the I75A hGSTA1-1 is more susceptible to proteolytic digestion as compared to the wild-type hGSTA1-1 proteins as judged by the differences in C_m values and is thus less stable.

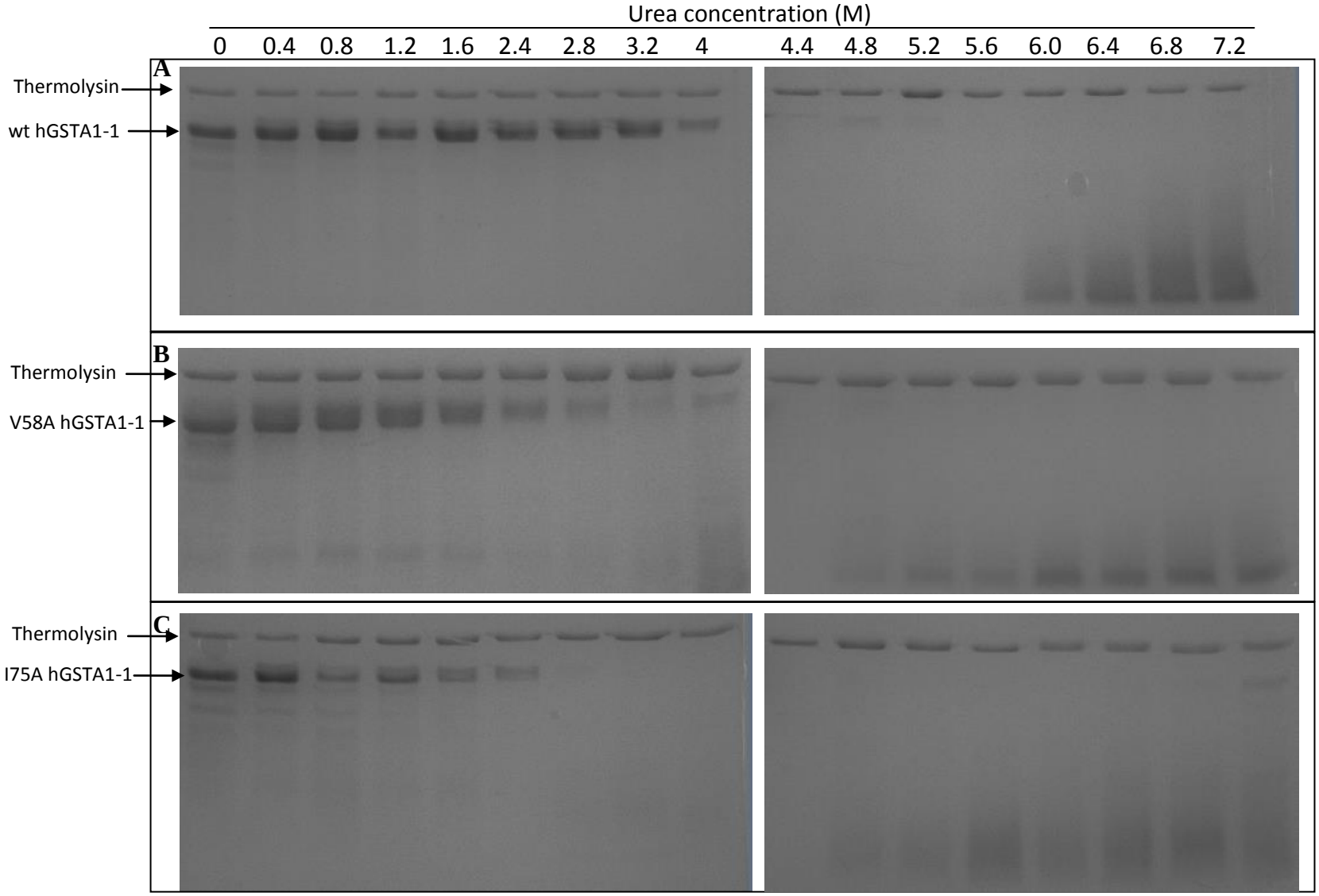


Figure 3-14. SDS-PAGE analysis of pulse proteolysis

Gel images (Coomasie stained) of the proteolytic cleavage by thermolysin of (A) wild-type (B) I75A and (C) V58A hGSTA1-1 at different urea concentrations (indicated on each lane) after equilibrium was reached. The upper bands in each sample lane represent thermolysin while the lower band represents the hGSTA1-1 protein. Digests of urea concentrations between 4.4 M and 7.2 M were electrophoresed on a separate gel. Digested fragments are visible at the bottom of the gels and are clearly visible at high (>5 M) urea concentrations.

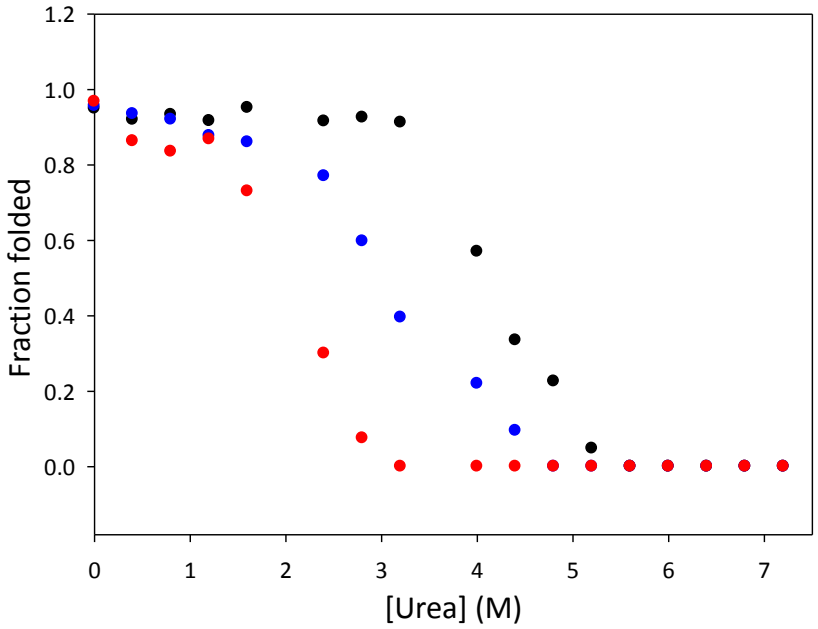


Figure 3-15. Densitometric analysis of pulse proteolysis
Optical density analysis of thermolysin digestion of the wild-type (•), V58A (•) and I75A (•) hGSTA1-1 proteins. The relative amounts of the hGSTA1-1 protein band on the gel (Figure 3-14) are plotted as a function of urea concentration.

3.10 Protein crystallography

Crystallisation trials were carried out for the V58A and I75A hGSTA1-1 mutant proteins. The initial conditions used for the crystallisation trials were the same as those previously reported (Achilonu *et al*., 2010) and the variations included protein and precipitant concentrations. Conditions that were optimum for the V58A hGSTA1-1 protein were: 7 mg/ml protein, 19% (w/v) PEG 3350, in 1 M Tris-HCl, pH 7.5, 10 mM DTT. Needle-like crystals were obtained (Figure 3-16A). Crystals of the apo-structure are required to draw a meaningful comparison between the mutant and wild-type protein as a ligand bound protein may mask the impact of the mutation. In the case of the I75A hGSTA1-1 mutant protein, relatively small crystals were obtained (Figure 3-16B). Crystals were obtained under similar conditions as those used by the V58A hGSTA1-1 protein, with only the PEG size and concentration being altered (section 2.2.13).

The structures of the V58A and I75A hGSTA1-1 mutants in the unliganded form were determined at 1.88 Å and 1.86 Å resolution, respectively. The electron densities are well defined for residues 2- 209 for both subunits in each of the mutant structures. The electron densities at the sites for mutation fit well to that of an alanine residue (Figure 3-17). The quality of the fit of the model to the diffraction data is determined by the R value. A value of at least 0.06 is desirable for the difference between the final R and the final R_{free}, which indicates a good quality fit (Brünger, 1992). The crystal structures for both mutants yielded a difference of 0.06 between the final R and final R_{free} values (Table 3-4). The other data-collection and refinement statistics are shown in Table 3-4.

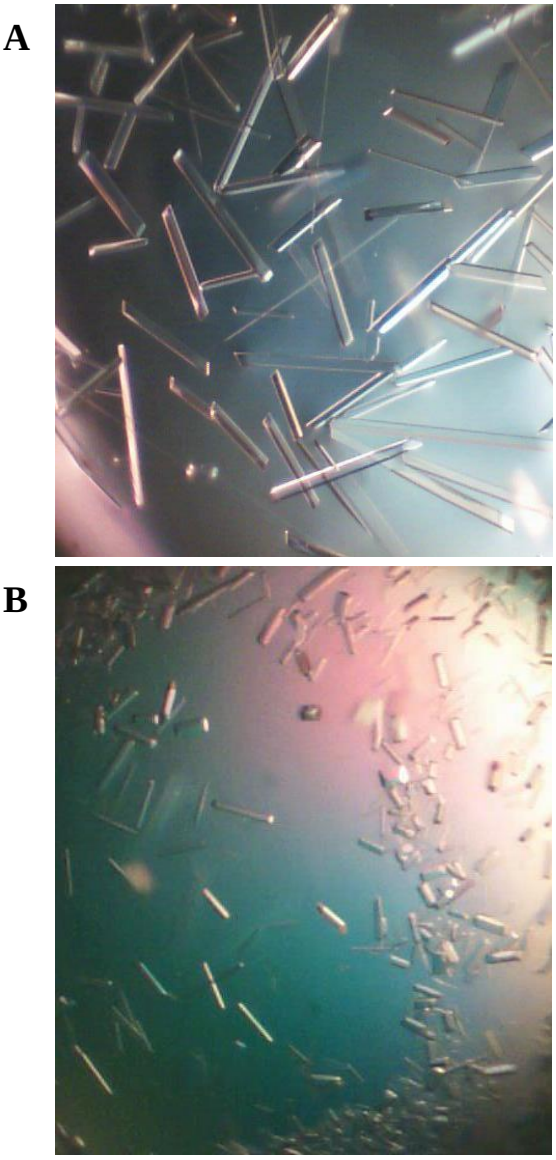


Figure 3-16. Crystal images of the mutant hGSTA1-1 proteins

(A) Crystals of V58A hGSTA1-1 obtained in 1 M Tris-HCl, pH 7.5, 10 mM DTT, 19% PEG 3350 using 7 mg/ml protein. (B) Crystals of I75A hGSTA1-1 obtained in 1 M Tris-HCl, pH 7.5, 10 mM DTT, 18% PEG 4000 using 7 mg/ml protein.

Table 3-5: Crystallographic X-ray data collection and refinement statistics of V58A and I75A hGSTA1-1.

Parameter	**V58A**	**I75A**
Buffer/salt	0.1 M Tris HCl pH7.5	0.1 M Tris HCl, pH7.5
PEG (% w/v)	18% (M_r 3350)	17% (M_r 4000)
PDB ID	3ZFL	3ZFB
Data Collection:	CCD	CCD
Wavelength (Å)	1.5418	1.5418
Space group	*C*2 (*C*121)	*C*2(*C*121)
Unit cell dimensions (Å)	a =99.89	a=99.87
	b =93.28	b=91.68
	c =51.57	c=51.25
Asymmetric unit content	Dimer	Dimer
Resolution range (Å)	51.30 – 1.88	51.18 – 1.86
No of reflection(obs)	205610	613762
No of unique reflections	42694	34481
Completeness (%)	98.8	99.4
I/σ (I)	4.06	4.36
Refinement:		
Resolution (Å)	49.87 – 1.88	51.18 – 1.86
R_{cryst}, R_{free}	0.221 (0.289)	0.228 (0.285)
Reflections	35960	36399
Protein atoms	3340	3362
Solvent molecules	350	362
Rmsd bond lengths[a] (Å)	0.016	0.016
Rmsd angles[a] (deg)	1.78	1.80
$\langle B \rangle$ protein (Å^2)	26.57	23.21
Matthews co-efficient (A^3.Da^{-1})	2.31	2.25
Ramachandran plot:		
Most favoured (%)	97.6	97.3
Generously allowed (%)	2.4	2.7
Disallowed (%)	0.05	0.05

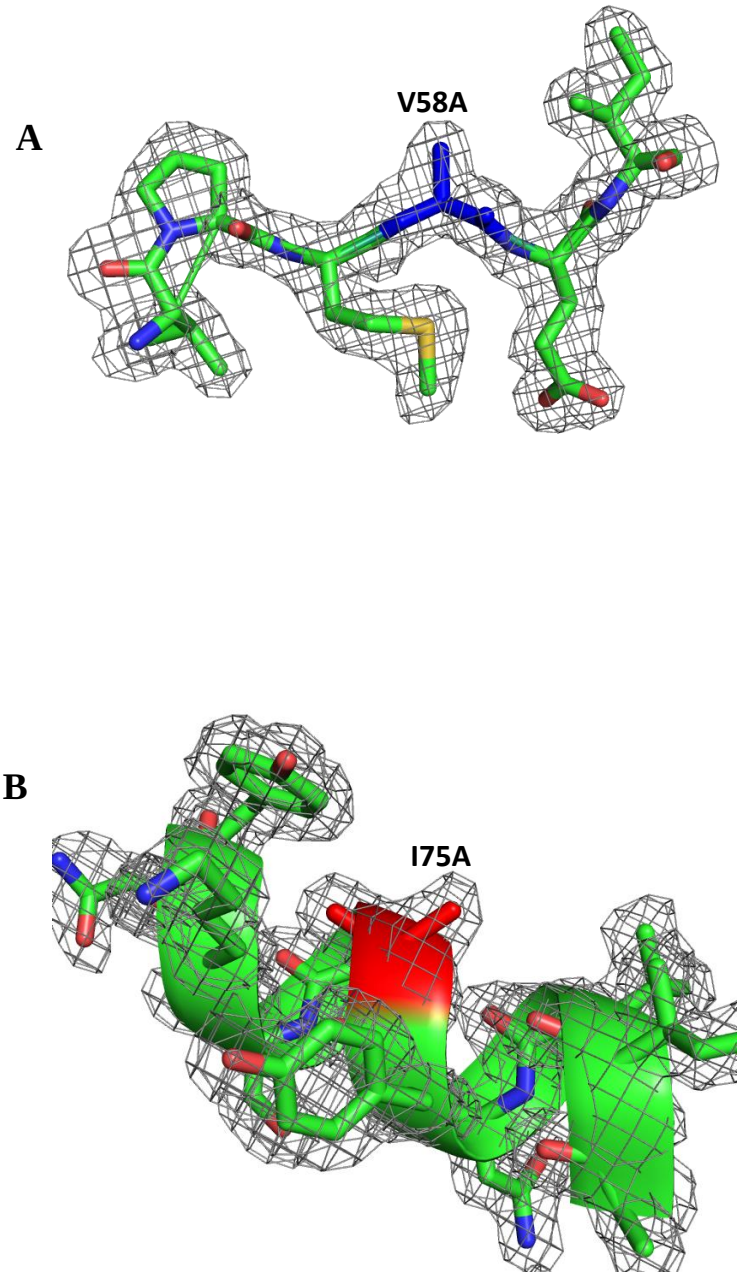


Figure 3-17. Fit of crystal models into electron density map
Electron density maps shown at the point of mutations, (A) V58A and (B) I75A hGSTA1-1. The mutated residues are highlighted in blue (V58A) and red (I75A) and fit well into the density of alanine residue. Images rendered using PyMOL^{TM}, DeLano Scientific, 2006; PDB codes: 3ZFL, and 3ZFB.

The Ramachandran plots of the solved structures (Figure 3-18) are considered to be satisfactory as only two non-glycine residues were found to be outside the allowed regions for each structure. In both crystal structures, the outliers are located on loop regions of the protein. The Gln67 residue is seen as an outlier in other hGSTA1-1 models (1K3O: Le Trong *et al*., 2002; 3KTL: Achilonu *et al*., 2010) and could be considered a feature of hGSTA1-1 crystals rather than an error in the model.

A parameter that indicates the extent of deviation of a crystal structure from a reference model is the root mean square deviation (r.m.s.d) value. (Maiti *et al*., 2004). A structural alignment of the polypeptide backbone of the mutant proteins against wild-type protein (Figure 3-19) provided r.m.s.d values of 0.58 Å and 0.61 Å for the V58A and I75A mutants, respectively. The crystal data concurs with the spectroscopic data (section 3.5) obtained for the native structures in solution. Thus the induced mutations caused no major structural alterations to the native protein.

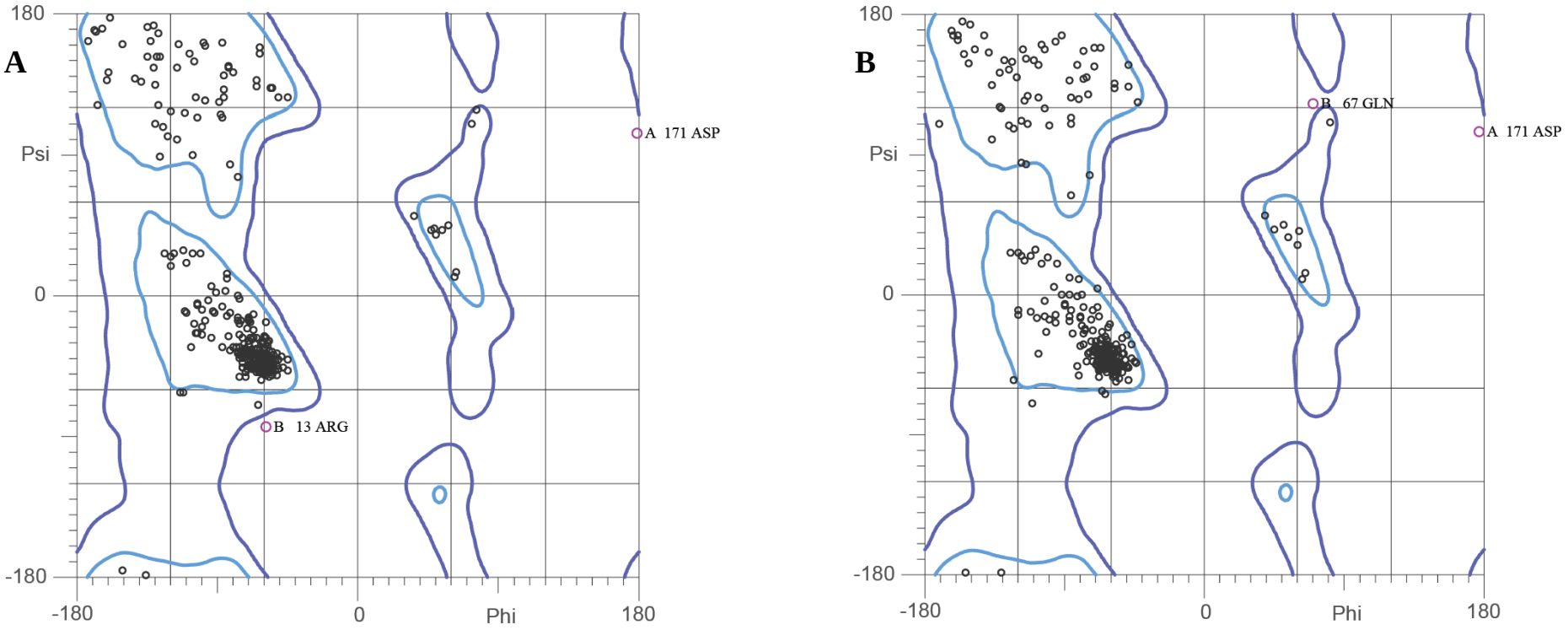


Figure 3-18. Ramachandran plots of the mutant hGSTA1-1 proteins
Ramachandran plots of the (A) V58A and (B) I75A hGSTA1-1 protein. Two residues are found outside the allowed regions in each structure represented by the purple open symbols.

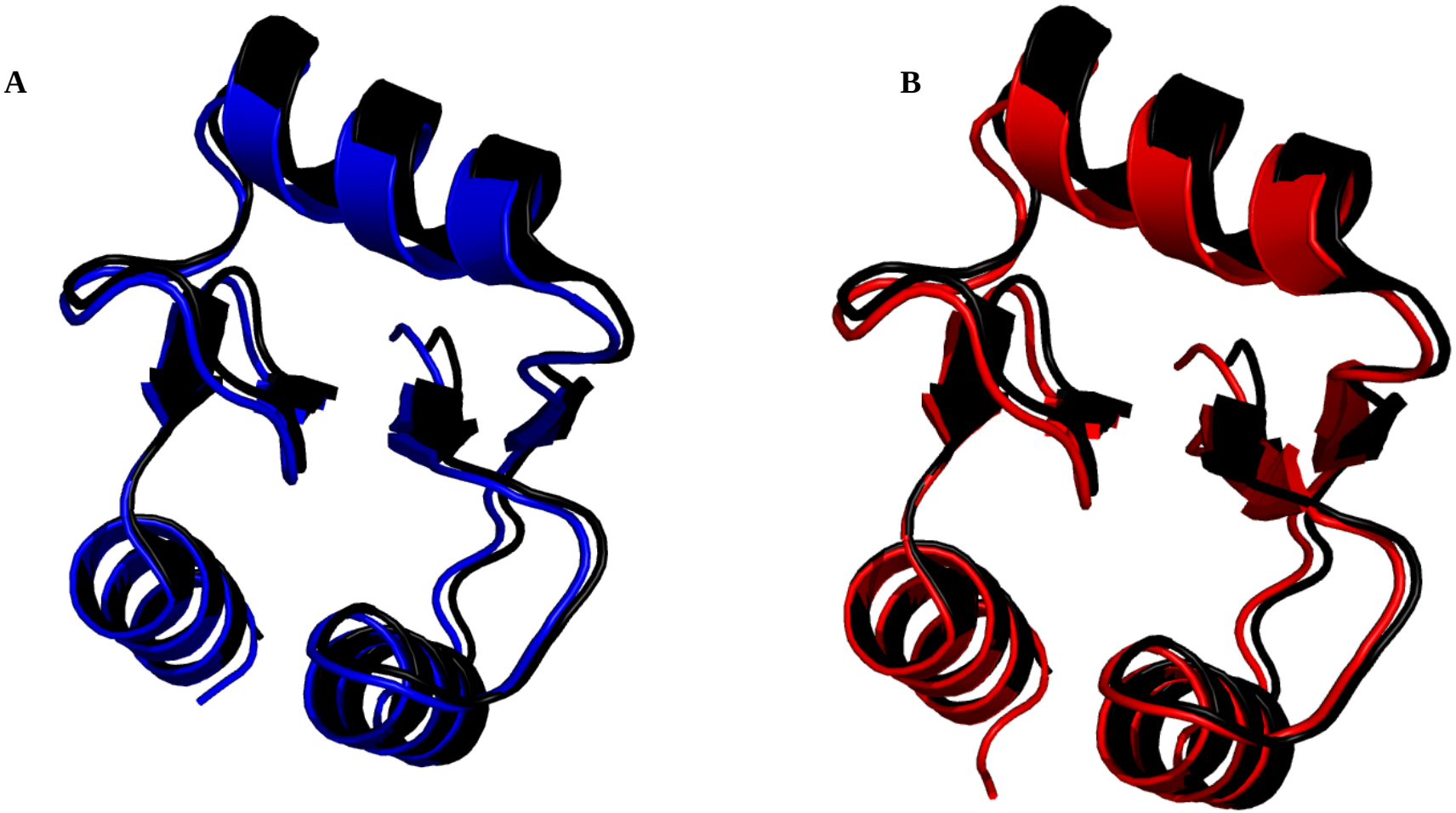


Figure 3-19. Structural alignment of wild-type and mutant hGSTA1-1 structures
Alignment of the Trx domain of (A) V58A (blue) and (B) I75A hGSTA1-1 (red) structures against the wild-type hGSTA1-1 protein (black). R.m.s.d values 0.58 Å (V58A) and 0.61 Å (I75A) were obtained for the alignment. Images rendered using PyMOL™, DeLano Scientific, 2006; PDB codes: 1K3O (Le Trong *et al*., 2002), 3ZFL, and 3ZFB.

CHAPTER 4: DISCUSSION

The GST fold, comprised of a thioredoxin (Trx) N-domain and an all α-helical C-domain, is present in at least 18 classes of proteins (SCOP database; Murzin *et al*., 1995) and found in both prokaryotes and eukaryotes. The Trx domain is further preserved in 22 other families of proteins. The primary role of GSTs is to detoxify endogenous and xenobiotic compounds by catalysing conjugating the tripeptide, GSH, to the electrophilic group thus rendering the compound inactive and soluble for excretion (reviewed in Armstrong 1997). The active site of GSTs spans the domain interface with the thioredoxin and C-terminal domains providing binding site for GSH and the xenobiotic compound, respectively (Armstrong 1997, Sheehan *et al*., 2001). While there has been substantial amount of work has been carried out in elucidating the reaction mechanism of the enzyme (Armstrong 1997; Dirr 1994b; Dirr and Wallace, 1999; Nilsson *et al*., 2002; Hegazy *et al*., 2004; Vararattanavech *et al*., 2006; Dourado *et al*., 2010; Gildenhuys *et al*., 2010a; Honaker *et al*., 2011) not much is known about some of conserved elements in contributing to the stability and function to the GSTs.

The Trx fold is made up of two subdomains: an N-termimal βαβ motif and a C-terminal ββα motif connected by an α-helix (Figure 4-1). These two subdomains form a tightly packed hydrophobic core (Figure 4-1). The Trx fold is highly adaptable for its function whilst maintaining its structure and this feature raises the question: which of its residues are responsible for maintaining the integrity of the fold? The role of some residues (Pro56 and Ile71), which are topologically conserved in the Trx fold, have been investigated in GSTs (Nathaniel *et al*., 2003; Achilonu *et al*., 2010), however the role of other topologically conserved elements of the hydrophobic core of the Trx domain in GSTs needs to be further explored.

A structure based sequence alignment of the different GST classes reveals several topologically conserved residues in the thioredoxin domain (Figure 1-5). This study has focused on further elucidating the role of non-active site topologically conserved residues of the hydrophobic core of the Trx domain.

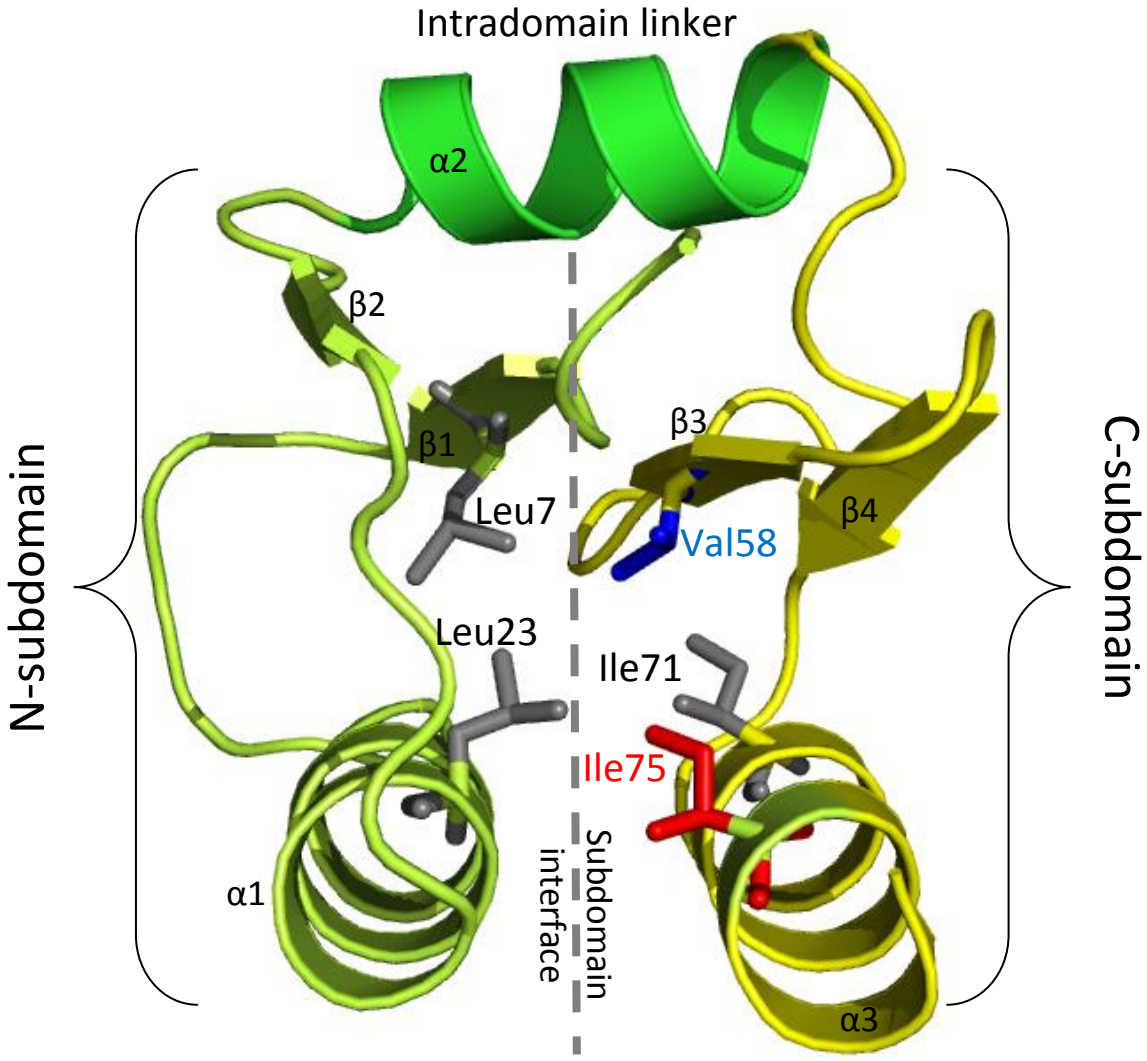


Figure 4-1. Trx domain in GSTs
Representation of the hydrophobic core of the Trx domain in GSTs. Topologically conserved residues of the N-subdomain (β1α1β2; lime green) and the C-subdomain (β3β4α3; yellow) are shown. The two subdomains are connected by a helix linker (α2; green). The residues that were mutated for this study are highlighted in blue (Val58) and red (Ile75).

Two additional residues on the C-subdomain have been found to be topologically conserved (Figure 1-5), namely Val58 and Ile75 in hGSTA1-1 and thus were the residues of choice to further investigate the role of the hydrophobic core in the Trx domain. The formation of favourable van der Waals interactions between such hydrophobic residues results in the formation of hydrophobic cores and is thought to be the nucleation event in the folding process of proteins (Dill, 1990). The resultant interactions are thus crucial in maintaining the stability of hydrophobic cores that stabilise the native protein structure.

The aforementioned residues were mutated to Ala to remove any interaction contributed to by the functional groups while not introducing a large degree of flexibility. A range of spectroscopic techniques were employed to analyse the native solution structure and function as well as to determine the conformational stability of the mutants and comparing them to the wild-type form of the protein.

4.1 Role of the C-subdomain in structural integrity

In order to tease out the function of a conserved residue, it is essential that its substitution does not affect the global structure of the protein so that mutant protein is structurally comparable to the reference (wild-type) protein.

The favourable comparison of the secondary structural content of both the V58A and I75A hGSTA1-1 proteins to the wild-type (Figure 3-4) confirmed that the cavity creating mutations caused no major alteration to the secondary structure. In the case of the V58A mutant, changes occurring on the β3 strand would be masked by the stronger α-helical signal [208 and 222 nm, (Woody, 1995)], thus any significant changes would result in a global structural change in the protein rather than a local alteration. An absence of any change in the secondary structural content is also witnessed in the I75A hGSTA1-1 protein. The crystal structures confirmed that the backbone of the protein remained unaffected as determined by the favourable r.m.s.d values of the mutant and wild-type structures (Figure 3-19).

Intrinsic tryptophan fluorescence is highly sensitive to the immediate environment around the indole ring thus providing low resolution in solution structural information

about this local environment. In the case of hGSTA1-1, a lone tryptophan (Trp21) per subunit is located at the domain interface with its indole ring protruding into domain 2 and is relatively solvent inaccessible [(SASA ~ 4Å^2, (Wallace *et al*., 2000)] therefore making it a suitable reporter of any changes occurring at the domain interface. The fluorescence emission spectra (Figure 3-5) showed that all the proteins maximally emit at 325 nm, characteristic of proteins with a buried indole ring in the native state (Lacowicz, 2006). An absence in a bathochromic or hyposchromic shift in the spectra (Figure 3-5) together with a good structural alignment (Figure 3-19), the induced mutations caused no structural alteration with respect to the domain interface.

A more global representation of the tertiary structure of the hGSTA1-1 proteins was determined by near-UV CD. The strong overlapping spectra (Figure 3-6) shows that there are no changes to the packing of the aromatic functional groups and again confirming that the tertiary environment is unaffected by the induced mutations.

The non-disruptive nature of the V58A and I75A mutations is consistent with the general observation of the structure seen in the wild-type being preserved in the single-point mutant proteins [(Matthews, 1991; Xu *et al*., 1998; P56G (Nathaniel *et al*., 2003); I71A (Achilonu *et al*., 2010); W21A (Balchin *et al*., 2010); R15L (Gildenhuys *et al*., 2010a).

In addition to the interaction of the Trx domain in GSTs with domain 2 of the same subunit, it also interacts with domain 2 of the adjacent subunit through the interaction of the α3 and α4 helices (Figure 4-2). Similar protein-protein interactions were disrupted when similar induced mutations were introduced in the GSK3β-axin interaction (Zhang *et al*., 2007). The cause of the disruption was a positional shift in the interacting helices. The I75A mutation does not seem to affect intersubunit domain interactions between domain 1 and domain 2 as there is no positional shift of the interacting helices as a result of the mutation (Figure 4-2) thus the dimeric nature of the protein has not been structurally compromised.

Other crystallographic studies have shown that the residues in the region around the point mutation are likely to be affected (Shortle, 1992; Eyal *et al*., 2003). In the case of the V58A and I75A mutations, the low resolution spectroscopic data in solution

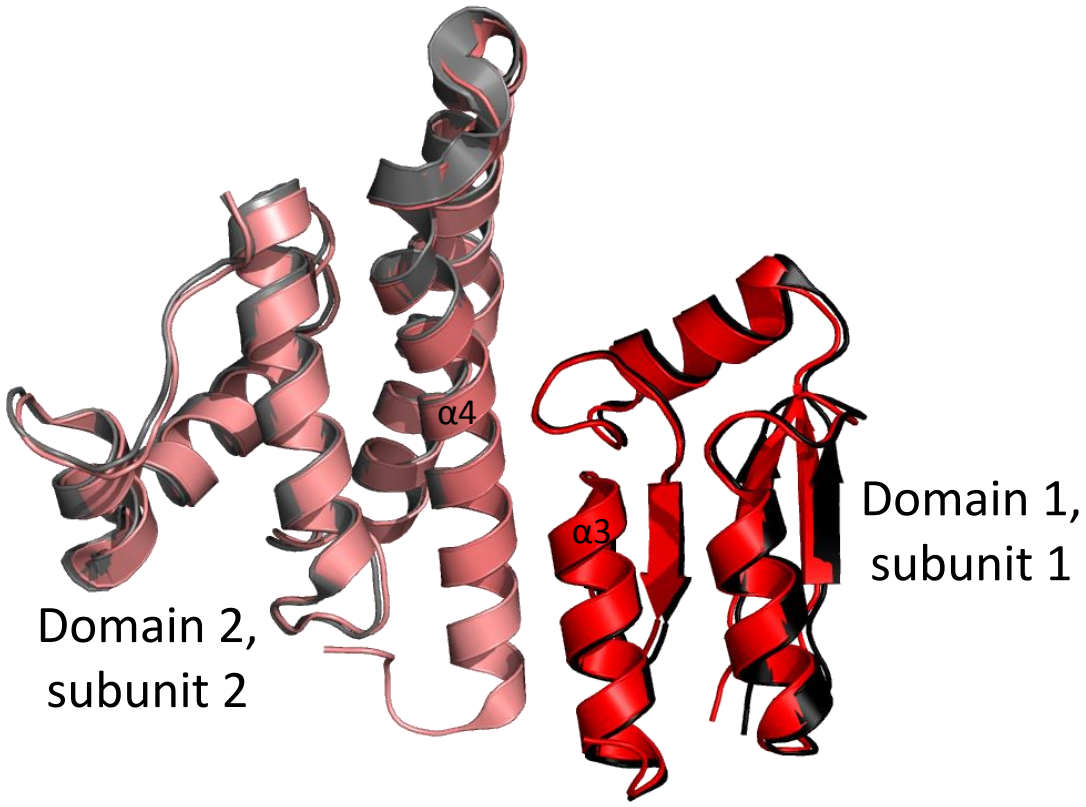


Figure 4-2. Domain interface of intersubunits of GSTs
Superimposition of the domain interaction of α3 of subunit 1 and α4 of subunit 2 between wild-type (black and gray) and I75A hGSTA1-1 (red and pink).

concurs well with the high resolution static crystal structural data, in that the global therefore the induced mutations are structurally non-disruptive and these residues alone are not crucial for maintaining the structural integrity in the native state.

4.2 Role in catalysis and ligandin function

The Trx domain of hGSTA1-1 constitutes the active site of the GSTs, thus any change to the Trx domain can impact on the catalytic function of the enzyme. The specific activities were determined (Table 3-1) and the enzymatic function of the proteins is still maintained even though cavities are created [V58A - 8Å^3; I75A - 55 Å^3; (cavities calculated using Swiss PDB viewer, Geux and Pietsch, 1997)] as a result of the induced mutations. This confirms that no structural alterations were induced and the active site dynamics are not affected as a result of the mutations.

This phenomenon has been seen in a previous study (Achilonu *et al*., 2010) involving another topologically conserved residue in the hydrophobic core of the Trx domain. The conserved Ile71 was mutated to alanine and valine and the enzyme was still active with slightly lower specific activity values; 86% and 97% for the alanine and valine substitutions, respectively (Achilonu *et al*., 2010). Furthermore, the domain interface constitutes the active site and thus together with the intrinsic tryptophan fluorescence data, indicate that there were no structural alterations at the interface.

In addition to its primary role of catalysis in detoxification, GSTs are also involved in the transport of many non-substrate ligands (Listowsky, 2005). The impact of its ligandin function by the mutations was assessed by ANS binding to apo-enzyme. Again, in line with the impact on catalytic function, the ligandin binding was insignificantly impacted on (Figure 3-7). This behaviour is seen when similar induced mutations are introduced at the Trx domain (Fisher, 2006, MSc dissertation, http://wiredspace.wits.ac.za/handle/10539/1726; Khoza, 2013, PhD thesis, http://wiredspace.wits.ac.za/handle/10539/12700). Mutations at the core of the Trx domain do not seem to affect the H-site as opposed to mutations at the domain interface (Balchin *et al*., 2010), subunit interface (Alves *et al*., 2006) and immediate vicinity of the active site (Gildenhuys *et al*., 2010a)

Although the ligandin function and enzymatic catalysis is maintained, these residues play a role in the optimal binding of the substrate as suggested by the lowered specific activity values (Table 3-1). In the case of the class Delta GST, mutations of the residues in the hydrophobic core of the N-subdomain of Trx directly affected GSH binding and as a result almost abolished enzymatic activity [1%, (Vararattanavech *et al*., 2006)]. Mutations of residues in the C-subdomain of class Delta GSTs also involved in a hydrophobic core, however, had only slightly lowered activity values (95%) and an decreased affinity for GSH (Winayanuwattikun and Ketterman, 2004). A similar trend is seen for mutations of residues in the C-subdomain of the Trx domain in class Omega GSTs, where the binding affinities for its substrates were lowered (Yamamoto *et al*., 2013).

Although the I75A mutation may be more deleterious in terms of conformational stability, the lower enzymatic activity of the V58A mutation can be attributed to an impact on long range interactions with the substrate (Gromiha and Selvaraj, 2004). Pro56, the *cis*-Pro that is also conserved in the Trx fold, is located on the same β strand as Val58 and the mutation of Pro56 knocked out enzymatic activity (Nathaniel *et al*., 2003). Thus, interactions mediated through β3, could be altered to compromise enzymatic activity.

Thus the induced mutations of both topologically conserved residues in the C-subdomain of the Trx domain do not alter the native structure as to abolish ligand binding and catalysis but are responsible for optimum binding for its substrates and hence catalysis.

4.3 Impact of the Trx C-subdomain on stability of GSTs

The conformational stabilities of the hGSTA1-1 proteins were determined by urea-induced equilibrium unfolding experiments. Extrinsic fluorescence, ANS binding, was used to probe for the presence of molten-globule intermediates. A spectroscopic independent probe, limited pulse proteolysis was employed to verify the results of the spectroscopic probes.

Although the proposed unfolding mechanism of the wild-type hGSTA1-1 is a three-state model (Wallace *et al*., 1998) due to the unfolding of α9 (section 1.4.3.1), the global unfolding of the protein follows a highly co-operative, two-state unfolding mechanism with the absence of a stable intermediate forming. The thermodynamic parameters (Table 3-3) compare favourably as previously reported (Wallace *et al*., 1998). The unfolding behaviour of V58A hGSTA1-1 closely resembles that of the wild-type protein in that the highly co-operative two-state behaviour is maintained and the thermodynamic parameters (Table 3-3) are similar to those of the wild-type hGSTA1-1 with a marginal destabilisation ($\Delta\Delta G(H_2O)$ ~2 kcal.mol^{-1}). Furthermore, the unfolding as monitored by the extrinsic fluorescent probe, ANS, follows the same trend as wild-type (Figure 3-10 and Figure 3-11) and no stable intermediates are formed. Thus the topologically conserved Val58 located on β3 has a small role in the stability of hGSTA1-1.

This behaviour can be attributed to the fact that Val58 has fewer contacts than the other the surrounding residues have with one another and consequently fewer van der Waals interactions in the hydrophobic core of the Trx domain (Figure 4-3) with only an 8 Å^3 cavity being created as a result of the mutation and hence the packing of the core is not severely impacted on.

Val58 is located on the β3 sheet in the GST-fold and the equivalent β-strand in thioredoxin was found to be relatively less rigid than the α-helices as determined by NMR measurements (Mottonen *et al*., 2009). Although the formation of the central β-sheet is thought to be the first step in the folding of Trx (Tasayco *et al*., 2000), these residues on these strands have not been shown to be crucial in the stability of the Trx fold (Garcia-Seisdedos *et al*., 2012). Furthermore, the adjacent β-strand is anti-parallel in direction (Figure 1-2), resulting in a more stable β-sheet structure than the parallel β-sheet structure as a result of the linear hydrogen bonds between the peptide groups of the backbone. Therefore the marginal destabilisation witnessed is a result of the orientation of the Val58 residue and the loss of interactions with surrounding residues is compensated by the backbone movements of the β-strand.

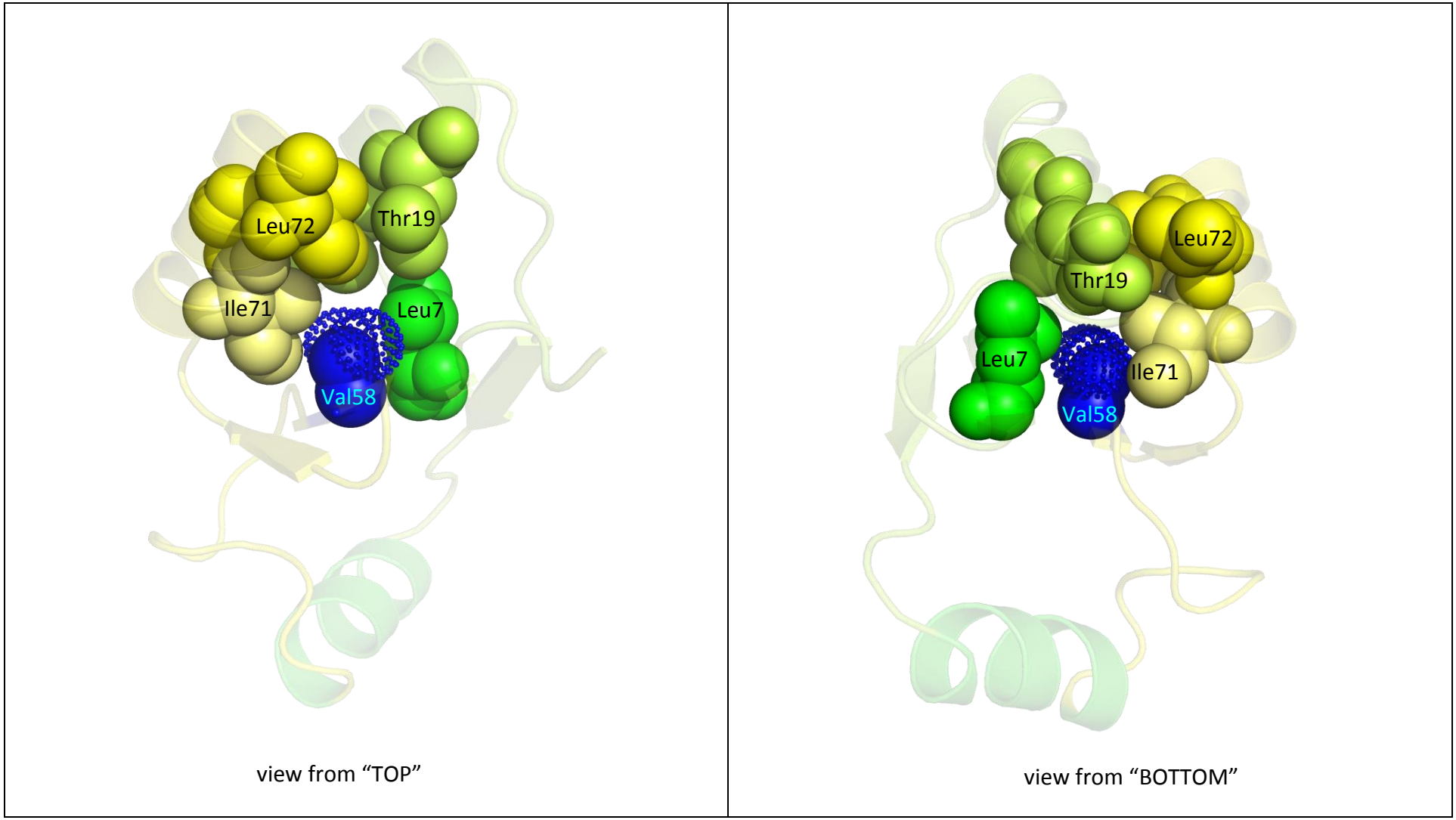


Figure 4-3. Contribution of Val58 to the hydrophobic core in hGSTA1-1

A superimposed space filling representation of the wild-type and V58A hGSTA1-1 of the Trx-domain core residues and their contacts. The contribution of the aliphatic side chain of Val58 is shown as dotted spheres (blue) while that of Ala58 is shown as solid spheres (blue). Images rendered using PyMOL^{TM}, DeLano Scientific, 2006; PDB codes: 1K3O (Le Trong *et al*., 2002) and 3ZFL.

Pulse proteolysis was used to corroborate the stability data seen by far-UV CD and fluorescence spectroscopy. The V58A hGSTA1-1 showed a similar pattern of proteolytic digestion as the wild-type hGSTA1-1 (Figure 3-14) confirming that the stability compares favourably with that of the wild-type hGSTA1-1 and that the stability of the Trx domain is not significantly impacted on by the V58A mutation.

Ile75, one of two topologically conserved Ile residues on α3, has a more pronounced role in the stability of GSTs. Substitution of Ile75 causes a dramatic shift in the unfolding mechanism from two-state to three-state, resulting in the formation of at least one stable intermediate. The strong unfolding co-operativity seen in wild-type hGSTA1-1 is also diminished as seen by the biphasic, non co-incident transitions of the CD and fluorescent spectral data (Figure 3-15). In this study, changes in the CD data can be attributed to the changes in the C-terminal domain of the protein since the α-helical content of this protein is predominantly found in this domain (Figure 1-3). In contrast, the Trp fluorescence unfolding data will provide information only to the local environment of the buried residue which is located at the interdomain interface of hGSTA1-1. The unfolding data displays a greater loss of tertiary structure at the interdomain interface whereas the secondary structural content remains higher at urea concentrations of 2 – 3 M. The data suggests that the mutation has had an impact on the stability of the N-terminal domain and has resulted in the uncoupling of the domain interface hence the exposure of Trp21 to the bulk solvent. The C-terminal domain however remains intact as seen by the relatively higher secondary structural content. The intermediate is a result of hydrophobic patches of the Trx domain interface being exposed to solvent and thus being probed by the hydrophobic ANS dye. A dimeric intermediate is proposed here as the existence of stable monomers remains questionable (Fabrini *et al*., 2009), especially at the concentrations used in this study (2μM) which is almost 10-fold higher than the dissociation constant of class Alpha GSTs [0.3 μM, (Vargo *et al*., 2004)]. Furthermore, the subunit interface is predominantly hydrophobic with a strong 'lock and key' interaction (Sayed *et al*., 2000) thus precluding the existence of stable monomers.

The destabilisation was again confirmed with the spectroscopic independent probe, pulse proteolysis and a distinct shift in proteolytic digestion was observed (Figure 3-15). The thermolysin enzyme targets hydrophobic residues [Ile, Val, Leu, Ala, Met,

Phe (Keil, 1992)] and thus the increased proteolytic cleavage at lower denaturant concentrations is a result of the destabilisation of the Trx domain by the I75A mutation causing the amphipathic core to be exposed. This further proves the more prominent role of Ile75 in the stability of hGSTA1-1.

The removal of a buried methyl group is predicted to result in a destabilisation by 1.1 kcal/mol (Eriksson *et al*., 1992) and in the case of an Ile → Ala mutation, three methyl groups are removed causing an estimated predicted loss of ~ 3.3 kcal/mol, however the loss is much higher in the case of the I75A mutation of hGSTA1-1 ($\Delta\Delta G(H_2O)$ ~ 10 kcal/mol). Previously reported data involving the Trx hydrophobic core in GSTs showed marked decreases in the conformational stability and unfolding co-operativity suggesting the formation of intermediates but were not stable enough to be detected by the spectroscopic probes (Nathaniel *et al*., 2003 Achilonu *et al*., 2010). Achilonu *et al*. (2010) found that a similar mutation on the same helix (α3), I71A, did not show such a dramatic decrease in stability but did suggest intermediates forming as suggested by the decrease in unfolding co-operativity. The lack of a marked decrease in stability is attributed to two ordered water molecules filling the cavity created in the I71A mutation, thus providing compensation by mediating interactions at the hydrophobic core. In the case of the I75A hGSTA1-1 crystal, no solvent molecules were found in the 55$Å^3$ created cavity thus leaving unsatisfied van der Waals contacts at the core (Figure 4-4). This unfolding behaviour is also seen in the case of mutating aliphatic residues at the hydrophobic core of γD crystalline proteins (Moreau and King, 2009). The alanine substitutions also resulted in lower melting temperatures as seen in this study with the formation of stable intermediates. Therefore Ile75 plays a role in maintaining the conformational stability of GSTs and domain co-operativity.

All enzymes require the active site to be more dynamic than the rest of the protein to allow substrate binding and even more so in promiscuous enzymes that bind a wide variety of substrates such as the GSTs (Nath and Atkins, 2008). Mutations of highly conserved residues in *Escherichia coli* Trx show that they are important in the conformation of the active site by providing a charge distribution which optimises interactions with its substrates and redox potential (Gleason, 1992) and ionic residues play a minor role in the stability in the fold (Esposito *et al*., 2012). The Trx domain in

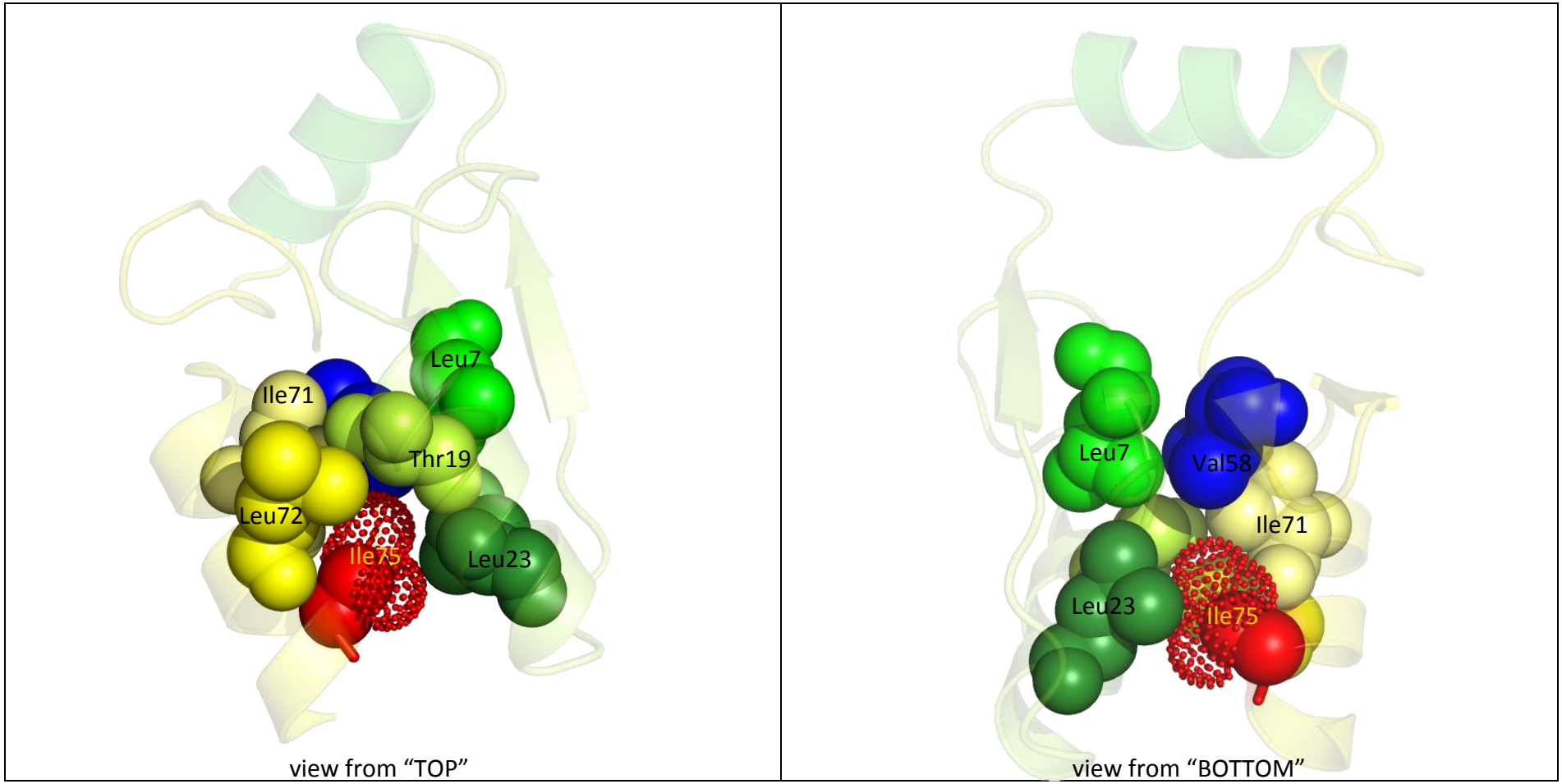


Figure 4-4. Contribution of Ile75 to the hydrophobic core in hGSTA1-1

A superimposed space filling representation of the wild-type and I75A hGSTA1-1 of the Trx-domain core residues and their contacts. The contribution of the aliphatic side chain of Ile75 is shown as dotted spheres (red) while that of Ala75 is shown as solid spheres (red). Images rendered using PyMOL^{TM}, DeLano Scientific, 2006; PDB codes: 1K3O (Le Trong *et al*., 2002) and 3ZFB.

GSTs is conformationally more dynamic than the C-terminal domain (Thompson *et al*., 2006; Stoychev *et al*., 2009) thus allowing the Trx domain to be the substrate binding domain. However, it has been shown that the Trx domain and the C-terminal all α-helical domain behave in a tightly cooperative manner (Dirr and Reneimer, 1991; Erhardt and Dirr 1995; Wallace *et al*, 1998; Stevens *et al*., 1998; Gildenhuys *et al*., 2010b; Parbhoo *et al*., 2011) and thus has a vital role in maintaining stability of GSTs.

Ile75 is located on α3 thus making it involved in subdomain interactions with α1 of the N-subdomain in the Trx domain (Figure 4-1). It is also more extensively interacting with surrounding residues in the hydrophobic core of the Trx domain (Figures 4-4). Ile75 therefore stabilises the Trx domain by maintaining the compactness of the subdomain interface. The destabilisation of the Trx domain occurs as a result of many van der Waals interactions being absent at the hydrophobic core. Unlike the β-strand structure, an α-helix is more rigid with fewer degrees of freedom in its conformational space and is thus unable to over-compensate for the resulting loss of interactions by the side-chains. Mottonen *et al*. (2009) found that the equivalent of α3 in thioredoxin is the most rigid helix of the protein hence any disruption in this area of the scaffold may be detrimental in overall stability. Hydrogen-exchange data of a monomeric GST homologue, hCLIC1 (Stoychev *et al*., 2009), as well as a dimeric class Mu GST (Thompson *et al*., 2006), showed the same relative rigidness with respect to the rest of the protein

The two subdomains of Trx were studied in isolation in which the two subdomains were synthesised as separate fragments (Tasayco *et al*., 2000; Santos *et al*., 2009). While the subdomains do assume some native structure, disordered regions were found to be at the subdomain interface and only assumed complete native conformation upon interaction with the opposite domain (Tasayco and Chao, 1995; Yu *et al*. 2000; Santos *et al*., 2007; Santos *et al*., 2009). In the full length protein, aliphatic residues at the domain interface were mutated to alanine maintaining the hydrophobic profile, and the result was a significant decrease in stability ($\Delta\Delta G$ 3.0 – 3.7 kcal.mol^{-1}) (Santos *et al*., 2009). The authors also proposed an energetic landscape for these subdomains in isolation in that the funnel ends with a basin populated with the molten globule state (Santos *et al*., 2009).

A study examined the contribution of the C-terminal helix to the stability and folding of Trx (Santos *et al*., 2007). Briefly, a large fragment (residues 1-93) of the full length protein (108 aa) and the truncated helix (residues 94-108) were examined in isolation and in complex. The large fragment showed molten globule like characteristics as determined by ANS binding in the absence of the C-terminal helix. The fragment also had higher susceptibility to limited proteolysis than the full length protein. Upon interaction with the C-terminal helix, native-like characteristics were witnessed as determined by circular dichroism and intrinsic fluorescence. Enzymatic activity was also restored as a result of the interaction. Santos *et al*. (2007) proposed that the C-terminal helix may be involved in the final events of folding resulting in the consolidation of the Trx fold.

In the case of the Trx domain in GSTs, α3 is the C-terminal helix. The resultant intermediate bound to ANS (Figure 3-12) thus displaying molten globule-like properties with the hydrophobic core being solvent exposed. The increased susceptibility to limited proteolysis (Figure 3-14) confirms the exposure of the hydrophobic residues.

A later study examined the dynamics of the interaction between the fragment of Trx and the truncated C-terminal helix using NMR (Binolfi *et al*., 2012). In isolation, it was found that α2 and α3 of the fragment and the C-terminal helix was disordered and contained no helical character. A native-like conformation occurred upon interaction with the stabilisation of the α2 and α3 helices and the complex also displayed native-like dynamics of the backbone. This data highlighted the role of the C-terminal helix in consolidating the Trx structure and that stabilisation and folding occurred through molecular recognition. Furthermore, interaction with the C-terminal helix allows Trx to act as a co-operative folding unit.

In GSTs, the C-terminal helix (α3) of the Trx domain can also be involved in maintaining the folding co-operativity. Achilonu *et al*. (2010) showed that the mutation of Ile71 reduced the *m*-value and therefore a reduction in folding co-operativity. This study shows that Ile75, located on α3 as well, shows a distinct deviation from a two-state to a three-state unfolding mechanism and thus a loss of

folding co-operativity. The α3 interaction with the rest of the Trx domain is suggested to be the co-operativity determinant and the removal of the interactions contributed by Ile75 allows a stable intermediate to be populated and detected spectroscopically.

The van der Waals interactions made by the chain of Ile75 on α3 and the residues occupying the hydrophobic pocket is far greater than any other residue in the Trx domain (Figure 4-3). The contacts span an area of 49.5Å^2, as computed by the Contact Map Analysis tool (http://ligin.weizmann.ac.il/cma/; Sobolev *et al*., 2005). An interaction schematic of subdomains in the Trx domain (Figure 4-5) shows that Ile75 is involved in only hydrophobic interactions with several residues, while Val58 is involved in fewer interactions with hydrophobic residues.

Mutations of residues in Trx and in the Trx-domain of GSTs have proven the importance of the van der Waals interactions at the core of the domain, which are required for maintaining the stability of the protein structure through optimised packing (Wallace *et al*., 2000; Tasayco *et al*., 2000; Nathaniel *et al*., 2003 Santos *et al*., 2007, 2009; Achilonu *et al*., 2010; Balchin *et al*., 2010, Parbhoo *et al*., 2011). Folding of the Trx domain is achieved through interactions between the N and C-subdomains (Tasayco *et al*., 2000; Santos *et al*., 2009). The initial step of folding is driven by hydrophobic interactions between the β1 and β3 strands with the final steps requiring interactions between α1 and α3 (Tasayco *et al*., 2000; Santos *et al*., 2009). Thus the folding and stability of the Trx domain is maintained through the hydrophobic network of interactions between the subdomains.

Similar induced mutations of C-subdomain residues, while not topologically equivalent to those in this study, involved in the core of the Trx domain of class Delta GSTs, are believed to play a role in the folding process as determined by refolding experiments (Winayanuwattikun and Ketterman, 2004).

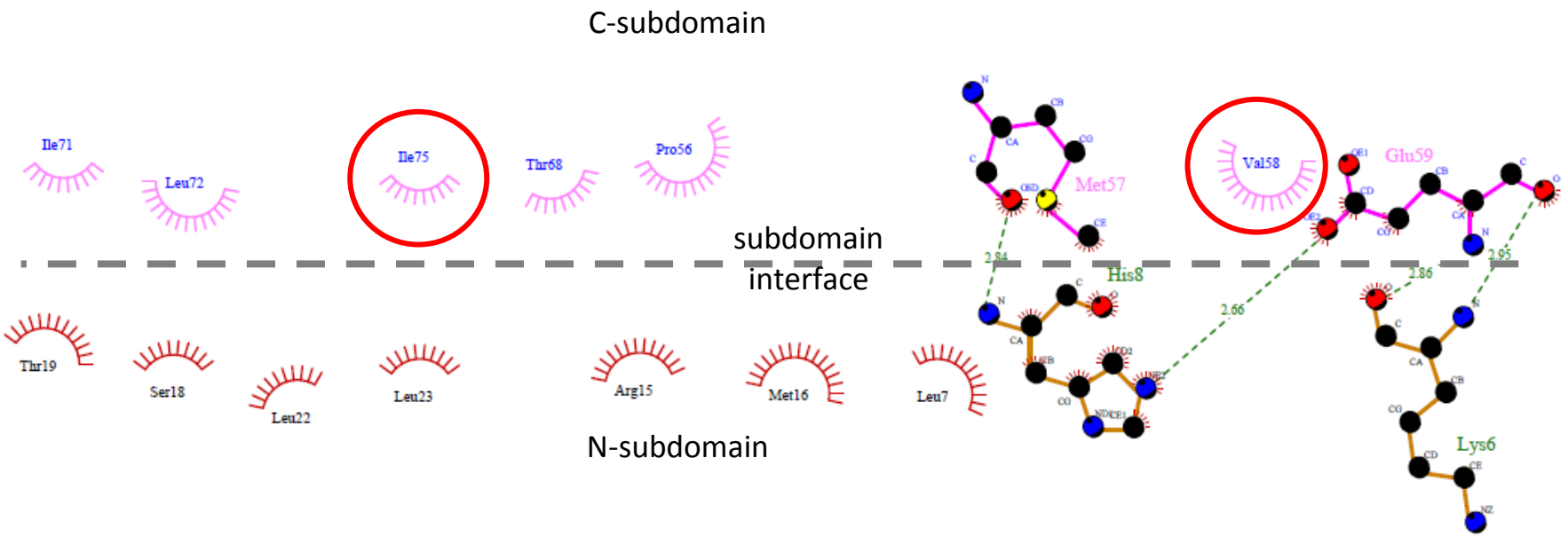


Figure 4-5. Thioredoxin subdomain interface interaction network
Schematic representation of the subdomain interaction between the N- and C-subdomains of hGSTA1-1. Hydrophobic interactions are shown by the spiked symbols, polar and charged residues are represented by ball and stick symbols and hydrogen bonds are shown by green dotted lines. The residues in question in this study are encircled in red. Image generated using Dimplot (Laskowski and Swindells, 2011)

4.4 Conclusion

The role of the C-subdomain conserved hydrophobic residues at the core of the Trx domain does not contribute toward ligand binding. The C-subdomain also does not contribute toward enzymatic function directly but does seem to affect catalysis indirectly by mediating long range interactions with the substrate. However, the C-subdomain conserved residues do have a prominent role in maintaining the stability of the enzyme which strongly correlates with Trx proteins where thermodynamic stability and function do not correlate (Hellinga *et al*., 1992; Assemat *et al*., 1995). Given the highly adaptable nature of the Trx-fold (Pan and Bardwell, 2006), there seems to be no major structural alteration of the neighbouring residues to compensate for cavity causing mutations. This tightly packed hydrophobic core of the Trx domain is conserved from archae to mammals and can be considered to be the stabilising force in proteins with a Trx fold which is line with conserved hydrophobic residues of other protein families lying in clusters rather than in isolation (Guharoy and Chakrabati, 2010).

Computational analyses of protein-protein interface hot spots which reveals that conserved residues occur in clustered, tightly packed regions and that their contribution to stability are co-operative (Guharoy and Chkrabati 2005; Keskin *et al*., 2005). Previous studies examining the Trx C-subdomain of GSTs [Pro56, (Nathaniel *et al*., 2003; Ile71, (Achilonu *et al*., 2010)] showed that mutations of other topologically conserved residues diminished the protein's stability and showed a decrease in unfolding co-operativity. The intermediates were spectroscopically undetectable, however mutation of Ile75 has populated a stable intermediate. The Ile75 residue, as is Ile71, is located on the C-terminal helix (α3) of the Trx domain. In the context of the Trx domain in GSTs, the α3 helix is proposed to be the co-operativity determinant of its unfolding.

Conserved hydrophobic cores are also implicated in the formation of folding nuclei therefore future studies should examine the kinetics of folding to ascertain the role of these residues in the folding pathway of GSTs. The dynamics of the Trx domain in GSTs should be examined at high resolution (e.g. NMR) to ascertain the role of these residues.

CHAPTER 5. REFERENCES

Abdalla, A.-M., Bruns, C.M., Tainer, J.A., Mannervik, B. and Stenberg, G. (2002). Design of a monomeric human glutathione transferase GSTP1, a structurally stable but catalytically inactive protein. *Protein Eng*. **15**, 827-834.

Achilonu, I., Gildenhuys, S., Fisher, L., Burke, J., Fanucchi, S., Sewell, B.T., Fernandes, M. and Dirr, H.W. (2010). The role of a topologically conserved isoleucine in glutathione transferase structure, stability and function. *Acta Crystallogr. Sect. F. Struct. Biol. Cryst. Commun.* **66**, 776-780.

Adler, V., Yin, Z., Fuchs, S. Y., Benezra, M., Rosario, L., Tew, K. D., Pincus, M. R., Sardana, M., Henderson, C. J., Wolf, C. R., Davis, R. J. and Ronai, Z. (1999a). Regulation of JNK signaling by GSTp. *EMBO J.* **18**, 1321-34.

Arnér, E.S.J. and Holmgren, A. (2006). The thioredoxin system in cancer. *Semin. Cancer Biol*. **16**, 420-426.

Alfadda, A.A. and Sallam, R.M. (2012). Reactive oxygen species in health and disease. *J. Biomed. Biotechnol*. doi:10.1155/2012/936486.

Alves, C. S., Kuhnert, D. C., Sayed, Y. and Dirr, H. W. (2006). The intersubunit lock-and-key motif in human glutathione transferase A1-1: role of the key residues Met51 and Phe52 in function and dimer stability. *Biochem. J.* **393**, 523-8.

Armstrong, R. N. (1997). Structure, catalytic mechanism, and evolution of the glutathione transferases. *Chem. Res. Toxicol.* **10**, 2-18.

Assemat, K., Alzari, P.M. and Clement-Metral, J. (1995). Conservative substitutions in the hydrophobic core of *Rhodobacter sphaeroides* thioredoxin produce distinct functional effects. *Protein. Sci*. **4**, 2510-2516.

Balchin, D., Fanucchi, S., Achilonu, I., Adamson, R.J. Burke, J. Fernandes, M., Gildenhuys, S. and H.W. Dirr. (2010). Stability of the domain interface contributes towards the catalytic function at the H-site of class alpha glutathione transferase A1-1. *Biochim. Biophys. Acta.* **1804**, 2228- 2233.

Benjwal, S., Verma, S., Rohm, K. H. and Gursky, O. (2006). Monitoring protein aggregation during thermal unfolding in circular dichroism experiments. *Protein Sci.* **15**, 635-9.

Berndt, C., Lillig, C.H. and Holmgren, A. (2008). Thioredoxins and glutaredoxins as facilitators of protein folding. *BBA – Mol. Cell Res*. **1783**, 641-650.

Bhattacharyya, S. and Varadarajan, R. (2013). Packing in molten globules and native states. *Curr. Opin. Struct. Biol.* **23**, 11-21.

Binolfi, A., Fernández C.O., Sica M.P., Delfino, J.M. and Santos, J. (2012). Recognition between a short unstructured peptide and a partially folded fragment leads to the thioredoxin fold sharing native-like dynamics. *Proteins*. **80**, 1448-64

Board P.G., Coggan, M, Chelvanayagam, G., Easteal, S., Jermiin, L.S., Schulte, G.K., Danley D.E., Hoth, L.R., Griffor, M.C., Kamath, A.V., Rosner, M.H., Chrunyk, B.A., Perregaux, D.E., Gabel, C.A., Geoghegan, K.F. and Pandit J. (2000). Identification, characterization, and crystal structure of the Omega class glutathione transferases. *J. Biol. Chem*. **275,** 24798-806.

Bordner, A.J. and Abagyan, R. (2005). Statistical analysis and prediction of protein-protein interfaces. *Proteins: Structure, Function and Genetics*. **60**, 353-366.

Bousset L, Belrhali H, Melki R, Morera S. (2001). Crystal structures of the yeast prion Ure2p functional region in complex with glutathione and related compounds. *Biochemistry*.**40**, 13564-73.

Braman, J., Papworth, C. and Greener, A. (1996). Site-directed mutagenesis using double-stranded plasmid DNA templates. *Meth. Mol. Biol*. **57**, 31-44.

Brünger, A.T. (1992). Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature*. **355**, 472-5.

Calamai, M., Chiti, F. and Dobson, C. M. (2005). Amyloid fibril formation can proceed from different conformations of a partially unfolded protein. *Biophys. J*. **89**, 4201-10.

Cameron, A. D., Sinning, I., L'Hermite, G., Olin, B., Board, P. G., Mannervik, B. and Jones, T. A. (1995). Structural analysis of human Alpha-class glutathione transferase A1-1 in the apo-form and in complexes with ethacrynic acid and its glutathione conjugate. *Structure.* **3**, 717-27.

Chaffotte, A. F., Cadieux, C., Guillou, Y. and Goldberg, M. E. (1992). A possible initial folding intermediate: the C-terminal proteolytic domain of tryptophan synthase β-chains folds in less than 4 milliseconds into a condensed state with non-native-like secondary structure. *Biochemistry*. **31**, 4303-8.

Chen, V.B., W.B. Arendall, 3rd, J.J. Headd, D.A. Keedy, R.M. Immormino, G.J. Kapral, L.W. Murray, J.S. Richardson, and D.C. Richardson. (2010). MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **66**,12-21.

Chern, M. K., Wu, T. C., Hsieh, C. H., Chou, C. C., Liu, L.F., Kuan, I. C., Yeh, Y. H., Hsiao, C. D. and Tam, M. F. (2000). Tyr115, gln165 and trp209 contribute to the 1, 2-epoxy-3-(p-nitrophenoxy)propane-conjugating activity of glutathione S-transferase cGSTM1-1. *J. Mol. Biol*. **300,** 1257-69.

Cho, S. G., Lee, Y. H., Park, H. S., Ryoo, K., Kang, K. W., Park, J., Eom, S. J., Kim, M. J., Chang, T. S., Choi, S. Y., Shim, J., Kim, Y., Dong, M. S., Lee, M. J., Kim, S. G., Ichijo, H. and Choi, E. J. (2001). Glutathione S-transferase mu modulates the stress-activated signals by suppressing apoptosis signal-regulating kinase 1. *J. Biol. Chem.* **276**, 12749-55.

Corrêa, D.H., Ramos, C.H. (2009). The use of circular dichroism spectroscopy to study protein folding, form and function. *African J. Biochem. Res*. **3**, 164-173

Chung, C. T., Niemela, S. L. and Miller, R. H. (1989). One-step preparation of competent *E. coli*: transformation and storage of bacterial cells in the same solution. *Proc. Natl. Acad. Sci. U. S. A.* **86**, 2172-5.

Dalle-Donne, I., Rossi, R., Giustarini, D., Colombo, R. and Milzani, A. (2007). S-glutathionylation in protein redox regulation. *Free Radical Biology and Medicine*. **43**, 883-898.

Dill, K. A. (1990). Dominant forces in protein folding. *Biochemistry*. **29**, 7133-55.

Dill, K. A., Alonso, D. O. and Hutchinson, K. (1989). Thermal stabilities of globular proteins. *Biochemistry*. **28**, 5439-49.

Dill, K. A. and Chan, H. S. (1997). From Levinthal to pathways to funnels. *Nat. Struct. Biol.* **4**, 10-9.

Dill, K.A., Ozkan, S.B., Weikl, T.R., Chodera, J.D. and Voelz, V.A. (2007). The protein folding problem: when will it be solved? *Curr. Opin. Struct. Biol*. **17**, 342–346.

Dirr, H.W. (2001) Folding and assembly of glutathione transfeases. *Chem. Biol. Interact*. **133**, 19-23.

Dirr, H.W. and Reinemer, P. (1991). Equilibrium unfolding of class π glutathione S-transferase. *Biochem. Biophys. Res. Commun*. **180**, 294-300.

Dirr, H., Reinemer, P. and Huber, R. (1994a). Refined crystal structure of porcine class Pi glutathione S-transferase (pGST P1-1) at 2.1 Å resolution. *J. Mol. Biol.* **243**, 72-92.

Dirr, H., Reinemer, P. and Huber, R. (1994b). X-ray crystal structures of cytosolic glutathione S-transferases. Implications for protein architecture, substrate recognition and catalytic function. *Eur. J. Biochem.* **220**, 645-61.

Dirr, H. W. and Wallace, L. A. (1999). Role of the C-terminal helix 9 in the stability and ligandin function of class Alpha glutathione transferase A1-1. *Biochemistry.* **38**, 15631-40.

Dirr, H. W., Little, T., Kuhnert, D. C. and Sayed, Y. (2005). A conserved N-capping motif contributes significantly to the stabilization and dynamics of the C-terminal region of class Alpha glutathione S-transferases. *J. Biol. Chem.* **280**, 19480-7.

Dobson, C.M. (2004). Principles of protein folding, misfolding and aggregation. *Semin. Cell Dev. Biol*. **15**, 3-1.

Dobson, C. M. and Karplus, M. (1999). The fundamentals of protein folding: bringing together theory and experiment. *Curr. Opin. Struct. Biol.* **9**, 92-101.

Dourado, D.F., Fernandes, P.A., Ramos, M.J. (2010). Glutathione transferase classes alpha, pi, and mu: GSH activation mechanism. *J. Phys. Chem. B.* **114**, 12972-80.

Dulhunty, A., Gage, P., Curtis, S., Chelvanayagam, G. and Board, P. (2001). The glutathione transferase structural family includes a nuclear chloride channel and a ryanodine receptor calcium release channel modulator. *J. Biol. Chem.* **276**, 3319-23.

Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr*. **60,** 2126-2132.

Engelhard, M., and Evans, P.A. (1995). Kinetics of interaction of partially folded proteins with a hydrophobic dye: evidence that molten globule character is maximal in early folding intermediates. *Protein Sci*. **4**, 1553-1562.

Epp O, Ladenstein R, Wendel A. (1983). The refined structure of the selenoenzyme glutathione peroxidase at 0.2-nm resolution. *Eur. J. Biochem*. **133**, 51-69.

Erhardt, J. and Dirr, H. (1995). Native dimer stabilises the subunit tertiary structure of porcine class pi glutathione S-transferase. *Eur. J. Biochem.* **230**, 614-20.

Eriksson A.E., Baase, W.A., Zhang, X.J., Heinz, D.W., Blaber, M., Baldwin, E.P. and Matthews, B.W. (1992). Response of a protein structure to cavity-creating mutations and its relation to the hydrophobic effect. *Science*. **255**, 178-83.

Esposito, L., Ruggiero, A., Masullo, M., Ruocco, M.R., Lamberti, A., Arcari, P., Zagari, A. and Vitagliano, L. (2012). Crystallographic and spectroscopic characterizations of *Sulfolobus solfataricus* TrxA1 provide insights into the determinants of thioredoxin fold stability. *J. Struct. Biol.* **177**, 506-512.

Eyal, E., Najmanovich, R., Edelman, M. and Sobolev, V. (2003). Protein side-chain rearrangement in regions of point mutations. *Proteins: Struct. Funct. Bioinf.* **50**, 272-282.

Fabrini, R., De Luca, A., Stella, L., Mei, G., Orioni, B., Ciccone, S., Federici, G., Lo Bello, M. and Ricci, G. (2009). Monomer-dimer equilibrium in glutathione transferases: a critical re-examination. *Biochemistry.* **48**, 10473-82.

Fanucchi, S., Adamson, R. J. and Dirr, H. W. (2008). Formation of an unfolding intermediate state of soluble chloride intracellular channel protein CLIC1 at acidic pH. *Biochemistry.* **47**, 11674-81.

Fernandes, A.P. and Holmgren, A. (2004). Glutaredoxins: Glutathione-Dependent Redox Enzymes with Functions Far Beyond a Simple Thioredoxin Backup System. *Antiox. Redox Sign.* **6**, 63-74.

Fisher L. (2006). Role of a topologically conserved isoleucine in the structure and function of glutathione transferases. MSc dissertation. University of Witwatersrand. [http://wiredspace.wits.ac.za/handle/10539/1726].

Freire, E. (1995). Thermodynamics of partly folded intermediates in proteins. *Annu. Rev. Biophys. Biomol. Struct.* **24**, 141-65.

Freire, E., Murphy, K. P., Sanchez-Ruiz, J. M., Galisteo, M. L. and Privalov, P. L. (1992). The molecular basis of cooperativity in protein folding. Thermodynamic dissection of interdomain interactions in phosphoglycerate kinase. *Biochemistry.* **31**, 250-6.

Fritz-Wolf K, Becker A, Rahlfs S, Harwaldt P, Schirmer RH, Kabsch W, Becker K. (2003). X-ray structure of glutathione S-transferase from the malarial parasite *Plasmodium falciparum*. *Proc Natl Acad Sci U S A.* **100**, 13821-6.

Garcia-Seisdedos, H., Ibarra-Molero, B. and Sanchez-Ruiz, J.M. (2012) How many ionizable groups can sit on a protein hydrophobic core? *Proteins*. **80**, 1-7.

Gallogly, M.M. and Mieyal, J.J. (2007). Mechanisms of reversible protein glutathionylation in redox signaling and oxidative stress. *Curr. Opin. Pharm*. **7**, 381-391.

Guex, N. and Peitsch, M.C. (1997). SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. *Electrophoresis*. **18**, 2714-2723.

Gianni, S., Guydosh, N.R., Khan, F., Caldas, T.D., Mayor, U., White, G.W., DeMarco, M.L., Daggett, V. and Fersht, A.R. (2003). Unifying features in protein-folding mechanisms. *Proc. Natl. Acad. Sci. USA*. **100**, 13286–13291.

Gildenhuys, S., Wallace, L. A. and Dirr, H. W. (2008). Stability and unfolding of reduced *E. coli* glutaredoxin 2: a monomeric structural homologue of the glutathione transferase family. *Biochemistry*. **47**, 10801-8.

Gildenhuys, S., Dobreva, M., Kinsley, N., Sayed, Y., Burke, J., Pelly, S., Gordon, G.P., Sayed, M., Sewell, T. and Dirr, H.W. (2010a). Arginine 15 stabilizes an S_NAr reaction transition state and the binding of anionic ligands at the active site of human glutathione transferase A1-1. *Biophys. Chem*. **146**, 118-125.

Gildenhuys, S., Wallace, L.A., Burke, J.P., Balchin, D., Sayed, Y. and Dirr, H.W. (2010b). Class Pi glutathione transferase unfolds via a dimeric and not monomeric intermediate: Functional implications for an unstable monomer. *Biochemistry*. **49**, 5074-5081.

Gleason, F. K. (1992) Mutation of conserved residues in *Escherichia coli* thioredoxin: Effects on stability and function. *Protein Sci*. **1**, 609–616.

Goodsell, D.S. and Olson, A.J., (2000). Structural symmetry and protein function. *Annu. Rev. Biophys. Biomol. Struct*. **29**, 105–153.

Gromer, S., Urig, S. and Becker, K. (2004). The thioredoxin system--from science to clinic. *Med. Res. Rev*. **24**, 40-89.

Gromiha, M.M. and Selvaraj, S. (2004). Inter-residue interactions in protein folding and stability. *Prog. Biophys. Mol. Biol*. **86**, 235-77.

Gunasekaran, K., Eyles, S. J., Hagler, A. T. and Gierasch, L. M. (2001). Keeping it in the family: folding studies of related proteins. *Curr. Opin. Struct. Biol.* **11**, 83-93.

Guharoy, M. and Chakrabarti, P. (2005). Conservation and relative importance of residues across protein-protein interfaces. *Proc. Nat. Acad. Sci. USA*. **102**, 15447-15452.

Guharoy, M. and Chakrabarti, P. (2010). Conserved residue clusters at protein-protein interfaces and their use in binding site identification. *BMC Bioinformatics*. **11**, art. no. 286.

Habeeb, A. F., Schrohenloher, R. E. and Bennett, J. C. (1972). Studies of the component chains of human IgM by citraconylation. *Biochim. Biophys. Acta.* **263**, 339-50.

Habig, W. H. and Jakoby, W. B. (1981). Assays for differentiation of glutathione S-transferases. *Method Enzymol.* **77**, 398-405.

Han, J. H., Batey, S., Nickson, A. A., Teichmann, S. A. and Clarke, J. (2007). The folding and evolution of multidomain proteins. *Nat. Rev. Mol. Cell. Biol*. **8**, 319–330.

Hartl, F. U., Hlodan, R. and Langer, T. (1994). Molecular chaperones in protein folding: the art of avoiding sticky situations. *Trend. Biochem. Sci.* **19**, 20-5.

Hartl, F. U. and Hayer-Hartl, M. (2009).Converging concepts of protein folding in vitro and in vivo. *Nat. Struct. Mol. Biol.* **16** (6), 574-581.

Harrop, S. J., DeMaere, M. Z., Fairlie, W. D., Reztsova, T., Valenzuela, S. M., Mazzanti, M., Tonini, R., Qiu, M. R., Jankova, L., Warton, K., Bauskin, A. R., Wu, W. M., Pankhurst, S., Campbell, T. J., Breit, S. N. and Curmi, P. M. (2001). Crystal structure of a soluble form of the intracellular chloride ion channel CLIC1 (NCC27) at 1.4-A resolution. *J. Biol. Chem.* **276**, 44993-5000.

Hawe, A., Sutter, M. and Jiskoot, W. (2008). Extrinsic fluorescent dyes as tools for protein characterization. *Pharm Res.* **25**, 1487-99.

Hayes, J. D. and Pulford, D. J. (1995). The glutathione S-transferase supergene family: regulation of GST and the contribution of the isoenzymes to cancer chemoprotection and drug resistance. *Crit. Rev. Biochem. Mol. Biol.* **30**, 445-600.

Hegazy, U. M., Mannervik, B. and Stenberg, G. (2004). Functional Role of the Lock and Key Motif at the Subunit Interface of Glutathione Transferase P1-1. *J. Biol. Chem*. **279**, 9586-9596.

Hellinga, H. W., Wynn, R. and Richards, F. M. (1992). The hydrophobic core of *Escherichia coli* thioredoxin shows a high tolerance to nonconservative single amino acid substitutions. *Biochemistry*. **31**, 11203-11209.

Honaker, M.T., Acchione, M., Sumida, J.P., Atkins, W.M. (2011). Ensemble perspective for catalytic promiscuity: calorimetric analysis of the active site conformational landscape of a detoxification enzyme. *J. Biol. Chem*. **286**, 42770-6.

Hornby, J. A., Luo, J. K., Stevens, J. M., Wallace, L. A., Kaplan, W., Armstrong, R. N. and Dirr, H. W. (2000). Equilibrium folding of dimeric class mu glutathione transferases involves a stable monomeric intermediate. *Biochemistry*. **39**, 12336-44.

Hornby, J. A., Codreanu, S. G., Armstrong, R. N. and Dirr, H. W. (2002). Molecular recognition at the dimer interface of a class mu glutathione transferase: role of a hydrophobic interaction motif in dimer stability and protein function. *Biochemistry*. **41**, 14238-47.

Horwick, A. (2002). Protein aggregation in disease: a role for folding intermediates forming specific multimeric interactions. *J. Clin. Invest.* **11,** 1221 – 32.

Jeppesen, M. G, Ortiz, P., Shepard, W., Kinzy, T. G., Nyborg, J., Andersen, G. R. (2003). The crystal structure of the glutathione S-transferase-like domain of elongation factor 1Bgamma from *Saccharomyces cerevisiae*. *J. Biol. Chem*. **278**, 47190-8.

Ji, X., Zhang, P., Armstrong, R. N. and Gilliland, G. L. (1992). The three-dimensional structure of a glutathione S-transferase from the Mu gene class. Structural analysis of the binary complex of isoenzyme 3-3 and glutathione at 2.2-Å resolution. *Biochemistry*. **31**, 10169-10184.

Kaplan, W., Husler, P., Klump, H., Erhardt, J., Sluis-Cremer, N. and Dirr, H. (1997). Conformational stability of pGEX-expressed *Schistosoma japonicum* glutathione S-transferase: a detoxification enzyme and fusion-protein affinity tag. *Protein Sci.* **6**, 399-406.

Keil, B. (1992). Specificity of proteolysis. Springer-Verlag, New York, USA.

Kellis, J. T., Jr., Nyberg, K., Sali, D. and Fersht, A. R. (1988). Contribution of hydrophobic interactions to protein stability. *Nature* **333**, 784-6.

Kelly, S. M., Jess, T. J. and Price, N. C. (2005). How to study proteins by circular dichroism (2005). *BBA - Proteins Proteom.* **1751**, 119-139.

Keskin, O., Ma, B. and Nussinov, R. (2005). Hot regions in protein--protein interactions: the organisation and contribution of structurally conserved hot spot residues. *J. Mol. Biol.* **345**, 1281-94.

Khoza, T. (2013). The N-subdomain of the thioredoxin fold of glutathione transferase is stabilised by topologically conserved leucine residues. PhD Thesis. University of Witwatersrand. [http://wiredspace.wits.ac.za/handle/10539/12700].

Kinsley, N., Sayed, Y., Mosebi, S., Armstrong, R. N. and Dirr, H. W. (2008). Characterization of the binding of 8-anilinonaphthalene sulfonate to rat class Mu GST M1-1. *Biophys. Chem.* **137**, 100-4.

Kuhnert, D. C., Sayed, Y., Mosebi, S., Sayed, M., Sewell, T. and Dirr, H. W. (2005). Tertiary interactions stabilise the C-terminal region of human glutathione transferase A1-1: a crystallographic and calorimetric study. *J. Mol. Biol.* **349**, 825-38.

Ladner, J. E., Parsons, J. F., Rife, C. L., Gilliland, G. L. and Armstrong, R. N. (2004). Parallel evolutionary pathways for glutathione transferases: structure and mechanism of the mitochondrial class kappa enzyme rGSTK1-1. *Biochemistry.* **43**, 352-61.

Lakowicz, J. R. (2006) Principles of fluorescence spectroscopy, Protein fluorescence, pp 445-465. Plenum Publishers, USA.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature.* **227**, 680-5.

Laskowski, R. A, Swindells, M. B. (2011). LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. *J. Chem. Inf. Model.* **51**, 2778-2786.

Laurent, T.C. Moore, E.C. and Reichard, P. (1964). Enzymatic synthesis of deoxyribonucleotides. IV. Isolation and characterisation of thioredoxin, the hydrogen donor from *Escherichia coli. J. Biol. Chem*. **239**, 3436-44.

Le Trong, I., Stenkamp, R. E., Ibarra, C., Atkins, W. M. and Adman, E. T. (2002). 1.3-A resolution structure of human glutathione S-transferase with S-hexyl glutathione bound reveals possible extended ligandin binding site. *Proteins.* **48**, 618-27.

Lian, H.-Y., Jiang, Y., Zhang, H., Jones, G.W., Perrett, S. (2006). The yeast prion protein Ure2: Structure, function and folding. *Biochimica et Biophysica Acta - Proteins and Proteomics*. **1764**, 535-545.

Lindorff-Larsen, K., Rogen, P., Paci, E., Vendruscolo, M. and Dobson, C. M. (2005). Protein folding and the organization of the protein topology universe. *Trend Biochem. Sci*. **30**, 13–19.

Lindberg, M. O. and Oliveberg, M. (2007). Malleability of protein folding pathways: a simple reason for complex behaviour. *Curr. Opin. Struct. Biol.* **17**, 21-9.

Line, K., Isupov, M.N., Lacourse, J., Brophy, P.M. and Littlechild, J.A. (2010). *Fasciola hepatica* sigma class GST. PDB Deposition [2WB9].

Listowsky, I. (2005). Proposed intracellular regulatory functions of glutathione transferases by recognition and binding to *S*-glutathiolated proteins. *J. Pept. Res*. **65**, 42-6.

Littler, D.R., Harrop, S.J., Fairlie, W.D., Brown, L.J., Pankhurst, G.J., Pankhurst, S., DeMaere, M.Z., Campbell, T.J., Bauskin, A.R., Tonini, R., Mazzanti, M., Breit, S.N. and Curmi, P.M. (2004). The intracellular chloride ion channel protein CLIC1 undergoes a redox-controlled structural transition. *J Biol Chem.* **279**, 9298-305.

Luo, J. K., Hornby, J. A., Wallace, L. A., Chen, J., Armstrong, R. N. and Dirr, H. W. (2002). Impact of domain interchange on conformational stability and equilibrium folding of chimeric class micro glutathione transferases. *Protein Sci.* **11**, 2208-17.

Lyon, R. P. and Atkins, W. M. (2002). Kinetic characterisation of native and cysteine 112-modified glutathione S-transferase A1-1: reassessment of nonsubstrate ligand binding. *Biochemistry.* **41**, 10920-7.

Maiti, R., van Domselaar, G.H, Zhang, H. and Wishart. D.S. (2004). SuperPose: a simple server for sophisticated structural superposition. *Nucleic Acids Res.* **32**, W590-594.

Mannervik, B. and Danielson, U. H. (1988). Glutathione transferases--structure and catalytic activity. *CRC Crit. Rev. Biochem.* **23**, 283-337.

Marqusee, S. and Baldwin, R.L. (1987). Helix stabilization by Glu-...Lys+ salt bridges in short peptides of de novo design. *Proc. Natl. Acad. Sci.U.S.A.* **84,** 8898-8902.

Martin, J.L. (1995). Thioredoxin - A fold for all reasons. *Structure*. **3**, 245-250.

Matthews, B.W. (1991). Mutational analysis of protein stability. *Curr. Opin. Struct. Biol.* **1**,17-21.

Meister, A. and Tate, S. S. (1976). Glutathione and related gamma-glutamyl compounds: biosynthesis and utilization. *Ann. Rev. Biochem*. **45**, 559-604.

Meister, A. and Anderson, M. E. (1983). Glutathione. *Annu. Rev. Biochem*. **52**, 711-60.

Moreau, K. L. and King, J. (2009). Hydrophobic core mutations associated with cataract development in mice destabilize human γD-crystallin *J. Biol. Chem*. **284**, 33285-33295.

Mosebi, S., Sayed, Y., Burke, J. and Dirr, H. W. (2003). Residue 219 impacts on the dynamics of the C-terminal region in glutathione transferase A1-1: implications for stability and catalytic and ligandin functions. *Biochemistry*. **42**, 15326-32.

Mottonen, J. M., Xu, M., Jacobs, D. J. and Livesay, D. R. (2009). Unifying mechanical and thermodynamic descriptions across the thioredoxin protein family. *Proteins*. **75**, 610-627.

Murshudov, G.N., Vagin, A.A. and Dodson, E.J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr*. **53**, 240-255.

Murzin, A. G., Brenner, S. E., Hubbard, T. and Chothia, C. (1995). SCOP: a structural classification of proteins database for the investigation of sequences and structures. *J. Mol. Biol.* **247**, 536-40.

Myers, J.K., Pace, C.N. and Scholtz, J.M. (1995). Denaturant *m*-values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding. *Protein Sci.* **4,** 2138-2148.

Nakamura, T., Nakamura, H., Hoshino, T., Ueda, S., Wada, H. and Yodoi, J. (2005). Redox regulation of lung inflammation by thioredoxin. *Antioxid. Redox Signal*. **7**, 60-71.

Nath, A. and Atkins, W. M. (2008). A quantitative index of substrate promiscuity. *Biochemistry*. **47**, 157-166.

Nathaniel, C., Wallace, L. A., Burke, J. and Dirr, H. W. (2003). The role of an evolutionarily conserved *cis*-proline in the thioredoxin-like domain of human class Alpha glutathione transferase A1-1. *Biochem. J.* **372**, 241-6.

Nilsson, L.O., Edalat, M., Pettersson, P.L., Mannervik, B. (2002). Aromatic residues in the C-terminal region of glutathione transferase A1-1 influence rate-determining steps in the catalytic mechanism. *Biochim Biophys Acta*. **1598,** 199-205.

Nordlund, P. and Reichard P. (2006) Ribonucleotide reductases. *Annu Rev Biochem*. **75**, 681-70.

Oakley, A. J., Harnnoi, T., Udomsinprasert, R., Jirajaroenrat K., Ketterman, A. J., and Wilce, M.C.J. (2001). The crystal structures of glutathione S-transferases isozymes 1–3 and 1–4 from *Anopheles dirus* species B. *Protein Sci*. **10**, 2176-2185.

Ohgushi, M. and Wada, A. (1983). 'Molten-globule state': a compact form of globular proteins with mobile side-chains. *FEBS Lett.* **164**, 21-4.

Pace, C. N. (1986). Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Method. Enzymol.* **131**, 266-80.

Pace, C. N. (2001). Polar group burial contributes more to protein stability than nonpolar group burial. *Biochemistry.* **40**, 310-3.

Pace, C. N., Shirley, B. A., McNutt, M. and Gajiwala, K. (1996). Forces contributing to the conformational stability of proteins. *Faseb J.* **10**, 75-83.

Pace, C. N., Fu, H., Fryar, K. L., Landua, J., Trevino, S. R., Shirley, B. A., Hendricks, M. M., Iimura, S., Gajiwala, K., Scholtz, J. M. and Grimsley, G. R. (2009) Contribution of hydrophobic interactions to protein stability. *J. Mol. Biol*. **408**, 514-528.

Pan, J.L., and Bardwell, J.C. (2006). The origami of thioredoxin-like folds. *Protein Sci*. **15,** 2217-2227.

Park, C. and Marqusee, S. (2005). Pulse proteolysis: A simple method for quantitative determination of protein stability and ligand binding. *Nat. Meth*. **2**, 207-212.

Parbhoo, N., Stoychev, S. H., Fanucchi, S., Achilonu, I., Adamson, R. J., Fernandes, M., Gildenhuys, S., Dirr, H. W. (2011). A conserved interdomain interaction is a determinant of folding cooperativity in the GST fold. *Biochemistry.* **50**, 7067-7075.

Pedone, E., Limauro, D., D'Ambrosio, K., De Simone, G. and Bartolucci, S. (2010). Multiple catalytically active thioredoxin folds: A winning strategy for many functions. *Cell. Mol. Life Sci.* **67**, 3797-3814.

Perbandt, M., Höppner, J., Betzel, C., Walter, R. D. and Liebau, E. (2005). Structure of the major cytosolic glutathione S-transferase from the parasitic nematode *Onchocerca volvulus*. *J. Biol. Chem*. **280**,12630-6.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C. and Ferrin, T.E. (2004). UCSF Chimera-a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605-12.

Powis, G., Mustacich, D. and Coon, A. (2000). The role of the redox protein thioredoxin in cell growth and cancer. *Free Radic. Biol. Med.* **29**, 312-322.

Privalov, P. L. (1996). Intermediate states in protein folding. *J. Mol. Biol.* **258**, 707-25.

Privalov, P.L., Mateo, P.L., Khechinashvili, N.N., Stepanov, V.M. and Revina, L.P. (1981). Comparative thermodynamic study of pepsinogen and pepsin structure. *J. Mol. Biol*. **152**, 445-64.

Ptitsyn, O.B. (1995) Structures of folding intermediates. *Curr. Opin. Str. Biol.* **5,** 74-78.

Ramagopal, U.A., Toro, R., Burley, S.K., Almo, S.C. (2010). Structure of stringent starvation protein A homolog from *Haemophilus influenza*. PDB Deposition [3LYK].

Ratnaparkhi, G. S. and Varadarajan, R. (2000). Thermodynamic and structural studies of cavity formation in proteins suggest that loss of packing interactions rather than the hydrophobic effect dominates the observed energetics. *Biochemistry*. **39**, 12365-12374.

Ravi, D., Muniyappa, H., Das, K. C. (2005). Endogenous thioredoxin is required for redox cycling of anthracyclines and p53-dependent apoptosis in cancer cells. *J. Biol. Chem*. **280**, 40084-40096.

Reinemer, P., Prade, L., Hof, P., Neuefeind, T., Huber, R., Zettl, R., Palme, K., Schell, J., Koelln, I., Bartunik, H. D. and Bieseler B. (1996). Three-dimensional structure of glutathione S-transferase from Arabidopsis thaliana at 2.2 Å resolution: structural characterization of herbicide-conjugating plant glutathione S-transferases and a novel active site architecture. *J. Mol. Biol*. **255**, 289-309.

Rossjohn, J., McKinstry, W. J., Oakley, A. J., Verger, D., Flanagan, J., Chelvanayagam, G., Tan, K. L., Board, P. G. and Parker, M. W. (1998). Human theta class glutathione transferase: the crystal structure reveals a sulfate-binding pocket within a buried active site. *Structure.* **6**, 309-22.

Rumfeldt, J. A., Galvagnion, C., Vassall, K. A. and Meiering, E. M. (2008). Conformational stability and folding mechanisms of dimeric proteins. *Prog. Biophys. Mol. Biol.* **98**, 61-84.

Sacchetta, P., Aceto, A., Bucciarelli, T., Dragani, B., Santarone, S., Allocati, N. and Di Ilio, C. (1993). Multiphasic denaturation of glutathione transferase B1-1 by guanidinium chloride. Role of the dimeric structure on the flexibility of the active site. *Eur. J. Biochem*. **215**, 741-745.

Sancho, J. (2013). The stability of 2-state, 3-state and more-state proteins from simple spectroscopic techniques... plus the structure of the equilibrium intermediates at the same time. *Arch. Biochem. Biophys*. **531**, 4-13.

Sancho, J., Bueno, M., Campos, L.A., Fernández-Recio, J., Irún, M.P., López, J., Machicado, C., Pedroso, I. and Toja, M. (2002). The 'relevant' stability of proteins with equilibrium intermediates. *ScientificWorldJournal*. **2**, 1209-15.

Santos, J., Marino-Buslje, C., Kleinman, C., Ermácora, M. R. and Delfino, J. M. (2007). Consolidation of the thioredoxin fold by peptide recognition: Interaction between *Escherichia coli* thioredoxin fragments 1-93 and 94-108. *Biochemistry*. **46**, 5148-5159.

Santos, J., Sica, M. P., Buslje, C. M., Garrote, A. M., Ermácora, M. R. and Delfino, J.M. (2009). Structural selection of a native fold by peptide recognition. Insights into the thioredoxin folding mechanism. *Biochemistry*. **48,** 595-607.

Santucci, R., Sinibaldi, F. and Fiorucci, L. (2008). Protein folding, unfolding and misfolding: Role played by intermediate states. *Mini Rev. Med. Chem.* **8**, 57-62.

Sayed, Y., Wallace, L. A. and Dirr, H. W. (2000). The hydrophobic lock-and-key intersubunit motif of glutathione transferase A1-1: implications for catalysis, ligandin function and stability. *FEBS Lett.* **465**, 169-72.

Semisotnov, G. V., Rodionova, N. A., Razgulyaev, O. I., Uversky, V. N., Gripas, A. F. and Gilmanshin, R. I. (1991). Study of the "molten globule" intermediate state in protein folding by a hydrophobic fluorescent probe. *Biopolymers* **31**, 119-28.

Sheehan, D., Meade, G., Foley, V. M. and Dowd, C. A. (2001). Structure, function and evolution of glutathione transferases: Implications for classification of non-mammalian members of an ancient enzyme superfamily *Biochem. J.* **360**, 1-16

Shortle, D. (1992). Mutational studies of protein structures and their stabilities. *Q. Rev. Biophys.* **25,** 205-250.

Slavik, J. (1982). Anilinonaphthalene sulfonate as a probe of membrane composition and function. *Biochim. Biophys. Acta.* **694,** 1-25.

Sluis-Cremer, N., Naidoo, N. N., Kaplan, W. H., Manoharan, T. H., Fahl, W. E. and Dirr, H. W. (1996). Determination of a binding site for a non-substrate ligand in mammalian cytosolic glutathione S-transferases by means of fluorescence-resonance energy transfer. *Eur. J. Biochem.* **241**, 484-8.

Sobolev V., Eyal E., Gerzon S., Potapov V., Babor M., Prilusky J. and Edelman M. (2005) SPACE: a suite of tools for protein structure prediction and analysis based n complementarity and environment. *Nucl. Acids Res*. **33**, W39-W43.

Stenberg, G., Bjornestedt, R. and Mannervik, B. (1992). Heterologous expression of recombinant human glutathione transferase A1-1 from a hepatoma cell line. *Protein Expr. Purif.* **3**, 80-4.

Stenberg, G., Dragani, B., Cocco, R., Mannervik, B. and Aceto, A. (2000). A conserved "hydrophobic staple motif" plays a crucial role in the refolding of human glutathione transferase P1-1. *J. Biol. Chem.* **275**, 10421-8.

Stevens, J. M., Hornby, J. A., Armstrong, R. N. and Dirr, H. W. (1998). Class sigma glutathione transferase unfolds via a dimeric and a monomeric intermediate: impact of subunit interface on conformational stability in the superfamily. *Biochemistry.* **37**, 15534-41.

Stevens, J. M., Armstrong, R. N. and Dirr, H. W. (2000). Electrostatic interactions affecting the active site of class sigma glutathione S-transferase. *Biochem J* **347 Pt 1**, 193-7.

Strickland, E.H. (1974) Aromatic contributions to circular dichroism spectra of proteins. *Crit. Rev. Biochem.* **2,** 113-75.

Stoychev S.H., Nathaniel, C., Fanucchi, S., Brock, M., Li, S., Asmus, K., Woods, V.L. Jr. and Dirr, H.W. (2009). Structural dynamics of soluble chloride intracellular channel protein CLIC1 examined by amide hydrogen-deuterium exchange mass spectrometry. *Biochemistry*. **48**, 8413-21.

Tanford, C. (1968). Protein denaturation. *Adv. Protein. Chem.* **23**, 121-282.

Tanford, C. (1970). Protein denaturation. C. Theoretical models for the mechanism of denaturation. *Adv. Protein. Chem.* **24**, 1-95.

Tars, K., Larsson, A. K., Shokeer, A., Olin, B., Mannervik, B. and Kleywegt G. J. (2006). Structural basis of the suppressed catalytic activity of wild-type human glutathione transferase T1-1 compared to its W234R mutant. *J. Mol. Biol*. **355**, 96-105.

Tasayco, M.L., Chao, K. (1995). NMR study of the reconstitution of the β-sheet of thioredoxin by fragment complementation. *Proteins*. **22,** 41-44.

Tasayco, M.L., Fuchs, J., Yang, X.-M., Dyalram, D., Georgescu, R.E. (2000). Interaction between two discontiguous chain segments from the β-sheet of *Escherichia coli* thioredoxin suggests an initiation site for folding. *Biochemistry* **39**, 10613-10618.

Teale, F. W. J. (1960). The ultraviolet fluorescence of proteins in neutral solution. *Biochem. J.* **76,** 381-8.

Thannhauser, T. W., Konishi, Y. and Scheraga, H. A. (1984). Sensitive quantitative analysis of disulfide bonds in polypeptides and proteins. *Anal. Biochem.* **138**, 181-8.

Thom, R., Dixon, D. P., Edwards, R., Cole, D. J. and Lapthorn A. J. (2001). The structure of a zeta class glutathione S-transferase from Arabidopsis thaliana: characterisation of a GST with novel active-site architecture and a putative role in tyrosine catabolism. *J. Mol. Biol*. **308**, 949-62.

Thom, R, Cummins, I, Dixon, D.P., Edwards, R, Cole, D. J, Lapthorn, A. J. (2002). Structure of a tau class glutathione S-transferase from wheat active in herbicide detoxification. *Biochemistry*. **41,** 7008-20.

Thompson, L. C., Walters, J., Burke, J., Parsons, J. F., Armstrong, R. N., Dirr, H. W. (2006). Double mutation at the subunit interface of glutathione transferase rGSTM1-1 results in a stable, folded monomer. *Biochemistry.* **45**, 2267-2273.

Tsytlonok, M. and Itzhaki, L.S. (2013). The how's and why's of protein folding intermediates. *Arch. Biochem. Biophys*. **531**, 14-23.

Vagin, A., and Teplyakov A. (2000). An approach to multi-copy search in molecular replacement. *Acta Crystallogr. D Biol. Crystallogr*. **56,** 1622-1624.

Vargo, M.A., Nguyen, L. and Colman, R.F. (2004). Subunit interface residues of glutathione S-transferase A1-1 that are important in the monomer-dimer equilibrium. *Biochemistry*. **43**, 3327-3335.

Vararattanavech, A., P. Prommeenate, and A.J. Ketterman. (2006). The structural roles of a conserved small hydrophobic core in the active site and an ionic bridge in domain I of Delta class glutathione S-transferase. *Biochem. .J*. **393**, 89-95.

Vreuls, C., Filee, P., Van Melckebeke, H., Aerts, T., De Deyn, P., Llabres, G., Matagne, A., Simorre, J. P., Frere, J. M. and Joris, B. (2004). Guanidinium chloride denaturation of the dimeric *Bacillus licheniformis* BlaI repressor highlights an independent domain unfolding pathway. *Biochem. J.* **384**, 179-90.

Wallace, L. A. and Dirr, H. W. (1999). Folding and assembly of dimeric human glutathione transferase A1-1. *Biochemistry.* **38**, 16686-94.

Wallace, L. A., Blatch, G. L. and Dirr, H. W. (1998). A topologically conserved aliphatic residue in α-helix 6 stabilizes the hydrophobic core in domain II of glutathione transferases and is a structural determinant for the unfolding pathway. *Biochem. J.* **336** (Pt 2), 413-8.

Wallace, L. A., Burke, J. and Dirr, H. W. (2000). Domain-domain interface packing at conserved Trp-20 in class Alpha glutathione transferase impacts on protein stability. *Biochim. Biophys. Acta.* **1478**, 325-32.

Wang, T., Arifoglu, P., Ronai, Z. and Tew, K. D. (2001). Glutathione S-transferase P1-1 (GSTP1-1) inhibits c-Jun N-terminal kinase (JNK1) signaling through interaction with the C terminus. *J. Biol. Chem.* **276**, 20999-1003.

Weber, G. and Young, L. B. (1964). Fragmentation of Bovine Serum Albumin by Pepsin. I. the Origin of the Acid Expansion of the Albumin Molecule. *J. Biol. Chem.* **239**, 1415-23.

Wilce, M. C. and Parker, M. W. (1994). Structure and function of glutathione S-transferases. *Biochim. Biophys. Acta.* **1205**, 1-18.

Winayanuwattikun, P. and Ketterman, A.J. (2004). Catalytic and structural contributions for glutathione-binding residues in a Delta class glutathione S-transferase. *Biochem. J.* **382**, 751-7.

Wolynes, P.G. (2005). Energy landscapes and solved protein-folding problems. *Philos. Transact. A Math. Phys. Eng. Sci*. **363**, 453–464 (discussion: 464–7).

Wolynes, P. G., Onuchic, J. N. and Thirumalai, D. (1995). Navigating the folding routes. *Science.* **267**, 1619-20.

Wongsantichon, J. and Ketterman, A. J. (2006). An intersubunit lock-and-key 'clasp' motif in the dimer interface of Delta class glutathione transferase. *Biochem. J.* **394**, 135-44.

Woody, R. W. (1995). Circular dichroism. *Method Enzymol.* **246**, 34-71.

Xia, B., Vlamis-Gardikas, A., Holmgren, A., Wright, P. E. and Dyson, H. J. (2001). Solution structure of *Escherichia coli* glutaredoxin-2 shows similarity to mammalian glutathione-S-transferases. *J. Mol. Biol*. **310**, 907-18.

Xu, J., W.A. Baase, E. Baldwin, and B.W. Matthews. (1998). The response of T4 lysozyme to large-to small substitutions within the core and its relation to the hydrophobic effect. *Protein Sci.* **7**,158-177.

Yamamoto, K., Suzuki, M., Higashiura, A. and Nakagawa, A. (2013). Three-dimensional structure of a *Bombyx mori* Omega-class glutathione transferase. *Biochem. Biophys. Res Commun*. **438**, 588-93.

Yassin, Z., Ortiz-Salmeron, E., Garcia-Maroto, F., Baron, C. and Garcia-Fuentes, L. (2004). Implications of the ligandin binding site on the binding of non-substrate ligands to *Schistosoma japonicum*-glutathione transferase. *Biochim Biophys Acta* **1698**, 227-37.

Yon, J. M. (2001). Protein folding: a perspective for biology, medicine and biotechnology. *Braz. J. Med. Biol. Res.* **34**, 419-35.

Yu, W. F., Tung, C.-S., Wang, H. and Tasayco, M.L. (2000). NMR analysis of cleaved *Escherichia coli* thioredoxin (1-73/74-108) and its P76A variant: Cis/trans peptide isomerisation. *Protein Sci*. **9**, 20-28.

Zhang, N., Jiang, Y., Zou, J., Zhuang, S., Jin, H. and Yu, Q. (2007). Insights into unbinding mechanisms upon two mutations investigated by molecular dynamics study of GSK3β-axin complex: Role of packing hydrophobic residues. *Proteins*. **67**, 941-949.