

https://dspace.org/

The role of a conserved interdomain salt bridge on the structure, function and stability of the Y-GSTs

Robertson, Gary Jay

2013-01-29
http://hdl.handle.net/10539/12315

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



The Role of a Conserved Interdomain Salt Bridge on the Structure, Function and Stability of the Y-GSTs

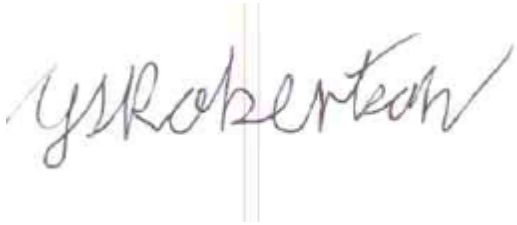
Gary Jay Robertson

A dissertation submitted to the Faculty of Science, University of the Witwatersrand, Johannesburg, in fulfilment of the requirements for the degree of Master of Science.

Johannesburg, 2012

Declaration

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



_____
Gary Jay Robertson

16th day of October, 2012

ABSTRACT

Domain interfaces are important to the folding, stability, structure and function of multidomain proteins. In the case of human glutathione *S*-transferase A1-1 (hGSTA1-1) site-directed mutagenesis studies have previously implicated the interdomain Arg13 residue of the protein in maintaining the proper catalytic function of the GST though its exact role was never determined (Stenberg *et al*., 1991). In this study it was shown by structural and sequence alignment of many representatives of the GST family and other thioredoxin-fold containing proteins that Arg13 is also highly conserved throughout the Alpha, Mu, Pi, *Plasmodium falciparum* and Sigma classes, all of which are Y-GSTs, and that it forms an interdomain salt bridge. This study therefore chose to evaluate the contribution of Arg13 towards the structure, stability and function of hGSTA1-1 by mutating the Arg residue to an Ala and performing comparative studies between wild-type and R13A hGSTA1-1. The spectral properties of R13A hGSTA1-1 monitored using far-ultraviolet circular dichroism and fluorescence indicated no significant changes in the secondary structure as compared to the native protein though fluorescence did indicate local tertiary structural changes around Trp21. Additionally, the catalytic activity of the R13A variant was reduced by 70% as compared to that of the wild-type enzyme further indicating local tertiary structural changes at and possibly near the active site which is located near the Trp21 residue. Conformational stability studies were performed by monitoring both thermal- and chemical-induced protein unfolding. The stability of the R13A variant was lower than that of the wild-type protein as revealed by a thermal-induced unfolding study which indicated that the melting point (T_m) of the R13A variant was 6 °C lower than that of the wild-type. Thermal-induced unfolding was shown not to be reversible however and the thermodynamic parameters of unfolding could not be determined. Urea-induced equilibrium unfolding studies on the other hand were reversible and displayed a variant-induced destabilisation of the conformation of the protein with a $\Delta\Delta G_{(H2O)}$ of 16.7 kJ.mol^{-1} between the mutant and native protein. Additionally urea-induced equilibrium unfolding studies in the presence of ANS indicated that the equilibrium unfolding of both wild-type and R13A hGSTA1-1 was three-state. In summary the Arg13 residue is more important to the function of the protein than it is for its global stability or structure. Also since the Arg13 residue was found to be highly conserved in all the Y-GSTs and that it forms an interdomain interaction, the residue most likely performs a similar role in each of the Y-GSTs as well.

Error never can be consistent, nor can truth fail of having support from the accurate examination of every circumstance.

— James Hutton

'Theory of the Earth', *Transactions of the Royal Society of Edinburgh* (1788), 1, 259

ACKNOWLEDGMENTS

My supervisor, Professor Heinrich W. Dirr for his support and guidance throughout my studies and for the opportunity for working in his laboratory.

My former co-supervisor, Dr. Samantha Gildenhuys, without whom none of this work would have been possible.

All of my current and past colleagues of the Protein-Structure Function Research Unit especially Nish Parbhoo and Derryn Legg E'Silva.

I would also like to thank the University of the Witwatersrand, Professor H.W. Dirr and the National Research Foundation for financial assistance.

Finally, all of my family and friends who have sustained me during this entire endeavour.

Table of Contents

ABSTRACT iii
LIST OF FIGURES ix
LIST OF TABLES xi
LIST OF ABREVIATIONS xii
RESEARCH OUTPUT xvi
CHAPTER 1. Introduction 1
1. Multidomain proteins 1
1.1. Domains and domain interfaces: defining the terms 1
1.1.1. Domains 1
1.1.2. Domain interfaces 2
1.2. The role of the domain interfaces 3
1.3. Domain interface architecture 4
1.4. Forces responsible for domain interface stability 5
1.4.1. Electrostatic interactions 6
1.4.2. Van der Waals force 7
1.4.3. Hydrogen bonding 7
1.4.4. Salt Bridge 8
1.4.5. Hydrophobic effect 9
1.5. Glutathione *S*-transferase superfamily 10
1.5.1. Classification of the glutathione *S*-transferase enzymes 11
1.5.2. Evolutionary relationship of the glutathione *S*-transferases 12
1.5.3. The structure of cytosolic glutathione *S*-transferases 13
1.5.4. Conformational stability of glutathione *S*-transferase 13
1.5.5. The active site 16
1.5.6. Catalytic mechanism 20
1.5.7. Role of the domain interface in glutathione *S*-transferase 23
1.5.8. The domain interface Arg13-Glu169 salt bridge 25
1.5.9. Objective and aims 26
Chapter 2. Experimental procedure 28
2. Materials 28
2.1. Structural alignment and characterization of domain interfaces in the GST superfamily 28
2.2. Preparation of the wild-type and mutant pKHA1 plasmids 30

2.3. Purification of the R13A mutant pKHA1 plasmid 31
2.4. Sequencing of the R13A mutant and wild-type pKHA1 plasmids 32
2.5. Preparation of R13A and wild-type hGSTA1-1 proteins 33
2.5.1. Induction studies 33
2.5.2. Protein production and purification 34
2.6. SDS-PAGE 35
2.7. SE-HPLC 37
2.8. Determination of protein concentration 37
2.8.1. Absorbance spectroscopy 37
2.8.2. Bradford assay 38
2.9. Spectroscopic structural studies of R13A and wild-type hGSTA1-1 proteins 38
2.9.1. Circular dichroism spectroscopy 38
2.9.2. Fluorescence spectroscopy 39
2.9.2.1. Intrinsic fluorescence-tryptophan fluorescence 40
2.9.2.2. Extrinsic fluorescence-ANS ligand binding 41
2.10. GSH-CDNB conjugation assay 42
2.10.1. GSH concentration determination 42
2.10.2. Specific activity 42
2.11. Stability studies of the R13A and wild-type hGSTA1-1 proteins 43
2.11.1. Thermal denaturation studies 43
2.11.2. Urea-induced equilibrium unfolding studies 44
2.11.2.1. Reversibility of unfolding 44
2.11.2.2. Urea-induced equilibrium unfolding/refolding 44
2.11.2.3. Data fitting 46
2.12. R13A hGSTA1-1 Crystallization 49
CHAPTER 3. Results 51
3. Multiple structural alignments 51
3.1. Purity and sequencing of the mutant and wild-type DNA 51
3.2. Purity of R13A and wild-type hGSTA1-1 55
3.2.1. Over expression and purification 55
3.2.2. Size and purity determination 55
3.3. Structural properties of the R13A variant and wild-type proteins 60
3.3.1. Secondary structural characterisation 60

3.3.2. Tertiary structural characterisation 60
3.3.2.1. Intrinsic fluorescence 60
3.3.2.2. Extrinsic fluorescence – ANS binding 62
3.4. Specific activity: GSH-CDNB conjugation assay 64
3.5. Conformational stability of R13A and wild-type hGSTA1-1 66
3.5.1. Thermal-induced unfolding 66
3.5.2. Chemical denaturant – induced unfolding 68
3.5.2.1. Reversibility of urea-induced unfolding 68
3.5.3. Urea-induced equilibrium unfolding in the presence and absence of ANS 71
3.6. R13A hGSTA1-1 Crystallization 75
CHAPTER 4. Discussion and conclusion 76
4.1. A conserved interdomain salt bridge in the Y-GSTs 76
4.2. The R13A mutation does not affect the global structure of the enzyme 78
4.3. Loss of the Arg13-Glu169 salt bridge compromises the active site of hGSTA1-1 78
4.4. The R13A mutation influences helix 9 dynamics 81
4.5. Role of Arg13 in the stability of hGSTA1-1 82
4.6. Conclusion 85
CHAPTER 5. References 87
Appendix 104

LIST OF FIGURES

Figure	Title	Page
1.1	A structural alignment comparison subunits from four classes of GSTs	14
1.2	The crystal structure of the homodimer hGSTA1-1	15
1.3	The crystal structure of the homodimer hGSTA1-1 highlighting the positions of the G- and H-site	17
1.4	The catalytic mechanism of hGSTA1-1	22
1.5	Representation of the Trp21 lock-and-key motif	24
2.1	Representation of a typical two-state urea denaturation curve	47
3.1	Structure based sequence alignments of 18 thioredoxin fold containing proteins	52
3.2	Comparing the structural role of the conserved Arg residue in the Alpha, Mu, Pi, *Plasmodium falciparum* and Sigma classes of GST	53
3.3	R13A mutant pKHA1 plasmid sequencing results	54
3.4	A 12% polyacrylamide SDS-PAGE gel showing the results of the induction studies performed to optimize production of R13A hGSTA1-1 protein	56
3.5	Elution profiles of the R13A variant and wild-type hGSTA1-1 proteins from purification	57
3.6	SDS-PAGE analysis of the purity and size of the R13A and wild-type hGSTA1-1 subunits	58
3.7	SE-HPLC of wild-type and R13A hGSTA1-1	59
3.8	Far-UV circular dichroism spectra of R13A and wild-type hGSTA1-1	61
3.9	Fluorescence emission spectra of wild-type and R13A variant hGSTA1-1	61
3.10	Fluorescence emission spectra of free, unbound ANS and ANS bound to the wild-type and R13A variant hGSTA1-1	63
3.11	Progress curve of the conjugation reaction of GSH with CDNB as catalysed by R13A hGSTA1-1 protein	65

3.12	Specific activity of wild-type and R13A variant hGSTA1-1 as determined using the GSH-CDNB conjugation assay	65
3.13	Thermal-induced unfolding transitions of wild-type and R13A variant hGSTA1-1	67
3.14	Recovery from unfolding of R13A monitored by fluorescence	69
3.15	Recovery from unfolding of R13A monitored by circular dichroism	69
3.16	Reversibility of unfolding of R13A hGSTA1-1 monitored by fluorescence	70
3.17	Urea-induced equilibrium unfolding of wild-type and R13A hGSTA1-1 in the presence and absence of ANS	72

LIST OF TABLES

Table	Title	Page
1	Listing the protein crystal structures used in the structure based sequence alignment, their PDB codes and the resolution they were solved at	29
2	Oligonucleotide primer sequences used for site-directed mutagenesis	30
3	Oligonucleotide primer sequences used for sequencing the ORF of the pKHA1 plasmid	32
4	Thermodynamic parameters of two-state equilibrium unfolding for both wild-type and R13A hGSTA1-1	74

LIST OF ABREVIATIONS

[Θ]	mean residue ellipticity
$[\Theta]_{xxx}$	mean residual ellipticity at xxx nm
ANS	8-anilino-1-napthalene sulfonate
A_{xxx}	absorbance at xxx nm
BLAST	basic local alignment search tool
BLASTx	nucleotide 6-frame translation-protein alignment search tool
BSP	bromosulphophthalein
CD	circular dichroism
cDNA	complementary DNA
CDNB	1-chloro-2,4-dinitrochlorobenzene
	chromatography
C_m	the denaturant concentration at the midpoint of the unfolding curve
CM	carboxymethyl
DNase	deoxyribonuclease
dNTP	deoxyribonucleotides
Dpnl	a type IIM restriction endonuclease
dsDNA	double stranded DNA
DTNB	5,5'-dithio-*bis*-(2-nitrobenzoic acid)
DTT	dithiothreitol
EDTA	ethylenediaminetetraacetic acid
Em_{xxx}	Emission of a fluorophore at xxx nm

Ex_{xxx}	Excitation of a fluorophore at xxx nm
GS^-	anion of GSH after the thiol moiety becomes deprotonated
GSH	glutathione (reduced form)
G-site	glutathione binding site
GST	glutathione *S*-transferase
HBA	hydrogen bond acceptor
HBD	hydrogen bond donor
hGSTA1-1	human Alpha class glutathione *S*-transferase
H-site	hydrophobic electrophilic substrate binding site
IPTG	isopropyl β-D-1-thiogalactopyranoside
ITC	isothermal titration calorimetry
k_{cat}	turnover number
k_{cat}/K_M	catalytic efficiency
K_d	dissociation constant
kDa	kilodalton
K_M	Michaelis-Menten constant
LB	lysogeny Broth
mdeg	millidegrees
Mr	relative molecular mass
mRNA	messenger RNA
m-value	dependence of free energy as a function of denaturant concentration
N	native conformation
N_2	folded dimer native conformation

OD_{600}	optical density at 600 nm
ORF	open reading frame
PDB	protein data bank
pK_a	acid dissociation constant
R13A	arginine to alanine mutation at residue 13
RNA	Ribonucleic acid
rpm	revolutions per minute
SCOP	structural classification of proteins
SDS	sodium dodecyl sulfate
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SE-HPLC	size exclusion-high performance liquid chromatography
T	temperature in kelvin
TEMED	N,N,N′,N′-tetramethylethylenediamine
T_m	melting point
Tris-HCL	tris(hydroxymethyl)aminomethane-Hydrochloric acid
U	unfolded monomer
UV	ultra violet
α9del	α helix 9 deletion
ΔG	Gibbs free energy
$\Delta G_{(H2O)}$	change in free energy of unfolding in the absence of denaturant
ΔH	enthalpy change
ΔS	entropy change
ε	molar absorption coefficient
ε_{xxx}	molar absorption coefficient at xxx nm

The one- and three-letter symbolism for amino acids (IUPAC-IUB, 1984) and the one-letter symbolism for deoxyribonucleic acids (IUPAC-IUB, 1974) have been used in accordance with the IUPAC-IUBMB system of nomenclature.

RESEARCH OUTPUT

1	MBRT Research Day-Oral presentation (07/12/2011) **The Role of the Domain Interface in the Stability, Structure and Function of GST A1-1 Examined by Site-Directed Mutagenesis of a Conserved Arg13** G.J. Robertson
2	SASBMB/FASBMB Conference-Poster presentation (29/01/2012-02/02/2012) **The Role of a Conserved Interdomain Salt Bridge on the Structure, Function and Stability of the GST superfamily** G.J. Robertson and H.W. Dirr

CHAPTER 1. Introduction

1. Multidomain proteins

The structure of a protein determines its biological function (known as the structure-to-function relationship). However, the structure of a protein itself is the result of a folding process and must be stabilized once formed by intra- and intermolecular forces. Multidomain proteins add another level of complexity to this structure-to-function relationship as any interdomain interactions must affect the folding pathway and the stability of the final product.

It is important to understand multidomain proteins since the majority of proteins fall into this category. In fact, most natural proteins encompass more than a single domain. It is estimated more than 50% of prokaryotic and archaeal proteins and more than 70% of eukaryotic proteins fall into this category (Apic *et al*., 2001; Han *et al*., 2007). Of these multidomain proteins, 95% have between 2 and 5 domains (Han *et al*., 2007).

1.1. Domains and domain interfaces: defining the terms

1.1.1. Domains

Protein scientists have not yet agreed on a universal definition for a protein domain. Instead, definitions vary with context, focusing on protein sequence, structural aspects, protein stability or protein function. (Jaenicke, 1999).

As this study will investigate the effect of domain-domain interactions on the structure, function and stability of glutathione S-transferase, it remains critical to derive an unambiguous definition of a protein domain and hence of the interfaces lying between spatially neighbouring domains. Domains have been described as stable units of protein structure that fold autonomously, thereby playing a central role as intermediates during folding (Wetlaufer, 1973). This definition is not fully comprehensive however, as some domains have been shown not to fold autonomously. Domains have also been described as co-operative thermodynamic units, detectable by distinct folding and unfolding transitions that are distinguishable by hydrodynamic and spectroscopic measurements (Privalov, 1979). Even this definition is not complete, as a protein could consist of multiple but identical

domains which might or might not be detectable by distinct folding and unfolding transitions. Another definition describes domains as compact, local, semi-independent structural units (Richardson, 1981). Looking at these definitions in context with the others we see they are often incomplete and misleadingly inaccurate.

For the purpose of this thesis the following definition is proposed based on the definitions discussed above: a protein domain would be a compact, semi-independent substructure of a protein that folds co-operatively and sometimes autonomously.

1.1.2. Domain interfaces

The definition of a protein domain applied to multidomain proteins implies that domains either exist as independent units or that they interact with each other to some degree through a shared interface involving specific atomic interactions between discrete amino acid residues.

Therefore, a domain interface may be defined as the surface area buried upon domain association as well as the specific contacts formed between an interacting pair of domains (Jones *et al*., 2000; Orengo *et al*., 1997). Analysis of dimer interfaces has shown that proteins often utilise residues with large hydrophobic side chains to anchor domains together by having the hydrophobic side chain of one domain inserted into hydrophobic pockets in another, often referred to as lock-and-key motifs (Jones and Thornton, 1995). It has also been shown by analysis that domain interfaces closely resemble protein surfaces by amino acid composition but are significantly different to protein interiors (Jones *et al*., 2000). Specifically evidence shows that domain interfaces contain charged and polar amino acid residues at levels more common to protein surfaces. The amino acid compositions of domain interfaces are remarkably similar to those of subunit interfaces and in general the two types of interface are very similar (Jones and Thornton, 1996; Jones *et al*., 2000). Arg was found to be common to both types of interface, probably due to its ability to form intermolecular hydrogen bonds or salt bridges (Jones *et al*., 2000).

1.2. The role of the domain interfaces

Since most natural proteins encompass more than a single domain it is logical to conclude multidomain proteins have an evolutionary advantage. Indeed multidomain proteins have several advantages over their single domain counterparts. Multidomain proteins show enhanced stability upon the association of multiple domains (Jaenicke, 1999) and new functionality often arises at the resulting domain interfaces (Larsen *et al*., 1998; Huang *et al*., 2008). Additionally, the organisation of proteins by structural domains represents an advantage for protein folding, with each domain being able to fold individually or co-operatively, accelerating the folding process and reducing a potentially large combination of residue interactions, while protecting the nascent polypeptide from proteolysis (Fersht, 1998; Jaenicke, 1999). Promiscuous domain interactions have also been implicated in the evolution of multimeric proteins in a process known as 'domain swapping', whereby domains in a single monomer switch to associations between domains on different protein molecules rather than with each other (Bennet *et al*., 1994; Bennet *et al*., 1995; Schlunegger *et al*., 1997).

There is a wide range of evidence to support the above conclusions. Studies have shown that while amino acid sequences coding for a domain (normally part of a multidomain protein) may fold and form stable protein structures (Teichmann *et al*., 1999) they almost in all cases do not produce a functionally active protein individually (Yon, 2001). Further domain interface residues are thought to be under evolutionary constraint and hence conserved due to the need to maintain favourable interactions (Valdar and Thornton, 2001; Elcock and McCammon, 2001). Evolutionary constraint always indicates importance. However, domain interfaces are only slightly more conserved than other residues in a protein that are not part of a interface and/or that have no direct roles in the function of a protein (Grishin and Phillips, 1994; Caffrey *et al*., 2004). The evidence suggesting domain interfaces are under evolutionary constraint is that they have been observed to combine in only a limited number of ways (Han *et al*., 2007). N to C-terminal domain order appears to be conserved, with the same domains occurring in the same sequence in different proteins (Bashton and Chothia, 2002). Shank *et al*. (2010) proposed that this was because a changing domain topology affects the coupling and folding co-operativity between the domains and that evolution would select for those topologies which fine-tune the folding landscape such that these molecules avoid regions that may lead to kinetic trapping. Bashton and Chothia (2002) proposed it was

because multidomain proteins likely evolved by duplication after a single ancestral recombination event set the domain order.

1.3. Domain interface architecture

It is now apparent that to fully understand the folding mechanism and stability of multidomain proteins, it is necessary to not only study the individual domains of proteins but the features and characteristics of any interactions between the domains (the domain interfaces) must also be analysed. Before the interactions occurring at a domain interface may be understood it is necessary to consider the unique architecture of domain interfaces. This is because the architecture of domain interfaces is a secondary and parallel element that shapes and alters those forces that contribute to domain interfaces. Firstly bioinformatics based studies of domain interfaces have shown that the size of the interface (the size of buried accessible surface area) is proportional to the extent of interface contacts and thus to protein stability (Jones and Thornton, 1995; Jones and Thornton, 1996; Tsai *et al*., 1996; Stites, 1997 Jones *et al*., 2000). Additional factors to consider are the planarity and structural complementarity of domain interfaces. They directly influence the energy contribution of the interactions between domains and hence the intrinsic stability of multidomain proteins.

Planarity is a gauge of how far interface residues deviate from a plane, indicating how flat or how warped an interface surface is (Laskowski, 1991). Planarity values therefore give an indication of protrusions and indentations at the interfaces (Jones and Thornton 1996). This is important because proteins often anchor interfaces together by inserting hydrophobic protrusion into hydrophobic hollows (Jones and Thornton, 1995). If the fit between protrusions and hollows is tight enough that solvent is excluded electrostatic forces will be stronger due to a lower dielectric constant. Planarity values increase with increasing size of the interface-accessible surface area (Jones and Thornton, 1995). This is caused by protruding side chains from the domain interface considerably increasing the surface area buried upon domain association. Therefore proteins with large interface-accessible surface area are likely to have more lock-and-key contacts providing added stability.

Complementarity gives an indication of the suitability of the shape of interacting surfaces to fit together. In terms of domain interfaces, the average strength of interactions between

interfaces is proportional to their complementarity (Jones and Thornton, 1996). This is because forces of protein stability are stronger at close proximity and at optimal angles. This can be seen for example in the poorly correlating antibody-antigen surfaces having smaller binding constants than the more complementary dimer interfaces (Jones and Thornton, 1996). A greater complementarity therefore, points to stronger binding affinity between a pair of interacting interfaces and hence a larger contribution toward the overall stability of the protein. Evidence suggests that the conservation of interdomain geometry is the primary evolutionary constraint of domain interfaces (Aloy *et al*., 2003). Aloy *et al*., (2003) examined the relationship between the conservation of interdomain geometry and protein sequence. They found that the more conserved two or more domains sequences were the greater the likelihood they would interact with a domain with similar geometry in other proteins. Indeed, a comparison of the domain interface geometry between many pairs of homologous domains from different proteins found that in 60% of cases the second domain occupied the same position relatie to the first domain (Han *et al*., 2006). A large complementary interface is therefore one of the prime determinants of the stabilizing effect domain interfaces have on a protein (Han *et al*., 2007).

The specific contributions of domain interface architecture to the forces which stabilize domain interfaces will be discussed in greater detail alongside the specific interaction as appropriate below:

1.4. Forces responsible for domain interface stability

The overall stabilizing effect of the formation of domain interfaces, similar to protein stability, is a balance between forces. In the case of domain interfaces it is a balance between those interactions that favour domain association and those interactions that favour domain dissociation. The difference is represented by the change in Gibbs free energy ($\Delta G°$) upon total unfolding, which itself is made up of two contributions under constant pressure as seen in the Gibbs–Helmholtz equation:

$$\Delta G° = \Delta H° - T\Delta S° \qquad (1)$$

where $\Delta G°$ is the change in Gibbs free energy of domain association, $\Delta H°$ is the change in enthalpy experienced upon associations (bond formation) and $\Delta S°$ is the corresponding change in the entropy of the system (freedom of a system to explore conformational space). *T*

represents the temperature of the system in kelvin. The overall stabilizing effect induced from domain interface contacts is therefore a balance of the two contributions of ΔH and ΔS that favour domain association or domain dissociation. ΔS itself is a balance of two components: ΔS_{conf} (entropy due to conformational changes) and ΔS_{solv} (the entropic contribution of the solvent). ΔS_{conf} is typically unfavourable in the case of domain association, as the number of possible states of the system is reduced. ΔS_{solv} on the other hand is dependent on the hydrophobicity of the interacting surfaces, and is often favourable. This is due to structuring of water molecules around hydrophobic side chains an entropically unfavourable occurrence. Burial of non-polar surface area upon the formation of a domain interface however, results in a more favourable increase in the entropy of the solvent as it becomes less ordered. The effects of ΔS_{solv} are discussed in greater detail as part of the hydrophobic effect (section 1.4.5.).

Upon domain-domain association the protein assumes a more compact and rigid structure. Its potential to occupy a greater conformational space is lost and hence conformational energy is lost resulting in unfavourable entropy change that will resist domain association. On the other hand the resulting domain-domain interactions formed upon domain association favour folding by causing a negative change in enthalpy. The dominant energy contributors to enthalpy are van der Waals interactions, electrostatic interactions and the hydrophobic effect (Dill, 1990).

1.4.1. Electrostatic interactions

Often the strongest of the interactions giving off 20-40 kJ.mol^{-1}, electrostatic interactions take place between functional groups having opposite and hence, attractive charges (Patrick, 2005). The strength of the interaction is dependent on firstly the distance between the charged ions and secondly the environment. That is to say the strength of the interaction falls off inversely with distance and is stronger in hydrophobic than in hydrophilic environments since the strength of the dielectric effect is weaker. This can be seen in Coulomb's law:

$$F = (k.q_1.q_2)/Dr^2 \qquad (2)$$

where F is the force between two electrical charges q_1 and q_2 that are separated by a distance r. k is the proportionality constant ($k = 2.14 \times 10^9$ cal.m.C^{-2}, where cal is calories, m is meters and C is coulombs) and D is the dielectric constant of the medium. An increase in dielectric constant will decrease the strength of the electrostatic interactions. Hence, a more polar environment such as at the surface of a protein will result in weaker interactions. Conversely, an increasingly non-polar or hydrophobic environment such as that in a protein interior or at domain and dimer interfaces (especially ones with a high complementarity) of proteins will strengthen electrostatic interactions.

1.4.2. Van der Waals force

Van der Waals interactions, also known as London dispersion forces, are non-covalent and occur between all atoms, both polar and non-polar (Lins and Brassuer, 1995). They are very weak interactions, giving off 2-4 kJ.mol^{-1}, that are the direct result of the uneven electronic distribution in otherwise neutral non-polar regions of molecules (Patrick, 2005). These uneven and constantly shifting distributions result in transient regions of high and low electron densities, resulting in the formation of temporary dipoles. These dipoles in one molecule can induce dipoles in neighbouring molecules leading to weak interactions between the opposite charges. The van der Waals force is highly dependent on distance, decreasing proportionally to the 7th power of the distance.

Hence van der Waals interactions are short ranged and functional only at very small distances. Therefore, for successful van der Waals interactions to contribute to domain-domain association, steric complementarity of the interface is essential.

1.4.3. Hydrogen bonding

Hydrogen bonds form between polar molecules when a hydrogen atom from one molecule (the hydrogen bond donor) is shared between its own electronegative atom and an electronegative atom belonging to another molecule (the hydrogen bond acceptor). These electronegative atoms are usually oxygen, nitrogen or fluorine. The strength of a hydrogen bond may vary greatly from 8-40 kJ.mol^{-1}. This is because the strength of a hydrogen bond is dependent on the electronegativity and orientation of the bonding (Dill, 1990). Although

hydrogen bonding can be thought of as just another kind of electrostatic bond the difference is that in H bonding there is orbital overlap and a weak form of Sigma bonding is taking place. This makes hydrogen bonding very dependent on the orientation and direction of the bond and is strongest when the hydrogen bond donor and hydrogen bond acceptor are opposite each other by 180° and the strength of the bond falls off markedly below 130° (Patrick, 2005).

Furthermore, hydrogen bonds are cooperative; a hydrogen bond to a carbonyl group in a peptide backbone will strengthen a second hydrogen bond to an amide group within that same peptide bond (Stickle *et al*., 1992). Likewise the contribution of hydrogen bonds toward protein stability is further enhanced by the existence of cooperative networks of hydrogen bonding wherein a hydrogen bond acceptor or hydrogen bond donor participates in a number of hydrogen bonds simultaneously (Stickle *et al*., 1992). Hydrogen bonds may be formed by both the peptide backbone and amino acids with polar side chains due to the presence of amine, carbonyl, thiol, and hydroxyl groups. That both the peptide backbone and polar residues can form hydrogen bonds and hydrogen bonds are so sensitive to distance and orientation, highlights the need for proper domain interface complementarity and for correct packing at domain interfaces of proteins. Otherwise intradomain local hydrogen bonds (meaning they are between neighbouring or near-neighbouring amino acids in sequence) would predominate over interdomain hydrogen bonds if they were not entirely impossible. Nonetheless, in terms of domain interfaces, the majority of hydrogen bond contributions are due to polar residues, as these contacts are most likely to be non-local, while backbone hydrogen bonds are predominantly local.

Even though hydrogen bonds provide a significant contribution toward protein stability they have been proven not to be the dominant folding force. If they were the dominant force behind protein folding, solvents that form strong hydrogen bonds with a protein should unfold it, since hydrogen bonds between amino acid residues would be abandoned in favour of binding to the solvent (Dill, 1990). This behaviour however is not observed for proteins.

1.4.4. Salt Bridge

Salt bridges are a specific kind of electrostatic interaction. They are formed between the acidic and negatively charged Asp and Glu residues and the basic as well as positively

charged amino acids Arg, Lys and His. Most salt bridges are stabilizing, though there have been proteins in which destabilizing salt bridges have been reported (Kumar and Nussinov, 1999). The energy contribution of salt bridges varies between 20-60 $kJ.mol^{-1}$.ion $pair^{-1}$ according to their location in the protein, their geometry, whether they are isolated or networked and if they are hydrogen bonded or not (Kumar and Nussinov, 1999).

In spite of the large energy gain to stabilization, the number of charged residues and hence the incidence of ion pairs in the interior of proteins is relatively low compared to the number of hydrophobic residues. This means that the numbers of ion pairs at domain and dimer interfaces of proteins are also very low. This is likely due to the large energy cost needed for the transfer of a charged ion from a polar to non-polar environment (the amino acids which comprise salt bridges are the most polar of all amino acids). This energy cost known as Born energy is approximately 80 $kJ.mol^{-1}$ (Dill, 1990). As a result, salt bridges in protein interiors, domains and dimer interfaces are thought to primarily be responsible for correct packing specificity over stability. Proper packing in non-polar environments is essential for salt bridges to have favourable geometries that will allow them to compensate for the high Born energy (Kumar and Nussinov, 1999; Hendsch and Tidor, 1994). Only under these favourable packing conditions will salt bridges in a non-polar environment stabilize a protein.

1.4.5. Hydrophobic effect

Water is the solvent of life. All molecules in a living organism are surrounded by it and this has an impact on their behaviour. In regards to molecular interactions water has two important effects depending on whether the molecules involved are polar or non-polar (Patrick, 2005). Polar molecules in solution are solvated by water. Before they can interact and bind with each other, resulting in an energy gain, they must be desolvated. Desolvation results in an energy loss. If the energy loss is greater than the energy gain from intermolecular bonds it will inhibit molecular recognition. However desolvation also results in an increase of entropy as a result of freeing the water molecules into a less ordered form will promote molecular recognition. Molecular recognition between Polar molecules in an aqueous environment is therefore a balance of these two forces.

For hydrophobic or non-polar molecules the consequences of an aqueous environment is very different. Water cannot solvate hydrophobic molecules, this instead forces the water molecules to form a greater number of bonds with each other than usual. This results in an abnormally ordered arrangement of water around hydrophobic molecules, meaning a negative entropy value. When two hydrophobic molecules interact the amount of hydrophobic surface area exposed to water decreases. This frees the water to become less organised causing an increase in entropy. Since the entropy of the universe wants to be always increasing binding will be promoted.

The hydrophobic effect is so powerful it is widely considered to be the major force behind protein folding and stability. Since domain interfaces however strongly resemble that of protein surfaces and are far less hydrophobic than protein interiors the hydrophobic effect is far less important in maintaining a domain interface (Jones *et al*., 2000). Rather it has been proposed that due to the enforced proximity between domains (they are a part of a single protein chain) a weaker hydrophobic effect along with networks of weak and less specific non-covalent interactions throughout the domain interface are sufficient to stabilize the interface (Jones *et al*., 2000).

1.5. Glutathione *S*-transferase superfamily

Cytosolic glutathione *S*-transferases (EC 2.5.1.18) constitute a protein superfamily of multi-functional detoxification enzymes comprising 18 subclasses (structural classification of proteins (SCOP): http://scop.mrc-lmb.cam.ac.uk/scop/data/scop.b.b.fh.b.A.html; Murzin *et al.*, 1995) and are found in almost all aerobic organisms including bacteria, yeasts, plants, insects and vertebrates. They are absent in the Archaea domain which almost universally does not even produce the substrate of the enzyme (Copley and Dhillon, 2002).

The detoxification enzymes carry out a vital role in the cellular defence against oxidative stress, explaining their widespread appearance in aerobic organisms, (Liang *et al*., 2005) as well as many electrophilic physiological and xenobiotic substances (Wilce and Parker, 1994; Mannervik and Danielson, 1988). Specifically the enzymes are involved in Phase II of enzymatic detoxification (Sheehan *et al*., 2001). In all cases this is achieved by catalysis of the conjugation of reduced glutathione (γ-glutamylcysteinylglycine, GSH) to a variety of

hydrophobic and electrophilic substrates. The conjugation of GSH to the substrate makes it far more soluble than it would be otherwise facilitating elimination via the mercapturate pathway (Boyland and Chasseuaud, 1969; Wilce and Parker, 1994). As a consequence of their detoxification roles GSTs have been discovered to be responsible in the development of resistance in some cancers towards various electrophilic anticancer drugs (Coles and Ketterer, 1990; Waxman, 1990; Tsuchida and Sato, 1992; Hayes and Pulford, 1995).

GSTs are also known as ligandins (ligand–binding proteins) and are involved in the intracellular storage and transport of a variety of hydrophobic, non-substrate compounds such as hormones, metabolites, heme, steroids and drugs (Mannervik and Guthenberg, 1981; Listowsky *et al.*,1988; Mannervik, 1985; Danger *et al*., 1992). GSTs also have important roles in cell signalling and other processes (Adler *et al*., 1999; Cho *et al*., 2001; Johansson and Mannervik, 2001; Kanaoka *et al*., 1997).

1.5.1. Classification of the glutathione *S*-transferase enzymes

The GSTs are categorized into classes designated by the names of Greek letters (e.g., Mu, Omega, and Sigma) according to their sequence, structure, substrate specificity and immunological reactions (Sheehan, *et al*., 2001). Currently, the cytosolic mammalian GSTs are grouped into nine species-independent gene classes: Alpha, Mu, Pi (Mannervik *et al*., 1985), Sigma (Buetler and Eaton, 1992), Theta (Meyer *et al*., 1991), Kappa (Pemble *et al*., 1996), Zeta (Board *et al*., 1997), Beta (Rossjohn *et al*., 1998) and Omega (Board *et al*., 2000). Additionally there are classes of cytosolic GSTs unique to insects (Delta), plants (Phi and Tau) and even one class from and named after *Plasmodium falciparum* (Sheehan *et al*., 2001; Hiller *et al*., 2006).

Recently, Atkinson and Babbit (2009) made the case that the GST enzymes should be further subdivided into two major subgroups, namely the S/C-GSTs and the Y-GSTs. They made an observation through global analysis of sequence and structural similarity between all the cytosolic GST classes that they all bound and activated GSH for conjugation through either one of two ways; by using either a Cys or Ser residue (S/C-GSTs) or a Tyr residue (Y-GSTs). The Y-GST subgroup is populated by the Alpha, Mu, Pi and Sigma classes of GST whereas the C/S-GST subgroup is represented by the Beta, Omega, Phi, Tau, Theta and Zeta GSTs. Atkinson and Babbit (2009) found that beyond simply sharing the same catalytic residue,

members that belonged to the same major subgroup were far more closely related to each other than they were to any GSTs outside of their subgroups.

1.5.2. Evolutionary relationship of the glutathione *S*-transferases

Our knowledge and understanding of the evolution of the GST proteins continues to change as more and more GST classes and isoenzymes within those classes are found. The various classes of GST are all thought to have evolved through a divergent pathway from a common single domain protein ancestor (Armstrong, 1997; Ladner *et al*., 2004). A comparison of the DNA and gene sequences of GSTs from the Alpha, Mu and Pi classes with that of the Sigma and Theta GST classes suggest that the Alpha, Mu and Pi GSTs evolved through the duplication of the Theta class gene (Pemble and Taylor, 1992; Pemble *et al*., 1996; Armstrong, 1997). Likewise, it is thought that the Sigma class GSTs must have also diverged from the Theta class or possibly an early variant of the Alpha, Mu and Pi classes of GST (Tomarev *et al*., 1993). Furthermore, it was once thought that the Theta class of GSTs itself must have diverged and evolved from the Kappa class of GSTs (Pemble *et al*., 1996) however after the unique protein fold of the Kappa class in comparison to the other cytosolic GSTs became apparent this is no longer thought to be the case (Ladner *et al*., 2004). More recent studies have even shown that the Kappa class of GSTs shares its most recent common ancestor with the DsbA-like enzymes (protein disulfide oxidoreductases) and that they are more closely related to each other than the Kappa class is to any other GST or thioredoxin fold protein (Ladner *et al*., 2004).

Nonetheless, despite the continually changing understanding of GST evolution there is sufficient evidence to show that the Alpha, Mu, Pi and Sigma classes of GST are four of the most recently evolved and closely related GST classes (Armstrong, 1997; Ranson *et al*., 2001 Sheehan *et al*., 2001 Fonseca *et al*., 2010). The taxonomic distribution and strong sequence and structural based alignment similarity as shown in the paper by Atkinson and Babbit (2009) goes a long way to lend credence to this theory. After all, the Alpha, Mu, Pi and Sigma classes of GST were assigned to the Y-GST major subgroup.

1.5.3. The structure of cytosolic glutathione *S*-transferases

Despite their proposed common ancestry GSTs share little sequence similarity. Nevertheless, crystal structure data shows that the overall structural fold of the proteins is well conserved (Figure 1.1). Specifically, the cytosolic GSTs have a dimeric quaternary structure (M_r ~ 50 kDa) and exist as stable homo- or heterodimers with each dimer subunit further subdivided into two distinct domains (Dirr *et al*., 1994; Wilce and Parker, 1994). Subunit-subunit contacts are vital in dimer assembly and protein stability of the GSTs (Erhardt and Dirr, 1995) but the composition of the dimer interface varies between classes and the dimerization process is highly specific and occurs only between subunits within the same gene class. Domain 1 or the N-terminal domain has a mixed α/β domain topologically similar to the thioredoxin fold (a mixed β -sheet of four β-strands and three α -helices with a folding topology corresponding to βαβαββα) while domain 2 or the C-terminal domain is entirely α-helical. That is with the exception of the Kappa class, whose C-terminal domain is inserted into the N-terminal domain (Ladner *et al*., 2004).

In the case of human glutathione *S*-transferase A1-1 (hGSTA1-1) enzyme, it is known to be a homodimer, comprising two identical A1 subunits. As mentioned each subunit consists of two domains; the thioredoxin like domain 1 and the purely α-helical domain 2 which makes up two thirds of the subunit (Sinning *et al*., 1993) (Figure 1.2). A structural feature unique to class Alpha GSTs though is the presence of a disordered amphipathic helix, α-helix 9 (Le Trong *et al*., 2002) (herein after unless specifically stated otherwise the word helix means an α-helix), at the C-terminus which forms an integral part of the active site of hGSTA1-1 (Sinning *et al*., 1993; Adman *et al*., 2001; Mosebi *et al*., 2003).

1.5.4. Conformational stability of glutathione *S*-transferase

Studies using equilibrium folding revealed that dimer formation of GSTs has significant impact on stabilisation of subunit tertiary structure for classes Alpha (Wallace *et al.*, 1998), Pi (Dirr and Reinemer, 1991; Erhardt and Dirr, 1995) and Sj (Kaplan *et al*., 1997; Andujar-Sanchez *et al.*, 2004). This is in part due to the burial of hydrophobic residues during dimer formation, in the case of hGSTA1-1 14% of the proteins otherwise solvent exposed surface area is buried upon dimerization (Dirr *et al*., 1994). These classes of GST are all shown to unfold via a highly co-operative two-state pathway. Additionally the proteins were found to have stabilisation energy of the native structure of ~100 $kJ.mol^{-1}$. However, in the case of

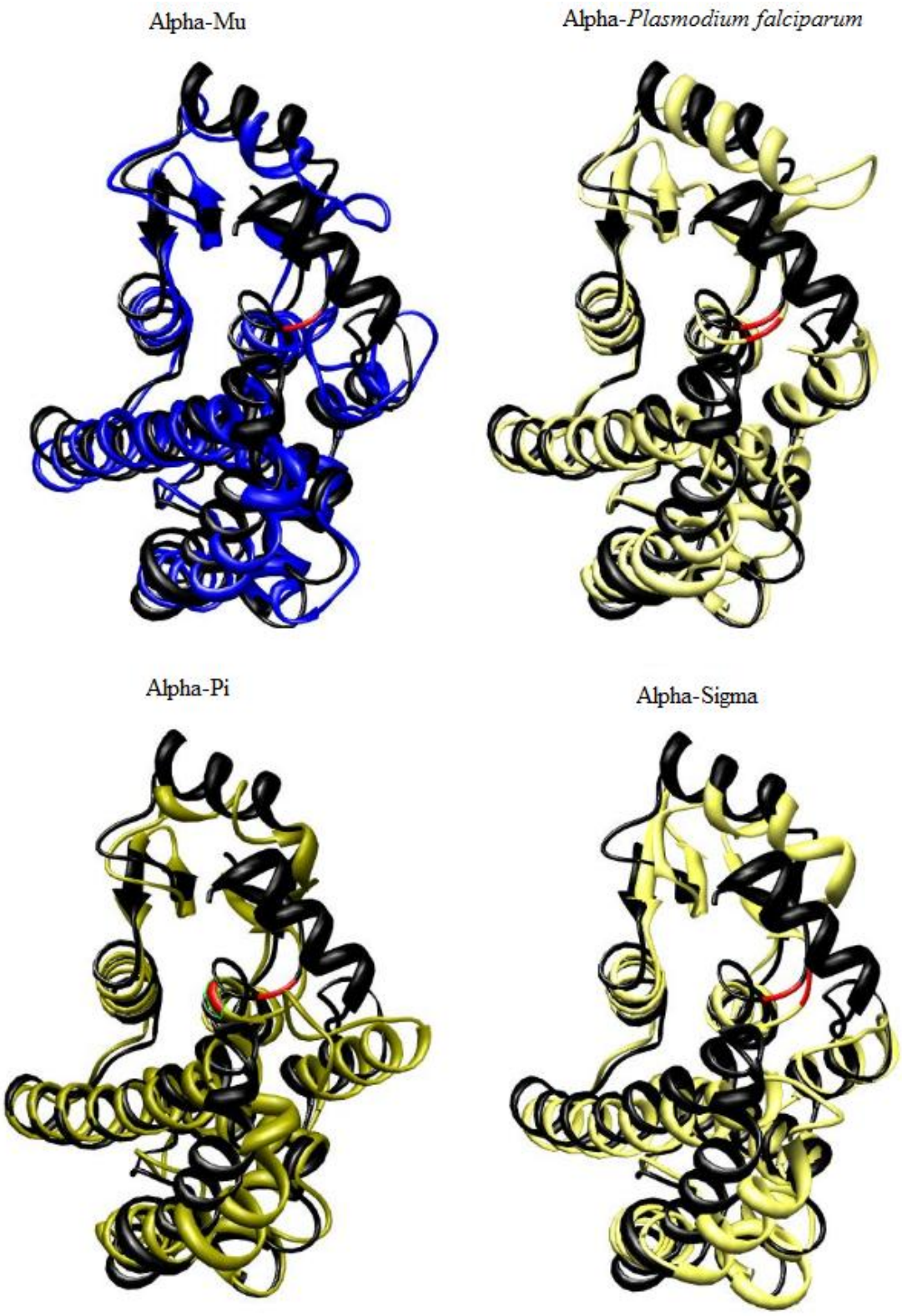


Figure 1.1: A structural alignment comparison subunits from four classes of GSTs. Despite sharing a maximum of forty percent sequence identity the structural folds of the subunits of Alpha, Pi, *Plasmodium falciparum* and Sigma class GSTs can be seen to have very similar structural folds. This figure was prepared using UCSF Chimera (Meng *et al*., 2006).

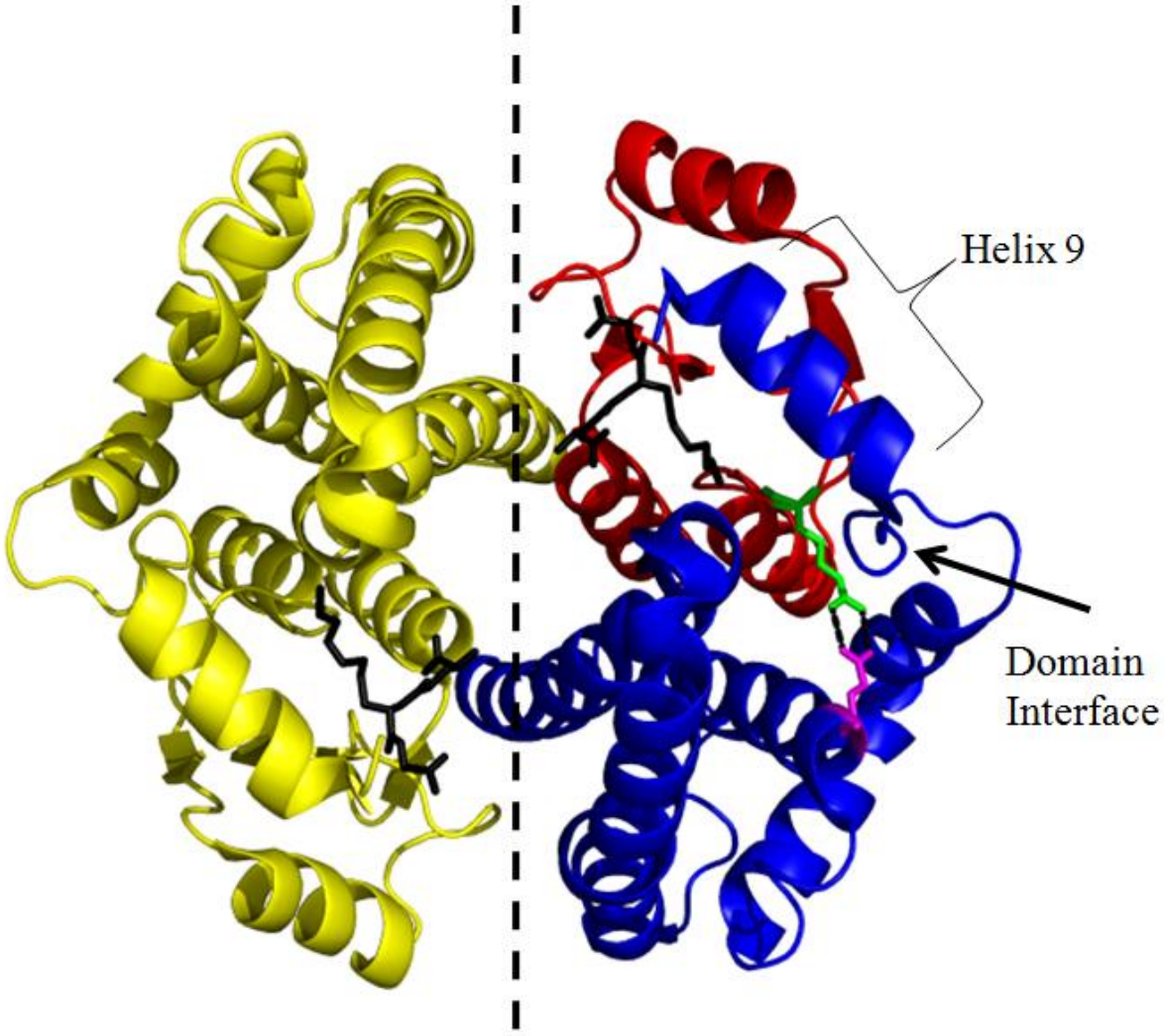


Figure 1.2: The crystal structure of the homodimer hGSTA1-1. Showing a complete A1 subunit in yellow, the second subunit is divided into domain 1 (red) and domain 2 (blue). *S*-hexyl glutathione (black), Arg13 on the active-site loop (green) and its salt bridge contact Glu169 on helix 6 (pink) are shown as stick models. The positions of Helix 9, the domain interface and the axis of symmetry at the subunit are signposted. Image generated from the PDB file 1K3L (Le Trong *et al*., 2002) using the program PyMol (www.pymol.sourceforge.net).

the classes of Sigma (Stevens *et al*., 1998) Beta (Sacchetta *et al*., 1993) and Mu (Hornby *et al*., 2000) dimerisation was shown to have a lesser influence on subunit stability.

These classes of GST unfold via a multi-step pathway and are able to form stable monomeric intermediates along the unfolding and refolding pathways. The class Mu and Sigma enzymes form molten globule-like inactive monomeric intermediates during unfolding indicating the extent of subunit stabilisation by quaternary interactions is less for these two classes than for the Alpha, Pi and Sj classes of GST which have been shown to be unable to form stable monomer units.

Though only the unfolded monomer and folded dimer of hGSTA1-1 are significantly populated at equilibrium, kinetic unfolding studies of hGSTA1-1 have indicated the existence of a three-state kinetic unfolding mechanism (Wallace *et al*., 1998). Additionally the presence of structured monomers possessing native-like secondary structures was observed in the kinetic refolding study (Wallace and Dirr, 1999). The intermediate during kinetic unfolding has not been assigned to a monomeric unit however; instead it is thought to be related to the dynamics of the unique α-helix 9 motif of the Alpha class of GSTs as well as changes at the domain interfaces of the hGSTA1-1 subunits around the Trp21 residue (Wallace *et al*., 1998).

1.5.5. The active site

Crystal structure data has verified that the GST dimers contain two active sites. The active sites are located at the confluence of the subunit and domain interfaces and are composed of structural elements of both domains of a subunit and from both subunits of the dimer (Dirr *et al*., 1994). A fully functional active site therefore requires domain association and dimerization for functionality (Dirr *et al*., 2004; Fabrini *et al*., 2009) (Figure 1.2). Evidence has also shown that the two active sites behave independently of each other (Danielson and Mannervik, 1985).

Each active site may be further sub-divided into two distinct binding sites: the hydrophilic G-site which binds glutathione and an adjacent H-site which binds the hydrophobic substrate (Figure 1.3)(Sinning *et al*., 1993). The G-site is highly specific for GSH while the H-site has

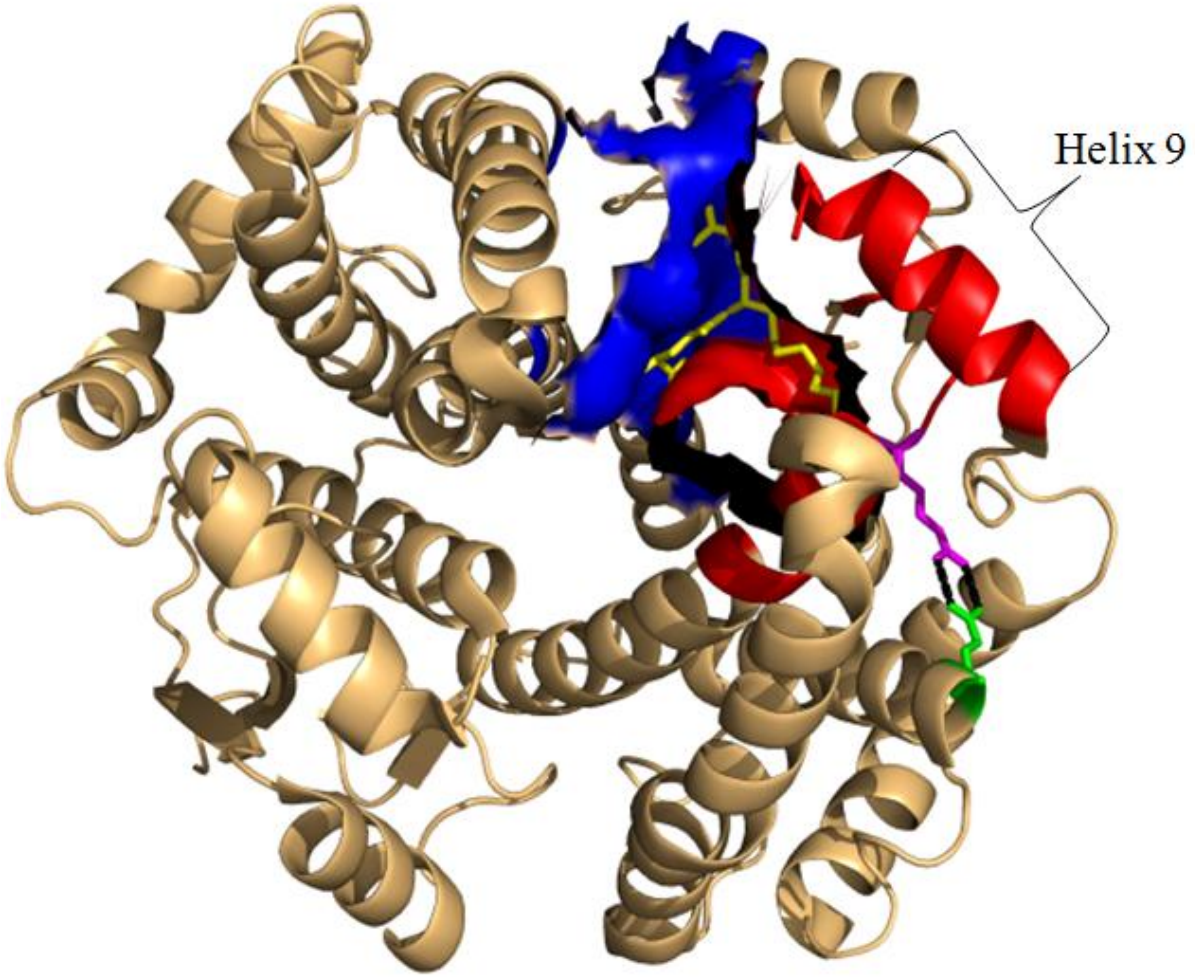


Figure 1.3: The crystal structure of the homodimer hGSTA1-1 highlighting the positions of the G- and H-site. The surface of the active site pocket is shown occupied with *S*-hexyl glutathione (yellow). The G-site is shown in blue while the H-site is shown in red. Arg13 on the active-site loop (pink) and its salt bridge contact Glu169 on helix 6 (green) are shown as stick models. Image generated from the PDB file 1K3L (Le Trong *et al*., 2002) using the program PyMol (www.pymol.sourceforge.net).

been found to bind a diverse range of hydrophobic, electrophilic substrates (Wilce and Parker, 1994).

Crystallographic data of hGSTA1-1 in complex with *S*-benzyl glutathione has allowed for the G-site to be defined (Sinning *et al*., 1993). The binding of GSH to hGSTA1-1 involves a network of specific polar interactions between the tripeptide and domain 1 of one subunit and two salt bridge interactions between both ends of the tripeptide and two amino acid residues (Asp101 and Arg131) from domain 2 of the other subunit (Dirr *et al*., 1994). In contrast there is only one salt bridge formed between the tripeptide and domain 2 in the cases of the Mu, Pi, Sigma or Sj classes or none as is the case for the Theta class. Additionally the sulphur atom of GSH forms two close contacts with the hydroxyl group of Tyr9 and to the N^{ε} of Arg15; both residues of the active site loop which connects β-strand 1 with helix 1 in domain 1 (Sinning *et al.,* 1993). These two interactions are vital in the catalytic action of the hGSTA1-1 (Sinning *et al*., 1993). Significantly there is an interdomain salt bridge formed between Arg15 and Glu104 which stabilizes the protein conformation at the G-site for hGSTA1-1, specifically it stabilizes the conformation of the active site loop (Gildenhuys *et al*., 2010). That the G-site is formed primarily from the highly conserved domain 1, thought to be descended from the single domain protein ancestor of all GSTs (Armstrong, 1997; Ladner *et al*., 2004), goes a long way to account for its high specificity for glutathione in hGSTA1-1 and all classes of GST.

Crystallographic data of hGSTA1-1 in complex with *S*-benzyl glutathione (Sinning *et al*., 1993), S-hexyl glutathione (Reinermer *et al*., 1992, 1996; Le Trong *et al*., 2002), ethacrynic acid (Cameron *et al*., 1995), glutathione-ethacrynic acid (Oakley *et al*., 1997), and *S*-(-p-nitrobenzyl) glutathione (Garcia-Sáez *et al*., 1994) has helped define the hydrophobic, electrophilic substrate binding site or H-site of GSTs.

In each case it can be seen that a highly hydrophobic pocket that is exposed to solvent is formed between three sides. Namely: the residues of the active site loop (Phe10, Ala12, Gly14 and the methylene groups of both Arg13 and Arg15); the C-terminus of helix 4 (Leu107, Leu108, Pro110, Val111 and the side chain of Glu104 which forms a salt bridge with Arg15); the region immediately prior to and a part of α-helix 9 (Met208, Leu213, Ala216, Ala219 Phe220 and Phe222) (Sinning *et al*., 1993). The C-terminal region of the

hGSTA1-1 protein has been shown to be an important part of the H-site (Board and Mannervik, 1991; Hoesch and Boyer, 1989; Sinning *et al*., 1993; Cameron *et al*., 1995).

Note that hGSTA1-1, when compared to other GST enzymes even other Alpha GST isoenzymes, has a far greater functional promiscuity, meaning that hGSTA1-1 is able to catalyse the conjugation of GSH to a far wider range of electrophilic substrates at its H-site. The cause was found to be due to structural heterogeneity not only at the active site (specifically the dynamic α-helix 9), but throughout the protein structure (Hou *et al.*, 2007). In particular flexibility at the domain interface around Trp21 was implicated in the broad catalytic range of hGSTA1-1.

Being unique to the Alpha class GSTs and due to its known role in forming the active sites of the Alpha class enzymes helix 9 has been the focus of much research. Helix 9 goes unobserved in the crystal structures of the apo hGSTA1-1 enzyme, due to its dynamic nature (Le Trong *et al.*, 2002; Sinning *et al*., 1993; Cameron *et al*., 1995). In the presence of an H-site ligand, however, the helix becomes highly ordered (Lian, 1998) and is visible in X-ray derived structures (Sinning *et al.*, 1993). Structures of hGSTA1-1 bound to ligands have therefore been used to examine helix 9 and have indicated the presence of numerous tertiary contacts between the helix and the protein (see Figure 8 in Dirr and Wallace, 1999). The helix 9 residue Ile219 contributes the majority of these tertiary contacts and it forms one with Phe10, an active site loop residue. Additionally Phe10 also forms tertiary contacts with the helix 9 residues Ser212, Glu215 and Ala216. These tertiary interactions between the helix and the protein stabilise the amphipathic helix but cannot do so in the absence of contributing stabilizing forces from an H-site ligand (Kuhnert *et al*., 2005).

Currently Helix 9 is known to exist in a wide range of conformational states and that it forms at times a lid over the active site (Sinning *et al*., 1993) (Figure 1.2 and Figure 1.3). The proof of this is that the apparent on-rate for glutathione binding to hGSTA1-1 is only 450 $mM^{-1}.s^{-1}$. Such a binding rate is far lower than the diffusion controlled value that could be expected for enzyme substrate association (Fersht, 1999). Variant proteins that destabilize or truncate the helix 9 from the protein often show a two-fold increase in the on-rate for glutathione binding (Gustafsson *et al*., 1999).

However, contrary to expectation while the removal or disruption of the helix 9 does improve the on-rate for glutathione binding it is accompanied by a decrease in enzyme performance (Dirr and Wallace, 1999). Though none of the residues in helix 9 contribute directly to the chemical mechanism of catalysis in hGSTA1-1, helix 9 and its conformational flexibility are known to indirectly influence the catalysis of GST in a number of ways: by maintaining the ability of Tyr9 (a vital catalytic residue of the active site loop) to favourably alter the pK_a of GSH at the active site (Gustafson *et al.,* 1999) and by forming a lid over and hence promoting the desolvation of the active site (Kuhnert *et al.,* 2005). Additionally Nilsson *et al.* (2002) suggested that the C-terminal region of helix 9 is responsible for orienting reactants and guiding them into the transition state

Further helix 9 appears to be involved in the binding of ligands at the H-site (Mosebi *et al*., 2003). It has been found that helix 9 contributes favourable entropy to ligand binding (Nielanik *et al.,* 2001) especially through interactions between the ligand and Phe222 (Kuhnert *et al*., 2005). Additionally the binding affinity for the ligands ANS and bromosulphophthalein (BSP) is enhanced when helix 9 is immobilized and stabilized by active-site ligands that occupy the G-site (i.e., glutathione and *p*-bromobenzylglutathione but not ethacrynic acid which occupies only a small portion of the adjacent G-site). The binding of ligands to hGSTA1-1 may stabilize the helix 9 not only through its interactions with Phe222 but because when a ligand binds to the H-site of the protein it causes Phe10 to swing into a new position that allows the side chain of Phe220 of the helix 9 to occupy its former position causing the C-terminal helix to close down and complete the formation of the binding site (Figure 6 in Le Trong *et al*., 2002).

Consequently it is no surprise, given the significant role helix 9 carries out in the catalytic function of hGSTA1-1 (Balchin *et al*., 2010; Dirr *et al*., 2005; Nillson *et al*., 2002) and its role in forming the active site an investigation by Hou *et al.,* (2007) found that conformational flexibility at helix 9 was partly responsible for the greater functional promiscuity of hGSTA1-1 as compared to other Alpha GST isoenzymes.

1.5.6. Catalytic mechanism

The essential event in the catalytic mechanism of GSTs is the deprotonation of the thiol group of cysteine in GSH, achieved through lowering its pK_a at the G-site, forming a GS^-

anion (Graminski *et al.*, 1989; Armstrong, 1991). This is facilitated by a Tyr, Cys or Ser residue. The nucleophilic GS^- anion then reacts with an aromatic electrophilic substrate, such as the widely used substrate CDNB, at the H-site by nucleophilic addition. The electrophilic substrate is thereby rendered harmless, and its water solubility is increased, promoting degradation of the product and subsequent excretion from the organism. The sequential mechanism referred to above may be represented as follows (Graminski *et al.*, 1989; Armstrong, 1991; Widersten *et al.*, 1996*):*

$$E + GSH \rightleftharpoons E{\cdot}GS^- + H^+ \rightleftharpoons E{\cdot}GS^-{\cdot}R\text{-}X \rightleftharpoons E{\cdot}GSR{\cdot}X^- \rightleftharpoons E + GSR + H{\cdot}X$$

where E is the GST enzyme, R-X is the hydrophobic electrophilic substrate, GSH is the reduced glutathione, X is the leaving group and GS^- is the highly reactive thiolate anion. GSH is believed to bind to the enzyme first due to the high GSH concentration in the cell (1-10 mM GSH) which is at least three orders of magnitude higher than its dissociation constant (Wilce and Parker, 1994; Dirr *et al*., 1994). The conjugation of GSH to 2, 4-dinitrochlorobenzene (CDNB) is shown in Figure 1.4 as an example.

The critical residue involved in this process in Alpha class GSTs is Tyr 9 (an active site loop residue), which acts as a hydrogen bond donor that bonds to the sulphur atom of GSH thus lowering the pK_a of the GSH sulfhydryl group from a pK_a of 9.0 to between 6.2 and 7.1 resulting in and hence stabilizing the thiolate anion (Wang *et al.*, 1992; Bjornestedt *et al*., 1995). Uniquely to class Alpha GSTs, in addition to Tyr9, Arg15 (also an active site loop residue), through the positive electrostatic potential generated by its guanidinium group is proposed to perform an important role in stabilizing the negatively charged thiolate form of GSH. Arg15 does this by forming a hydrogen bond with the hydroxyl group of Tyr9 and thus aids it in stabilizing the thiolate anion (Bjornstedt *et al*., 1995; Widersten *et al*., 1996; Gildenhuys *et al*., 2010). Site directed mutagenesis studies examining the replacement of Tyr9 with Phe (Allardyce *et al*., 1999; Stenberg *et al* 1991b) or Arg15 with either a leu or Ala (Dourado *et al*., 2010; Gildenhuys *et al*. 2010) have shown how vital both residues are to the proper catalytic function of hGSTA1-1.

The enzyme catalysed reaction of GSTs has been hypothesised to pass through a transition state which is an analogue of a Meisenheimer complex (Graminski *et al.*, 1989), a step which is typical of nucleophilic addition reactions. Further kinetic studies on the GST enzymes have

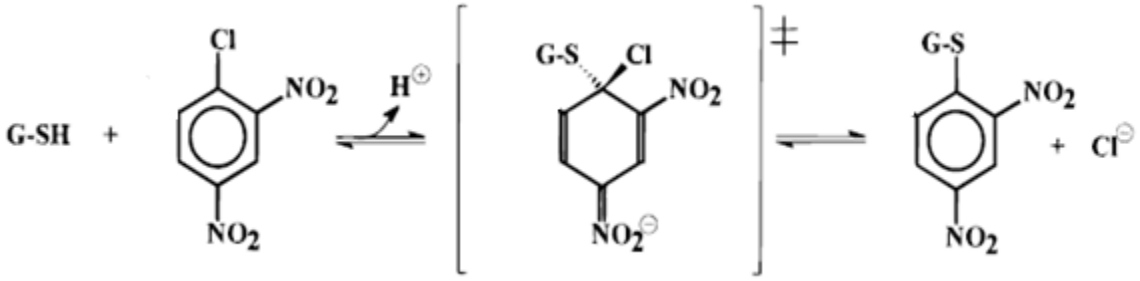


Figure 1.4: The catalytic mechanism of hGSTA1-1 Nucleophilic aromatic substitution reaction of GSH with CDNB to form S-(2,4-dinitrophenyl) glutathione, showing the putative transition state. The image is adapted from Widersten *et al*., (1996).

proposed that the formation of this transition step is the rate limiting step of GST catalysis (Chen *et al*., 1988; Graminski *et al.*, 1989; Ji *et al*., 1993).

1.5.7. Role of the domain interface in glutathione *S*-transferase

Evidence shows that the domain interface is crucial to the proper folding and stability of GSTs. It has been shown for instance that domains 1 and 2 of GST proteins unfold cooperatively in the case of class Alpha (Wallace *et al*., 1998), Sigma (Stevens *et al*., 1998), Pi (Erhardt and Dirr, 1995) and Sj (Kaplan *et al*., 1997) GSTs. Likewise, there is no evidence to support that the domains of GST proteins ever do unfold independently as indicated by the monophasic and coincident transitions of GSTs in several experiments (Hornby *et al*., 2000; Gildenhuys *et al*. 2008; Thompson *et al*., 2006; Wallace *et al*., 1998). Analysis of the domain interface in GST proteins indicates the domain interface to be large and predominantly hydrophobic with a high packing density explaining the cooperative behaviour (Han *et al*., 2007; Parbhoo *et al* 2011).

Further Luo *et al*. (2002) investigated the equilibrium unfolding and conformational stability of two domain-exchanged chimeric isoenzymes. The domains of the class Mu GST isoenzymes M1-1 and M2-2 were exchanged resulting in chimera GST subunit proteins M12 and M21. Each protein had one domain from GSTM1-1 and one domain from GSTM2-2 (Luo *et al*., 2002). The experiment showed that the resultant GSTM12 and M21 monomers were less stable than their wild-type counterparts (Luo *et al*., 2002). This indicates how critical domain interface complementarity is for correct domain-domain interaction which in turn will affect protein stability. The effect of disrupting the domain interface on stability and function was also investigated in the case of hGSTA1-1 (Wallace *et al*., 2000). A conserved residue in the class Alpha GST proteins, Trp21, was chosen for mutagenesis. Similarly to the lock-and-key motif found in the dimer interface of many GST proteins, the indole ring of Trp21 on helix 1 in domain 1 protrudes into a hydrophobic pocket formed by helix 6 and helix 8 in domain 2 where it is completely buried (Wallace *et al*., 2000) (Figure 1.5). Trp21 was replaced with either a Phe (W21F) or Ala (W21A). The side chain of phenylalanine is similar in size and hydrophobicity to that of tryptophan and as expected the W21F mutation did not affect the structure and function of hGSTA1-1, nonetheless it had a slight destabilizing effect on the protein (Wallace *et al*., 2000; Adman *et al*., 2001). However when

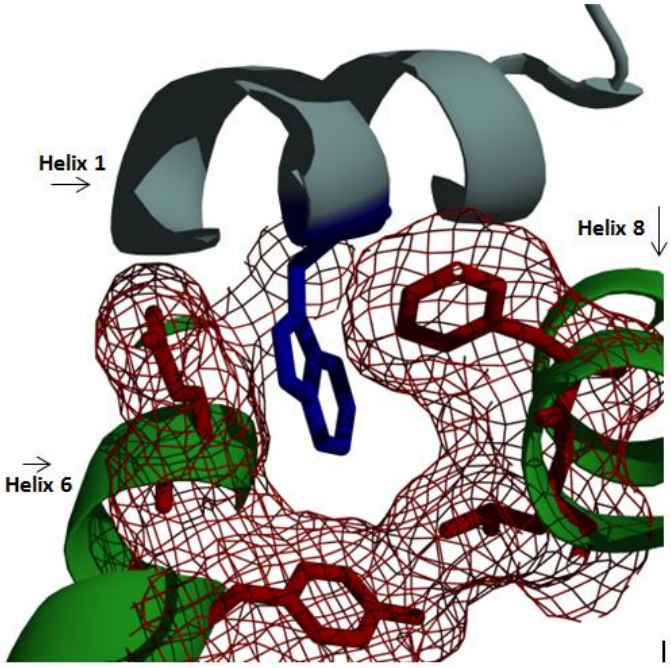


Figure 1.5: Representation of the Trp21 lock-and-key motif. Packing of Trp21 (blue) in domain 1 (grey) of hGSTA1-1 into the hydrophobic pocket formed by Ile153, Glu162, Tyr165, Phe197 and Leu198 (all red) in domain 2 (green). The mesh represents the van der Waals radii of the residues forming the binding pocket. Figure provided through Personal communication with David Balchin (Balchin *et al*. 2010). Generated from the PDB file 1k3l (Le Trong *et al*., 2002) using the program PyMol (www.pymol.sourceforge.net).

Trp was replaced with Ala, an amino acid with a much shorter side chain, a cavity forming mutation was the result. The W21A mutation was found to be both disruptive and destabilizing withthe equilibrium unfolding results also now indicating the presence of one or more intermediate species (Wallace *et al*., 2000). Bioinformatic studies have also revealed that this hydrophobic lock-and-key interdomain interface is highly conserved throughout the Alpha GSTs (Wallace *et al*., 2000) as well as other classes of GST (Stoychev, PhD thesis, 2008 (http://wiredspace.wits.ac.za/handle/10539/5604)) in addition to other thioredoxin fold containing proteins such as Clic (Stoychev, PhD thesis, 2008 (http://wiredspace.wits.ac.za/handle/10539/5604) ; Parbhoo *et al*., 2011) and Grx2 (Parbhoo, MSc Dissertation, 2010 (http://wiredspace.wits.ac.za/handle/10539/8557); Parbhoo *et al*., 2011). These results including the burial of a noteworthy amount of hydrophobic surface area at the domain interfaces of hGSTA1-1 highlights how the domain interface and its correct packing performs a significant role in the stability and function of the hGSTA1-1 protein. It is worth noting that hGSTA1-1 possesses a significantly larger domain interface compared to other GSTs (Stoychev, PhD thesis, 2008 (http://wiredspace.wits.ac.za/handle/10539/5604)).

Further, since the active site of GST is located at the domain interface it is clear that the structural integrity and conformational flexibility of the domain-domain interface is important to catalysis. Indeed in addition to the work done on Trp21 (Hou *et al.,* 2007; Balchin *et al*., 2010), it was found that site directed mutagenesis of Arg15 itself a domain interface residue also resulted in changed enzyme activity (Gildenhuys *et al*., 2010). Specifically in the case of the R15L variant enzyme it had less than 90% of the specific activity of the wild-type enzyme but showed only slight changes to its global structure and a minor loss of stability.

1.5.8. The domain interface Arg13-Glu169 salt bridge

It has long been understood from analysis of proteins that there is a good correlation between conserved amino acids and experimentally-identified amino acids which are important in structure, stability, function and folding (Keskin *et al*., 2005). In 1991 Stenberg *et al*. showed that a R13A variant hGSTA1-1 had compromised catalytic activity. The Arg13 residue was chosen for site directed mutagenesis studies because it was known to be conserved in the sequences of the Alpha, Mu and Pi classes of GSTs. Specifically the study by Stenberg *et al*.

(1991) showed that the R13A variant of hGSTA1-1 had lost 90% of its specific activity and had an 8-fold higher K_M value for GSH and a 1.4 times greater K_M value for CDNB and yet had greater affinity for the H-site inhibitors *S*-hexyl glutathione and BSP.

At the time of the study by Stenberg *et al*. (1991) the crystal structures of the GST isoenzymes and the method of their catalytic function had not yet been elucidated. The study hypothesized that the Arg13 may contribute directly to the binding of GSH by neutralising the negative charge of the carboxylate groups of GSH through its positive guanidium group. A reasonable postulate for the time but today it can be seen that the Arg13 residue though indeed part of the active site loop is in fact a domain interface residue and that specifically it is involved in a salt bridge interaction that reaches across from domain 1 to Glu169 on domain 2. Additionally it is clear that Arg13 is not in close proximity to the G-site and may not be directly involved in the binding of GSH (Figure 1.3). Rather the R13A mutation would instead result in the breaking of a salt bridge interaction at the domain interface and furthermore would introduce a cavity due to the loss of bulk. The exact nature of the relationship between the loss of the Arg13-Glu169 salt bridge and the loss in enzyme activity is therefore not clear and worthy of further study. Nonetheless based on what is known about hGSTA1-1 it is not surprising that a change in the domain interface of the enzyme results in a change in catalytic activity (section 1.5.7.) (Balchin *et al*., 2010). Additionally it is possible to speculate that the Arg13-Glu169 salt bridge may be important for maintaining the correct conformation of the active site loop and therefore its loss is likely to alter the H-site of the protein, perhaps affecting catalytic activity by changing the interactions between GSH and Tyr9 and Arg15 (both important catalytic and active site loop residues) and it may even disrupt the conformation of C-terminal helix 9 which is known to interact with Phe10 (also an active site loop residue). Alternatively the changes at the domain interface may result in an altered stability for the protein leading to the loss of catalytic activity.

1.5.9. Objective and aims

This study used hGSTA1-1 protein as a model multidomain protein in order to examine the interactions between the two domains of the protein and assess their influence on the structure, function and stability of the protein. The wide range of literature available on the

protein and the cytosolic GST superfamily makes it an excellent model protein to use in the investigation of the properties of the domain-domain interface.

The domain-domain interface in GST is topologically conserved in many classes of the GST superfamily and furthermore the domain interface Arg13 residue is highly conserved in the sequences of at least the Alpha, Mu and Pi classes of GST. It is perhaps to be expected that these three classes of closely related GSTs should share the Arg13 residue in their sequences though it raises the question if the Arg13 residue might also be conserved in the Sigma class of GST, as it along with the others form the major subgroup of the Y-GSTs (Atkinson and Babbit, 2009). It is worth considering if the Arg13 residue is conserved in any of the other classes of GSTs and if so does the residue perform the same role. Additionally experiments on R13A variant hGSTA1-1 have already shown the importance of the Arg13 residue to the enzymes proper catalytic activity and function (Stenberg *et al*., 1991) and while it was proposed at the time that Arg13 was probably a catalytic residue today it is clear that this is not so (Figure 1.3). Hence, the exact nature of the change that results in the loss of catalytic activity was never elucidated. The objective of this work will therefore be to determine what the effects are of disrupting the conserved interdomain Arg13-Glu169 salt bridge in a R13A variant hGSTA1-1 on the structure, stability and function of hGSTA1-1 and to assess the extent to which the role of this salt bridge might be shared amongst other classes of GST.

The research will aim to:

- Investigate through bioinfomatic techniques the extent to which the Arg13 residue is conserved throughout the GSTs.
- Use site-directed mutagenesis to create the R13A hGSTA1-1 variant protein.
- Determine the effect of the mutation on the structure of hGSTA1-1.
- Determine the influence of this mutation on specific activity.
- Determine the effect of the mutation on the proteins ability to bind the non-substrate ligand ANS.
- Determine the effect of the mutation on the stability of the protein.

Chapter 2. Experimental procedure

2. Materials

The pKHA1 plasmid encoding recombinant wild-type hGSTA1-1 was a generous gift from Professor B. Mannervik (Department of Biochemistry, University of Uppsala, Uppsala, Sweden: Stenberg *et al*., 1992). The QuikChange® II Site-Directed Mutagenesis kit was purchased from Stratagene (La Jolla, CA, USA). Ampicillin and chloramphenicol were purchased from Roche Diagnostics (Mannheim, Germany). Inqaba Biotech (Pretoria, South Africa) conducted synthesis of the necessary oligonucleotide primers and performed all sequencing to confirm the identity of the plasmids. The *Escherichia coli* BL21 (DE3) and BL21 (DE3) pLysS competent cells were obtained from Stratagene (La Jolla, CA, USA). The GeneJET™ Plasmid Miniprep Kit was obtained from Inqaba Biotech (Pretoria, South Africa). The compounds reduced L-glutathione (GSH), glutathione sulfonic acid and 8-anilino-1-naphthalenesulfonic acid (ANS) were of ultrapure grade and were purchased from Sigma-Aldrich Co. (St. Louis, MO, USA). 1-Chloro-2, 4-dinitrobenzene (CDNB) and urea was of ultrapure grade and were purchased from Merck Chemicals (Darmstadt, Germany). Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) protein molecular weight markers (SM0431) and DTT and IPTG were obtained from Fermentas Life Sciences (St. Leon-Rot, Germany). All other reagents were of analytical grade. All solutions and media were prepared using distilled water and filtered through a 0.45 µm filter unless otherwise stated.

2.1. Structural alignment and characterization of domain interfaces in the GST superfamily

To draw as much information as possible from known crystal structures of GST superfamily proteins, a structural alignment of representative GST and thioredoxin fold proteins was undertaken and analyzed in order to establish common trends and characteristic. The thioredoxin fold proteins were included in the alignment because the GST N-terminal domain is a thioredoxin fold domain. Since this study is concerned with multidomain proteins all thioredoxin fold containing protein examples were multidomain.

First, Proteins belonging to the GST protein family or which contained thioredoxin folds were identified using the Structural Classification of Proteins database (SCOP) (Murzin, 1995). Subsequently crystal structures of the GST and thioredoxin fold proteins were downloaded from the Protein Data Bank (PDB) (http://www.rcsb.org.pdb). Priority was given to those structures which had a high resolution (a cut of 3 Å was used but a resolution of 2 Å or better was preferred as at this resolution hydrogen bonds may be solved with certainty) and which were solved using apo protein. Apo structures were chosen because the Arg 13 residue under investigation in this thesis is so close to the hGSTA1-1 H-site, meaning that if any ligand was bound to the active site the induced conformational change at or near the H-site might alter the findings. For this thesis, eighteen crystal structures were used. See Table 1 for a list.

Table 1: Listing the protein crystal structures used in the structure based sequence alignment, their PDB codes and the resolution they were solved to.

	Protein	**PDB Code**	**Resolution (Å)**
Y-GSTs	**hGSTA1-1**	**1K3O**	**1.80**
	Mu GST	**1C72**	**2.80**
	Pi GST	**1TU7**	**1.50**
	***Plasmodium falciparum* GST**	**1K3O**	**1.80**
	Sigma GST	**2WB9**	**1.59**
S/C-GSTs	**Beta GST**	**1PMT**	**2.50**
	Delta GST	**1JLV**	**1.75**
	Omega GST	**1EEM**	**2.00**
	Phi GST	**1GNW**	**2.20**
	Tau GST	**1GWC**	**2.25**
	Theta GST	**2C3N**	**1.50**
	Zeta GST	**1E6B**	**1.65**
Thioredoxin fold proteins	**Clic**	**1K0M**	**1.40**
	Grx2, *Escherichia coli*	**1G7O**	**NMR**
	Grx2, T4 bacteriophage	**1DE2**	**NMR**
	eEF1Bgamma	**1NHY**	**3.00**
	SspA	**3LYK**	**2.20**
	Ure2p	**1G6W**	**2.50**

Before structure based sequence alignment was performed the crystal structure files of the proteins were first checked and subsequently modified using SwissPDB viewer V4.01 (Guex and Peitsch, 1997), so that each file would consist of only one subunit. Single subunits from each crystal structure were aligned using the inbuilt superpositioning function of UCSF Chimera and hGSTA1-1 as the reference structure. After structural alignment was performed structures with an Arg residue comparable to Arg13 of hGSTA1-1 were checked for salt bridges to residues in domain 2.

2.2. Preparation of the wild-type and mutant pKHA1 plasmids

The pKHA1 plasmid coding for wild-type hGSTA1-1 (provided by B. Mannervik) was modified by QuickChange mutagenesis (Stratagene, CA, USA) to replace the codon for Arg13 to code for an alanine instead. Primers were designed in accordance with the published wild-type nucleotide sequence of GST (Stenberg *et al*., 1992) using Primer-X (http://bioinformatics.orf/primerx) and analyzed for hairpins and loops by Generunner (V3.01, Hastings Software Inc., NY, USA). They are listed in Table 2.

Table 2: Oligonucleotide primer sequences used for site-directed mutagenesis

Protein name	Fw/Rev	Sequence
hGSTA1-1	Fw	5’ CACTACTTCAATGCAGCGGGCAGAATGGAGTC 3’
R13A	Rev	5’ GACTCCATTCTGCCCGCTGCATTGAAGTAGTG 3’

The position of the desired altered nucleic acid sequence is highlighted in yellow.

The mutant plasmid encoding R13A hGSTA1-1 was generated according to the protocol described by Papworth *et al*. (1996) in conjunction with the QuikChange site-directed mutagenesis kit. This three-step procedure is reported to generate mutants with greater than 80% efficiency in a single reaction. Briefly the procedure began with a reaction volume of 51 µl which comprised 5 µl (10x concentrate) reaction buffer, 1 µl (50 ng) double stranded DNA template (the pKHA1 plasmid), 1 µl (125 ng) forward primer, 1 µl (125 ng) reverse primer, 1 µl dNTP mix, 41 µl milli-Q water and 1 µl (2.5 U/µL) Pfu DNA polymerase. In the first step, the pKHA1 plasmids which were created in a natural cell environment and hence has methylated DNA, was denatured by heating the reaction volume to 95 °C for 30 s causing the strands of the double stranded DNA (dsDNA) plasmid to seperate. In the next step the

primers were permitted to anneal to the template DNA by lowering the reaction temperature to 55 °C for 60 s and in the final step DNA polymerase extends the primers to complement the template strands after the reaction is heated to 68 °C for 60 s. Fifteen additional amplification cycles were performed so the plasmids coding for the variant protein were built up in large numbers. Finally the parental DNA template was digested with 1 μl (10 U/μl) DpnI, for one hour at 37 °C and one hour at 20 °C. DpnI targets methylated DNA only and as such only the DNA which codes for wild-type hGSTA1-1 was digested. Afterwards cells were transformed with the remaining plasmids for the purposes of producing the variant protein (section 2.3.).

2.3. Purification of the R13A mutant pKHA1 plasmid

The final DNA product along with wild-type pKHA1 was used to transform *Escherichia coli* BL21 (DE3) pLysS cells. All inoculation and growth procedures were conducted under sterile conditions, using pre-sterilised media and containers, as well as aseptic techniques. Both wild-type and R13A modified pKHA1 were used to transform *Escherichia coli* BL21 (DE3) pLysS cells by a one-step heat shock (Chung *et al*. 1989). This method renders the cell membrane of *Escherichia coli* more porous to foreign DNA by subjecting it to rapidly changing temperature extremes of first ice and subsequently heat. The inclusion of the pLysS plasmid is important as it encodes T7 lysozyme which reduces basal polymerase transcription allowing the plasmid to become established (Studier and Moffat, 1986). The λ DE3 lysogen also carries the gene for T7 RNA polymerase required for transcription of the insert. 5 μl of each plasmid was incubated on ice with 50 μl *Escherichia coli* cells and were subsequently heat shocked at 42 °C on a heating block for 45 s before being transferred back to ice for two minutes. 900 μl SOC expression recovery medium (Super optimal broth with catabolite repression: 2% (w/v) tryptone, 0.5% (w/v) yeast, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 20 mM glucose) was added to each preparation, which was aftewards incubated at 37 °C for 1 hour on a shaker (250 rpm). Successful transformants were selected by plating the *Escherichia coli* cells on lysogeny broth (LB) agar (1% tryptone, 0.5% yeast, 0.5% NaCl) containing 100 and 30 $\mu g.ml^{-1}$ ampicillin and chloramphenicol respectively and incubating overnight (12-16 hours). Ampicillin and chloramphenicol select for the cells containing the pKHA1 and pLysS plasmids respectively, hence only successfully transformed cells will had the suitable resistances to thrive on the agar plate.

Successful colonies were chosen randomly and grown in 2xYT medium (1.6% (w/v) tryptone, 1.0% (w/v) yeast extract, 0.5% (w/v) NaCl) supplemented with 100 and 30 $\mu g.ml^{-1}$ of ampicillin and chloramphenicol respectively. The cells were cultured overnight in a shaking incubator under aerobic conditions at 37 °C at 250 rpm for 12-16 hours. The plasmid DNA was extracted from the turbid overnight cultures using the GeneJET™™ plasmid miniprep kit (Fermentas, St. Leon-Rot, Germany) in accordance with the protocol of the kit. The procedure is adapted from the standard alkaline-sodium dodecyl sulfate (SDS) lysis procedure (Ish-Horowicz and Burke, 1981) and based on the procedure Birnboim and Doly (1979) used to extract and purify plasmid DNA from bacterial cells.

2.4. Sequencing of the R13A mutant and wild-type pKHA1 plasmids

10 μL samples of both wild-type and presumed R13A modified pKHA1 plasmid was sent to Inqaba Biotech (Pretoria, South Africa) for sequencing using the sequencing primers listed in table 3. The resulting DNA sequences were checked to ensure the sole incorporation of the modified codon for the R13A substitution.

Table 3: Oligonucleotide primer sequences used for sequencing the ORF of the pKHA1 plasmid

Plasmid name	Fw/Rev	Sequence
pKHA1	Fw	5′ CGGCTCGTATAATGTGTGG 3′
	Rev	5′ CAGACCGCTTCTGCGTTC 3′

The resultant nucleotide sequences determined by Inqaba Biotech were aligned against the known mRNA nucleotide sequence encoding for hGSTA1-1 (GENE ID: 2938 GSTA1) using the BLASTN alignment tool (Altschul *et al.*, 1990). The nucleotide sequences were also translated using the Translate Tool of the ExPASy server (Swiss Institute of Bioinformatics) to confirm the presence of the R13A mutation.

2.5. Preparation of R13A and wild-type hGSTA1-1 proteins

The heterologous over-expression of both R13A and wild-type hGSTA1-1 encoding pKHA1 plasmids was performed through the method described by Wallace *et al.* (1999) and Alves *et al.* (2006). The method was designed to incorporate the optimal growth conditions of the *Escherichia coli* BL21 (DE3) pLysS cells to ensure optimal expression of each plasmid. The genes that encode for R13A and wild-type hGSTA1-1 on the pKHA1 plasmids are positioned on a lac operon and are downstream of the lac promoter. They are therefore subject to induction using IPTG (Perito *et al*., 1996). This occurs through the binding of IPTG to the lacI repressor, inhibiting it and allowing transcription via the lacZ promoter. Unlike the natural inducer of this gene (lactose), IPTG cannot be broken down in the cell, resulting in the genes becoming over-expressed which in turn leads to the production of large amounts of their gene-products. In this case either R13A or wild-type hGSTA1-1.

All further inoculation and growth periods were conducted in the presence of 100 $\mu g.ml^{-1}$ ampicillin and 30 $\mu g.ml^{-1}$ chloramphenicol, to create an environment that was only hospitable to the cells of interest, as well as using Erlenmeyer flasks that were five times greater in volume than that of the culture to ensure adequate aeration during growth.

2.5.1. Induction studies

According to Stenberg *et al*. (1991) the yield of R13A hGSTA1-1 was up to fifty times lower than that of the wild-type hGSTA1-1. Induction studies were therefore performed in order to determine if any alterations to the method used to over-express wild-type hGSTA1-1 would result in improved induction conditions for R13A hGSTA1-1.

Cultures of *Escherichia coli* BL21 (DE3) pLysS cells already transformed with the R13A mutant pKHA1 plasmid were incubated at 37 °C and 250 rpm until the start of the late-log phase (OD_{600} = 0.6). At this point IPTG was added to the cultures to different final concentrations (0.1-1.0 mM) to induce the over-expression of the pKHA1 plasmid. After induction the cultures were left to incubate a further 16 hours; some at 37 °C and others at 20 °C. Cells were then harvested by centrifugation at 5000 *g* for 20 minutes at 10 °C. The pellets were re-suspended in equilibration buffer (10 mM sodium phosphate, 1 mM ethylenediaminetetraacetic acid (EDTA), 0.02% sodium azide, pH 7.45), frozen at -20 °C and

thawed at 4 °C to facilitate cell lysis. Additionally 0.02 mg/ml DNAse, 0.02 mg/ml lysozyme and 2 mM $MgCl_2$ were added to further promote cell lysis and break down DNA, before lysis by sonication at 4 °C. The lysate was centrifuged at 13 000 *g* at 4 °C for 30 minutes and the supernatant and pellet collected. Samples of the supernatant and pellet were assessed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) according to the method of Laemmli (1970) (section 2.6.).

2.5.2. Protein production and purification

Escherichia coli BL21 (DE3) pLysS cells containing the pKHA1 plasmid for either wild-type or R13A mutant hGSTA1-1 were inoculated in 100 ml of sterile 2YT media (1.6% (w/v) tryptone, 1.0% (w/v) yeast extract, 0.5% (w/v) NaCl) and incubated for 12-16 hours at 250 rpm and 37 °C. Afterwards a 50-fold dilution of 20 ml of the *Escherichia coli* culture was performed into fresh and sterile LB media in a 5 L Erlenmeyer flask. The cultures were again incubated at 250 rpm and 37 °C until grown to critical concentration. This is determined spectroscopically to be when the culture enters the late-log phase (OD_{600} = 0.6). IPTG was subsequently added to the culture to a final concentration of 1 mM. Induction of the R13A mutant variant was carried out using the same procedure as above but with two exceptions. Firstly, eight times as much culture was used and secondly, IPTG was only added to a final concentration of 0.1 mM in accordance with the results of the induction studies (section 2.5.1.) that found this to be the optimal induction conditions for the R13A mutant. After the addition of IPTG over expression was allowed to continue overnight (12 - 16 hours, 37 °C, 250 rpm) before the cells were harvested as described previously (section 2.5.1.).

The wild-type and R13A mutant hGSTA1-1 proteins were purified from their respective supernatants using CM-Sepharose cation-exchange chromatography in accordance with a protocol adapted from the method proposed by Wallace *et al.* (1999). The respective supernatants, containing either soluble R13A or wild-type hGSTA1-1, were diluted to a final volume of 50 ml for wild-type and 400 ml for the mutant and filtered through a 0.2 μm syringe filter. This was done to avoid clogging the chromatographic resin with any large aggregates of cleaved lipid membranes.

Finally the diluted supernatant was loaded onto a CM-Sepharose cation exchange column pre-equilibrated with 10 column volumes of equilibration buffer (10 mM sodium phosphate, 1 mM EDTA, 0.02% sodium azide, pH 7.45) using an AKTA automated chromographic system. The column was washed with more equilibration buffer until an A_{280} reading of 0.000 was reached as recorded using the onboard Jasco V-550 UV/Vis spectrometer of the AKTA. An A_{280} reading of 0.000 indicates that all weakly interacting, unbound, protein had already been eluted from the column. The remaining bound protein was eluted from the column using a 0 – 300 mM NaCl gradient in equilibration buffer. The eluent was collected in fractions of 3 ml and spectroscopically assessed at 280 nm, again by the onboard spectrometer, for the presence of protein.

Samples from those fractions that were suspected of containing pure hGSTA1-1 had aliquots of 10 µL s subjected to discontinuous SDS-PAGE (Laemmli, 1970) (section 2.6.) using a 12% separating and 4% stacking gel alongside a molecular weight marker (Fermentas, St. Leon-Rot, Germany) and stained with Coomasie Blue to confirm the protein was the correct size.

Fractions containing pure protein were pooled and concentrated on ice to a volume of 10 ml by ultra-filtration using a membrane with a molecular weight cut-off of 10 kDa. Finally, the solution was dialysed against storage buffer (20 mM sodium phosphate buffer containing 1 mM EDTA and 0.02% sodium azide at pH 7.45), filtered through a 0.22 µm membrane and aliquoted in 1 ml volumes into cryogenic tubes. The contents of the cryogenic tubes were snap-frozen in liquid nitrogen and stored at -80 °C until use.

2.6. SDS-PAGE

The protein fractions obtained during the purification procedure (section 2.5.2.) were assessed by SDS-PAGE (Laemmli, 1970). SDS-PAGE is an electrophoretic technique wherein molecules are separated according to their electrophoretic mobility as they travel through a polyacrylamide matrix by the generation of an electric potential difference between either ends of the gel. However, since under normal conditions a protein molecule will migrate as a function of their charge, shape and size, SDS an anionic surfactant, is added to the solution to denature the proteins by disrupting non-covalent bonds and to mask their natural charges. Thus the protein molecules electrophoretic mobility becomes a function of

the size of the polypeptide chain alone, with shorter polypeptides migrating further through the gel than larger ones in the same time (Maurer, 1971). Since the distance travelled on the gel is a function of the size of the peptide only it is a useful and powerful technique to use to determine the purity and size of a protein, which goes a long way in allowing the identity of a protein to be inferred.

Purity is indicated by a single band on the gel and identity of the proteins can be inferred by the distance the protein band travelled alongside a control lane containing a known pure solution of the protein, or protein markers of known molecular weight. Pure wild-type or R13A hGSTA1-1 is indicated by a single band that correlates to the correct size of 27 kDa (the weight of a single hGSTA1-subunit) (Wallace *et al*., 2000).

A discontinuous gel system consisting of a 4% (w/v) acrylamide/bis-acrylamide stacking gel (0.1% (w/v) SDS, 0.05% (w/v) ammonium persulfate, 0.1% (w/v) N,N,N′,N′-tetramethylethylenediamine (TEMED) and 0.125 M tris(hydroxymethyl)aminomethane-Hydrochloric acid (Tris-HCl) buffer, pH 6.8) laid on top of a 12% (w/v) acrylamide/bis-acrylamide (w/v) separating gel (0.1% (w/v) SDS, 0.05% (w/v) ammonium persulfate, 0.1% (w/v) TEMED and 0.375 M Tris-HCl , pH 8.8) was used. Protein samples of 10 µL were diluted two-fold with sample buffer (10% (w/v) glycerol, 2% (w/v) SDS, 5% (w/v) β-mercaptoethanol, 0.05% (w/v) bromophenol blue). Samples were then boiled for 5 minutes to ensure that the proteins were fully denatured. The electrode buffer used contained 1% (w/v) SDS, 0.192 M glycine and 0.025 M Tris, pH 8.5. The protein samples (now 20 µl) were applied to the SDS-PAGE wells and electrophoresed at 180 V with discontinuous ampage for 45 minutes using a Bio-Rad electrophoresis system (Bio-Rad, California, U.S.). A molecular weight marker (Fermentas Life Sciences, (cat. no.: SM0431), Ontario, Canada) was run alongside the protein samples. The marker contains 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). After running 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 40% (v/v) ethanol and 10% (v/v) glacial acetic acid until the background was clear to visualize the results.

2.7. SE-HPLC

The dynamic volume and homogeneity of purified wild-type and R13A hGSTA1-1 was assessed using size exclusion–high performance liquid chromatography (SE-HPLC). The procedure was carried out using a TSK G2000 SWXL size exclusion column (TOSOH Corporation, Tokyo, Japan) along with a TOSOH TSK SWXL guard column. The column was pre-equilibrated with equilibration buffer that was pumped at an isocratic pressure of 40 bars through the column using a Dionex Ultimate 3000 pump (Sunnyvale, CA,). After equilibration 5 µM R13A or wild-type protein was loaded onto the column and washed with more equilibration buffer. Protein eluting from the column was detected using absorbance and fluorescence. The absorbance detector (Spectra-Physics (model: UV100SP), Fremont, CA, USA) was set to record absorbance at 280 nm while the fluorimeter (Jasco Inc., (model: FP 2020), Tokyo, Japan) was set to an excitation wavelength of 295 nm and emission wavelength of 345 nm. The experiment was carried out at lab temperature (25 °C) and data was collected using DTS software. (Malvern Instruments Ltd. Worcestershire, UK).

2.8. Determination of protein concentration

2.8.1. Absorbance spectroscopy

The concentration of the wild-type and variant GST proteins and ANS were determined spectrophotomically using a Jasco V-630 UV-VIS spectrophotometer (Jasco Inc., Tokyo, Japan) and by applying the Beer-Lambert law:

$$\mathbf{A = \varepsilon c} \boldsymbol{l} \qquad (3)$$

where A is the absorbance of the sample, ε is the molar extinction coefficient ($M^{-1}.cm^{-1}$), c is the concentration of the sample in solution (M) and *l* is the path length of the light passing through the solution (cm). The extinction coefficients used for determining ANS concentration was 5000 $M^{-1}.cm^{-1}$ at 350 nm (Weber and Young, 1964). The molar extinction coefficients for the wild-type and variant GST were calculated by using the formulae described by Perkins *(*1986):

$$\boldsymbol{\varepsilon = 5550\,\Sigma\,Trp + 1340\,\Sigma\,Tyr + 150\,\Sigma\,Cys} \qquad (4)$$

The molar extinction coefficients for the wild-type and variant proteins are the same since the mutation does not change any tryptophan, tyrosine or cysteine residues, the residues responsible for absorbance of light at 280 nm. The equation calculates an extinction coefficient of 38 200 $M^{-1}cm^{-1}$ for both the wild-type and R13A hGSTA1-1 proteins.

The concentrations were determined by fitting a linear regression to 5 or more points from a serial dilution. All readings were corrected for the signal caused by the buffer alone.

2.8.2. Bradford assay

The accuracy of the absorbance method to calculate the concentration of the GST proteins was confirmed by use of a Bradford protein assay. The Bradford assay is based on the shift in pK_a of an acidic solution of Coomasie blue G-250 dye upon binding to Arg and aromatic residues in protein. The pK_a shift causes a change in the colour of the dye allowing the protein concentration to be measured as a function of absorbance at 595 nm (Bradford, 1976). Samples of both wild-type and variant protein shown to have an equal A_{280} were diluted 100 times, reacted with one hundred µl of Bradford reagent (BioRad, California, USA) and has their A_{595} recorded.

While the R13A variant has one less Arg residue than the wild-type it is not thought the loss of a single, mostly buried, Arg residue will significantly affect the results.

2.9. Spectroscopic structural studies of R13A and wild-type hGSTA1-1 proteins

2.9.1. Circular dichroism spectroscopy

Circular dichroism (CD) is a spectroscopic method which relies on the fact that chiral molecules react differently with right and left circularly polarized light. That is to say a chiral molecule is optically active and will absorb left and right circularly polarized light even of the same wavelength differently. Since the majority of biological molecules are chiral this technique is useful in their examination. Since the technique is able to discern the difference between non super-imposable mirror images it can be seen to be highly sensitive to the three dimensional features of molecules and their structure (Woody, 1995). Therefore CD is a

powerful technique that can be applied to proteins. Chirality in proteins arises from disulfide groups, aromatic side chains, and the peptide backbone (Woody, 1995). CD can be further subdivided into two categories. Near-UV CD uses wavelengths of 250 - 300 nm and is sensitive to the disulphide groups and aromatic amino acids of a protein that have characteristic absorption bands in this range. Far-UV CD (180 - 250 nm) is sensitive to the peptide backbone conformation of a protein. Far-UV CD spectra recorded for proteins are therefore distinctive for and indicative of the unique secondary structure of a protein (Woody, 1995). For example proteins that have a high α-helical content display characteristic minima at 208 and 222 nm, as well as a strong positive band near 190 nm (Woody, 1995).

For this study Far-UV CD spectra (180 - 250 nm) were recorded using 2 μM wild-type and R13A variant GST. The protein was in GST storage buffer (20 mM sodium phosphate buffer containing 1 mM EDTA and 0.02% sodium azide at pH 7.45). All far-UV CD spectra were recorded at 20 °C and represent an average of 10 accumulations, at a scan speed of 200 $nm.min^{-1}$. The bandwidth used was 1 nm and the data pitch 0.2 nm. Measurements were obtained on a Jasco J-810 spectropolarimeter with Spectra Manager software v1.5.00 (Jasco Inc., Tokyo, Japan) and a path length of 2 mm. All spectra were buffer corrected. The spectra were normalised by calculating the mean residue ellipticity, [Θ] ($deg.cm^2.dmol^{-1}$) using the following equation (Woody, 1995):

$$[\Theta] = \frac{\mathbf{100} \times \Theta}{\boldsymbol{Cnl}} \qquad (5)$$

where, Θ is the measured ellipticity signal in mdeg, *C* is the protein concentration in mM, *n* is the total number of residues in the protein chain, and *l* is the path length in cm.

2.9.2. Fluorescence spectroscopy

Fluorescence is the emission of a photon that results from the return of an unpaired electron from an excited to ground state after it has been excited by the absorption of light energy (Lakowicz, 1999). In returning to its ground state the electron loses energy to molecular vibrations, heat and rotations. The energy lost between excitation and emission, known as Stokes' shift, results in the bathochromic (red) shift of emission spectra. Thus when the

electron emits a photon it is now lower in energy and hence longer in wavelength (Stryer, 1986).

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, 1999). By contrast the fluorescence of most proteins is dominated by tryptophan, with its quantum yield being more than double that of tyrosine. Additionally in the native state, tyrosine emission is often quenched by energy transfer to tryptophan, and to quenching due to nearby charged carboxyl and uncharged amino groups (Lakowicz, 1999). What makes tryptophan even more useful is that its indole ring, which is responsible for the fluorescence, is highly sensitive to solvent polarity (Lakowicz, 1999). It experiences a red shift in fluorescence as it is exposed further to an aqueous environment and a blue shift as it is further occluded into the protein and hence a hydrophobic environment. Therefore, tryptophan fluorescence is used to monitor tertiary structural changes in proteins which occur locally around a tryptophan residue.

All fluorescence measurements were recorded in a quartz cuvette with a 10 mm path-length using a Jasco FP-6300 spectrometer with Spectra Manager software v1.5.00 (Jasco Inc., Tokyo, Japan).

2.9.2.1. Intrinsic fluorescence-tryptophan fluorescence

hGSTA1-1 contains only one tryptophan residue, in domain 1 at position 21, and 10 tyrosine residues (three in domain 1, five in domain 2 and two in the connecting sequence linking the domains) per subunit. Trp21 is by far the dominant fluorophore and can be selectively excited by light at 295 nm while tyrosine residues are excited along with tryptophan residues upon exposure to light at 280 nm (Lakowicz, 1983). Additionally it has been shown that in wild-type hGSTA1-1 tyrosine fluorescence is masked by energy transfer to the tryptophan residues (Stryer, 1968).

Since Trp21 happens to be a domain interface residue and is located near the residue of interest, Arg13, it makes an excellent probe to detect any changes in the local tertiary structure around the mutation. Fluorescence emission spectra were recorded using protein

concentrations of 2 μM for both the wild-type and R13A variant in 20 mM phosphate buffer containing 1 mM EDTA and 0.02% sodium azide, pH 7.45. Spectra were recorded using excitation wavelengths of 280 nm and 295 nm, using slit widths of 5 nm (excitation) and 5 nm (emission) and a scan speed of 200 nm.min^{-1}. The spectra were buffer corrected and recorded between the excitation wavelength (to show Rayleigh scattering) and 500 nm and averaged from three samples with 3 accumulations each.

2.9.2.2. Extrinsic fluorescence-ANS ligand binding

The amphipathic ligand 8-anilino-1-naphthalene sulfonate (ANS) is a hydrophobic dye used as an extrinsic fluorescence probe (Engelhard and Evans, 1995). ANS is known to bind to the hydrophobic patches on proteins and further it has been shown by molecular docking and ligand displacement studies (Dirr *et al.*, 2005) and implicated by FRET (Sluis-Cremer *et al.*, 1996), to bind the H-site of hGSTA1-1. Specifically it is known that the hydrophobic anilino and naphthyl rings of the ANS molecule occupy the H-site and its polar sulfonate moiety binds at the interface of the G- and H-sites (Dirr *et al.*, 2005).

Free ANS in an aqueous environment emits maximally at 540 nm. Upon binding to a hydrophobic surface the dye experiences an increase in its fluorescence quantum yield and a blue-shift in its maximum emission wavelength, with the exact wavelength varying according to the hydrophobicity of the binding pocket (Engelhard and Evans, 1995; Stryer, 1965; Slavik, 1982). The fluorescence properties of ANS were therefore used to probe for changes to the conformation of the H-site between the wild-type and R13A variant hGSTA1-1 protein

Stock solutions of ANS (10 mM) were prepared in hGSTA1-1 storage buffer (20 mM sodium phosphate buffer, pH 7.45, containing 1 mM EDTA and 0.02% sodium azide). The concentration of the ANS stock was confirmed spectroscopically by measuring the absorbance at 350 nm and using extinction coefficient of ε_{350} = 5000 M^{-1}cm^{-1} (section 2.8.1.) (Weber and Young, 1964). ANS stock solutions were stored at 4 °C in foil wrapped containers to prevent photodecomposition and were used within 14 days of preparation.

Sample solutions of 2 μM wild-type and R13A variant hGSTA1-1 proteins were prepared in the presence of 200 μM ANS in hGSTA1-1 storage buffer at different concentrations of Urea (0 – 8 M) . Samples were incubated for 1 hour at 20 °C in the dark to promote the complete

binding of ANS to the protein. The samples were excited using a Jasco FP-6300 fluorimeter (Jasco Inc., Tokyo, Japan) at 390 nm, excitation and emission slit widths were at 5 nm, and emission spectra were recorded from 390 to 600 nm. All spectra were recorded at 20 °C; buffer corrected, and are an average of three accumulations at a scan speed of 200 nm.min^{-1}. An excitation wavelength of 390 nm was used even though this is not the optimal excitation wavelength for ANS in order minimise the inner-filter effect, the quenching of fluorescence by competing fluorophores namely the protein and ANS molecules (Birdsall *et al.*, 1983; Kubista *et al.*, 1994).

2.10. GSH-CDNB conjugation assay

2.10.1. GSH concentration determination

To accurately determine the concentration of stock solutions of the GST substrate GSH samples were treated with Ellman's reagent (Ellman, 1959) or DTNB (5, 5'-dithio-*bis*-(2-nitrobenzoic acid)). DTNB reacts with sulfhydryl groups to produce 2-nitro-5-thiobenzoic acid, a yellow-coloured product which absorbs at 412 nm and has a molar extinction coefficient of 14 150 M^{-1}.cm^{-1} (Riddles *et al.*, 1979), allowing for the concentration of GSH to be determined using absorbance spectroscopy (section 2.8.1).

2.10.2. Specific activity

The specific activity of the wild-type and R13A variant hGSTA1-1 enzyme was calculated using the CDNB conjugation assay described by Habig and Jakoby (1981) (Habig *et al.*, 1974). The enzyme catalysed reaction of GSH with 1-chloro-2, 4-dinitrobenzene (CDNB) produces 1-(*S*-glutathionyl)-2, 4-dinitrobenzene, which can be detected by absorbance at 340 nm (ε = 9 600 M^{-1}.cm^{-1}), allowing the progress of the reaction to be monitored spectroscopically.

Triplicate sample solutions of 1–10 μM variant and wild-type 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. The assay was performed at a pH of 6.5 because at this pH the non-catalysed reaction of GSH with CDNB is minimal. The conjugation reaction was initiated by the addition of 90 μL of a 33.3 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 amount of ethanol in the assay solution never exceeded 3% (v/v). All reactions were performed at 20 °C and followed using a Jasco V-630 spectrophotometer (Jasco Inc., Tokyo, Japan). The rate of product formation was calculated by monitoring the increase in the A_{340} reading after correcting for the non-enzymatic reaction. Finally the specific activity (μmol.min^{-1}.mg^{-1}) was determined by linear regression, using Sigmaplot v11.0. (Systat Software Inc; Chicago, IL, USA), of the slope of a plot between the initial velocities of complex formation (μmol.min^{-1}) versus the mass of protein (mg). Sigmaplot was also used to calculate the first standard deviation of a plot point represented by error bars (a measure of an experiments precision) as well as the standard error of the calculated fit (a measure of the experiments accuracy). The same was done for all fitted data.

2.11. Stability studies of the R13A and wild-type hGSTA1-1 proteins

2.11.1. Thermal denaturation studies

Thermal unfolding is widely used as a measure of protein stability. Thermal induced protein unfolding studies were conducted to assess the relative stabilities of wild-type and R13A hGSTA1-1 by monitoring the extent of α-helix denaturation in response to changes in temperature.

This was done by monitoring the ellipticity at 222 nm (section 2.9.1.) of samples of 2 μM R13A and wild-type hGSTA1-1 protein in 20 mM sodium phosphate buffer, pH 7.5, containing 1 mM EDTA and 0.02% sodium azide. Readings were taken as samples were subjected to an increasing and subsequent decreasing temperature over a 20 °C to 80 °C range. The temperature was controlled by a Jasco PTC-423S Peltier-type temperature control system (Jasco Inc., Tokyo, Japan) attached to a water-cooling bath using a continuous temperature gradient of 1 °C.min^{-1}. Samples were irradiated with 222 nm circular polarised light using a Jasco J-810 spectropolarimeter with Spectra Manager software v1.5.00 (Jasco Inc., Tokyo, Japan) employing a 1 nm bandwidth filter and a 2 mm path length. All measurements were recorded using a 0.5 s response time, standard (100 mdeg) sensitivity and a 0.1 °C data pitch. The temperature unfolding profiles were normalised by calculating the mean residue ellipticity [θ] deg.cm^2.dmol^{-1}.residue^{-1}

2.11.2. Urea-induced equilibrium unfolding studies

2.11.2.1. Reversibility of unfolding

In order to assess the conformational stability of the protein it is essential to first determine the degree of reversibility of the unfolding process as it will give an indication whether equilibrium between the folded and unfolded states exists. This is necessary as the calculation of the change in Gibbs free energy requires that the unfolding event is reversible.

The reversibility of unfolding was therefore determined for wild-type and R13A hGSTA1-1. The refolding of samples of 20 μM unfolded protein, incubated in GST storage buffer with 8 M urea for one hour at 20 °C, was achieved by a 10-fold dilution of each sample with GST storage buffer (20 mM sodium phosphate buffer containing 1 mM EDTA and 0.02% sodium azide at pH 7.45) for 1 hour at 20 °C. The refolded state of the protein was assessed using far-UV CD (section 2.9.1.) and intrinsic tryptophan fluorescence (section 2.9.2.1.). The amount of protein that was able to return to its native state was calculated by the comparison of the spectra of refolded protein with spectra of a control sample of 2 μM protein, either wild-type or R13A variant as appropriate, in 0.8 M urea.

Reversibility of unfolding is only an indication of equilibrium between the folded and unfolded states. To prove that equilibrium between the folded and unfolded states does exist it is necessary to show that the protein refolds along the same path that it unfolds.

All of the urea used for experiments was prepared by the method of Pace (1986) using GST storage buffer as the solvent. Following preparation, the pH of the stock urea solution was adjusted to pH 7.5 using phosphoric acid, filtered and the concentration of 10 M confirmed using an Atago R5000 pocket refractometer (Tokyo, Japan).

2.11.2.2. Urea-induced equilibrium unfolding/refolding

Protein unfolding transitions are very effective tools for calculating the stability of a protein. The effect of site-directed mutagenesis on the stability of a protein can be studied by comparison of the equilibrium unfolding curves of wild-type and variant protein. A denaturant can be used to shift the equilibrium of a protein from the native to the unfolded state. The equilibrium constant (*K*eq) can then be calculated and hence the conformational

stability parameters $\Delta G_{(H2O)}$ and *m*-value can be determined provided reversibility has been established (section 2.11.2.1.).

Urea-induced unfolding of 2 μM wild-type and R13A hGSTA1-1 was conducted in GST storage buffer over a range of urea concentrations from 0 M to 8 M urea in the absence and presence of 200 μM ANS. Samples were incubated at 20ºC for 1 hour to allow equilibrium to be reached but all experiments were performed within three hours of mixing to avoid possible protein aggregation interfering with the results. Samples were monitored using far-UV CD (section 2.9.1.) and fluorescence spectroscopy (section 2.9.2.1.). CD was used to monitor the increase of $[\Theta]_{222}$, while fluorescence spectroscopy monitored the combined fluorescence spectra of the Trp and Tyr residues (Ex_{280}) of hGSTA1-1 as a means of tracking protein unfolding with increasing concentrations of urea. The experiment was repeated in the presence of ANS as it is a probe to monitor the conformational changes induced in proteins during the unfolding process (Semistonov *et al*., 1987; Semistonov *et al*., 1991). Likewise ANS is used as a probe for the presence of intermediates during protein unfolding events (Semistonov *et al*., 1991), it has been shown that ANS when bound to a protein in its molten globule state (a common intermediate found during unfolding transitions) shows higher emission intensities than when bound to protein in either the folded or unfolded state (Ptitsyn, 1995). The samples were also excited at 340 nm and fluorescence recorded at 340 nm as a way to test for Rayleigh scattering, as this is a test for the presence of protein aggregates.

Refolding of the R13A variant hGSTA1-1 was also conducted over the same range of urea concentrations to prove equilibrium and the absence of hysteresis (section 2.11.2.1.). A stock solution of 20 μM unfolded R13A variant hGSTA1-1 protein in GST storage buffer with 8 M urea was diluted, with stocks of 10 M urea and GST storage buffer as appropriate, to a protein concentration of 2 μM in various concentrations of urea and left to incubate for 1 hour at 20ºC. The resultant unfolding and refolding curves were compared to identify whether the unfolding of the R13A variant hGSTA1-1 does indeed take place under equilibrium conditions.

2.11.2.3. Data fitting

All unfolding and refolding data obtained for the proteins were analysed according to a two-state unfolding process for a dimeric protein.

During a reversible two-state unfolding of a dimeric protein equilibrium is reached between the native (N) and the unfolded species (U):

$$N_2 \rightleftharpoons 2U \quad (6)$$

Furthermore during a two-state unfolding transition, only the unfolded and native states are present at significant concentrations (Pace, 1986). Hence, for a two-state mechanism:

$$f_N + f_U = 1 \quad (7)$$

where f_N is the fraction of folded or native protein and f_U is the fraction of unfolded protein. The recorded spectroscopic signal at any point during the unfolding of the samples is therefore in fact the sum of the signal of both species:

$$y_0 = y_N f_N + y_U f_U \quad (8)$$

where y_0 is the signal obtained for the respective spectroscopic probe, f_N represents the fraction of folded protein, and f_U represents fraction unfolded protein. Further y_N represents the y value for the folded state and can be extrapolated from the linear pre-transition region of the unfolding data, while 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 7 and 8, the fraction of unfolded protein can be obtained:

$$f_u = \frac{(y_N - y)}{(y_N - y_U)} \quad (9)$$

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

$$K_{eq} = \frac{[f_U]^2}{[f_N]} \quad (10)$$

The total protein concentration, P_t, expressed in terms of the monomeric species is:

$$P_t = 2N_2 + U \qquad (11)$$

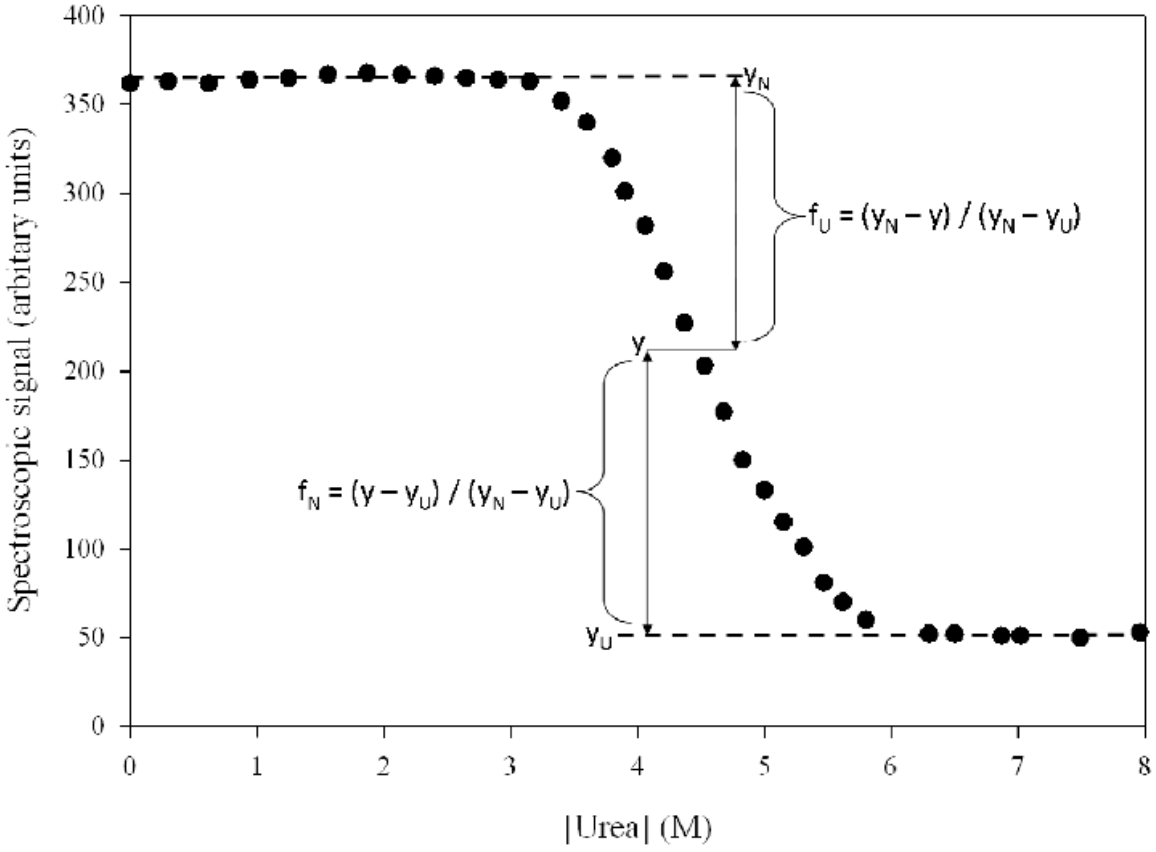


Figure 2.1: Representation of a typical two-state urea denaturation curve. The pre-transition region (plot of Y_N) where the protein remains folded, post transition region (plot of Y_U) where the protein is fully unfolded and the transition region where the protein is a state of equilibrium between the folded and unfolded states as a function of urea concentration are shown. Image adapted from Shirley (1995).

and the fraction of native dimeric species is:

$$F_N = 1 - F_U \quad (12)$$

where:

$$F_U = \frac{[U]}{P_t} \quad (13)$$

Combining these equations and solving for F_U in terms of the equilibrium constant K_{eq} and P_t, one obtains the following equation:

$$F_U = \frac{\sqrt{{K_{eq}}^2 + 8K_{eq}P_t} - K_{eq}}{4P_t} \quad (14)$$

Rearranging equation 9 yields the equation used for fitting the denaturation transitions:

$$y_0 = y_N(1 - f_U) + y_U(f_u) \quad (15)$$

Substituting equation 15 allows the k_{eq} of unfolding to be calculated and hence ΔG° according to the equation:

$$\Delta G^\circ = -RT\ln K_{eq} \quad (16)$$

where $\Delta G°$ is the free energy of unfolding, R is the gas constant and T is temperature in Kelvin. In order to determine $\Delta G_{(H2O)}$ it is assumed that according to the linear free-energy model $\Delta G°$ has a linear dependence on denaturant concentration [D] for all urea concentrations (Tanford, 1968, 1970). Therefore:

$$\Delta G^\circ = \Delta G_{(H20)} + \mathrm{m[D]} \quad (17)$$

where $\Delta G_{(H2O)}$ 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 (calculated from the linear equation 17 which has the formulae y = mx+c)

which is also an indicator of co-operativity and can be related to the change in solvent-accessible surface area ($\Delta SASA$) (Pace, 1975), and [D] is the denaturant concentration.

The equilibrium unfolding data obtained was fitted with these equations using SigmaPlot version 11.0 (Systat Software Inc; Chicago, IL, USA) and the parameters $\Delta G_{(H2O)}$ and *m* were obtained.

2.12. R13A hGSTA1-1 Crystallization

X-ray crystallography is a powerful technique that can solve the three dimensional structure of a protein to a high atomic resolution. The technique requires the protein to be solved first be crystallized into a large pure crystal with a high regularity. This step in many cases is the chief barrier to solving the atomic resolution structure of proteins (Geerlof *et al*., 2006). It is extremely difficult to predict the conditions for nucleation and growth of protein crystals (Rupp and Wang, 2004). In practice it is generally required to identify these conditions by screening the protein of interest against a large variety of crystallizing solutions (Chayen 2004).

In order to be able to carry out X-ray crystallography on the R13A variant hGSTA1-1 enzyme several attempts were made to crystallize the protein. Firstly a method similar to the one used by le Trong *et al*. (2002) to crystallize the wild-type enzyme was attempted. Specifically, the hanging drop diffusion method (McPherson, 1990) was used at 25°C and with a reservoir volume of 1.5 ml. Hanging drops were made using 2, 3 and 4 times dilutions of protein (10 mg/ml in 0.1 M Tris-HCl, 5 mM dithiothreitol (DTT) and 0.02 % NaN_3, pH 7.5) with reservoir buffer (0.1M Tris-Cl, 15-30 % (w/v) polyethylene glycol 3350, 0.02 % NaN_3 and 1 mM DTT, pH 7.5). Additional attempts were carried out using buffers from a Hampton index (Index HR2-144)(Hampton Research, Aliso Viejo, CA, USA) with which the Protein Structure Function Research Unit (School of Molecular and Cell Biology, Witwatersrand University, South Africa) has had previous success crystallizing GST proteins. See Appendix Table A for a list of the solutions used. Finally a third attempt was made, again using the method of Le Trong *et al*. (2002) but this time trying to crystallize the variant protein bound to its inhibitor glutathione sulfonate. The inhibitor was chosen because it is has a inhibitor strong affinity for the G-site and low overall bulk (Adman *et al*. 2001) and would

provide additional rigidity to the protein. Ligands bound to the active site of the enzyme are known to reduce the conformation flexibility of helix 9, something which is known to enhance the likelihood of crystallization by reducing entropy. Additionally this specific inhibitor binds mostly to the G-site of the protein so it is reasonable to expect that its binding will not overly mask any changes occurring to the protein at or around the R13A substitution or to the H-site where it is likely most of the changes to the protein structure are taking place (Adman *et al*. 2001).

CHAPTER 3. Results

3. Multiple structural alignments

The multiple structural alignments of the GST and dimeric thioredoxin fold proteins evident in Figure 3.1 show that the Arg13 residue of hGSTA1-1 is topologically conserved in the Alpha, Mu, Pi, *Plasmodium falciparum* and Sigma classes of GST. These findings match the results of the structural alignment performed by Atkinson and Babbitt (2009) with the additional finding that the Arg is conserved in the *Plasmodium falciparum* class.

Inspection of the GST proteins from the Alpha, Mu, Pi, *Plasmodium falciparum* and Sigma classes showed that in all cases the Arg residue forms a salt bridge interaction with either a Glu or a Asp residue of domain 2 (Figure 3.2).

3.1. Purity and sequencing of the mutant and wild-type DNA

The plasmid purification procedures produced 50 μL of 147 ng/μL mutant pKHA1 plasmid deoxyribonucleic acid (DNA) and 50 μL of 154 ng/μL wild-type pKHA1 plasmid DNA. The ratio of the calculated A_{260} (absorbance contributed by the DNA) and the A_{280} (absorbance contributed by protein) (i.e. the A_{260}/A_{280}) was equal to 1.89 for the mutant DNA and 1.94 for the wild-type. These A_{260}/A_{280} values exceed the minimum requirement for DNA purity (Maniatis *et al.*, 1982). Similarly the ratio of A_{260} to the A_{230} (absorbance contributed by carbohydrate impurities) (i.e. the A_{260}/A_{230}) was equal to 2.15 for the mutant DNA and 2.1 for the wild-type. The wild-type and R13A mutant pKHA1 plasmids were of sufficient purity and concentration for sequencing and transformation procedures.

The sequencing results of the open reading frame (ORF) for wild-type hGSTA1-1 encoding plasmid corresponded exactly to the sequence identified by Stenberg *et al*. (1992) (Figure 3.3). The sequencing of the ORF of the mutant plasmid indicated the desired mutation CGG → GCG had been incorporated and that no undesirable mutations had occurred.

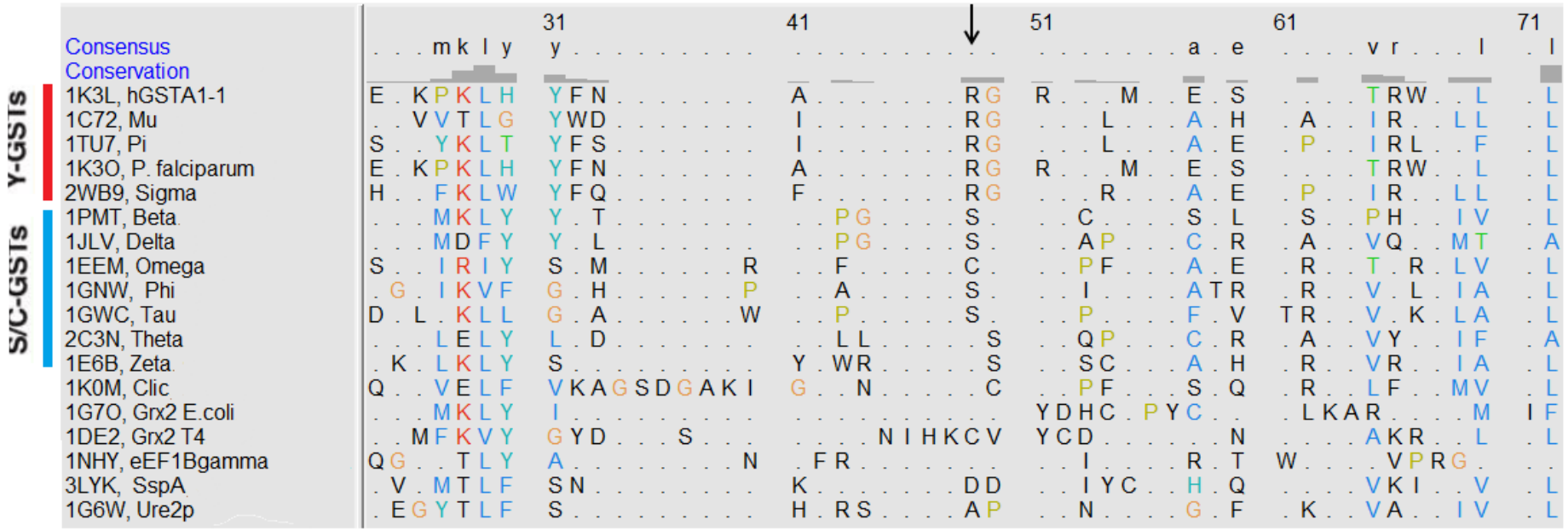


Figure 3.1: **Structure based sequence alignments of 18 thioredoxin fold containing proteins.** This alignment is created from representatives spanning the cytosolic GSTs as well as structural homologues. It lists both the PDB entry and GST class label for each structure. The arrow indicates the position of the Arg13 residue of hGSTA1-1 and the topologically equivalent residues of the other proteins. From the diagram it is possible to see that the Arg13 residue of hGSTA1-1 is highly conserved in Alpha, Pi, Mu, Sigma and *Plasmodium falciparum* GSTs. This alignment was prepared using UCSF Chimera (Meng *et al*., 2006).

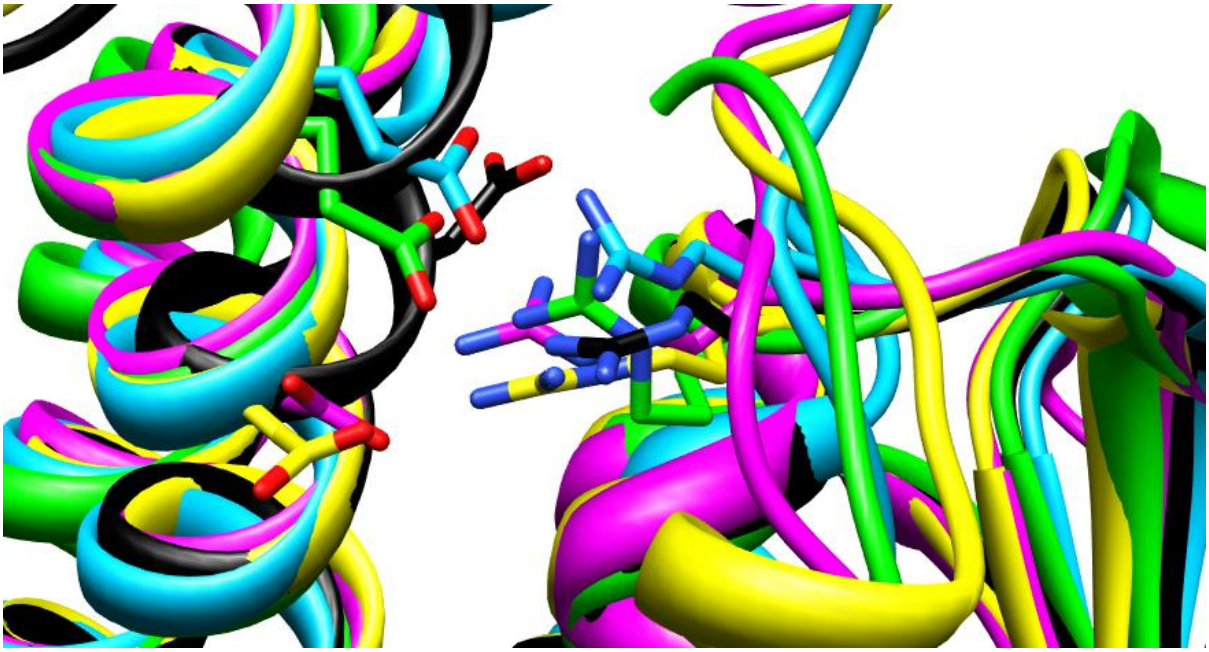


Figure 3.2: Comparing the structural role of the conserved Arg residue in the Alpha, Mu, Pi, *Plasmodium falciparum* and Sigma classes of GST. Diagrammatic representation showing the conserved interdomain interaction in Alpha GSTs (cyan, between Arg13 and Glu169), Mu (pink, between Arg10 and Asp 164, Pi (yellow, between Arg11 and Asp159), *Plasmodium falciparum (*black, between Arg13 and Asp171) and Sigma GSTs (green, between Arg14 and Glu172).The alignment and distance between the residues are all indicative of Hydrogen and salt bridge interactions. This figure was prepared using UCSF Chimera (Meng *et al*., 2006).

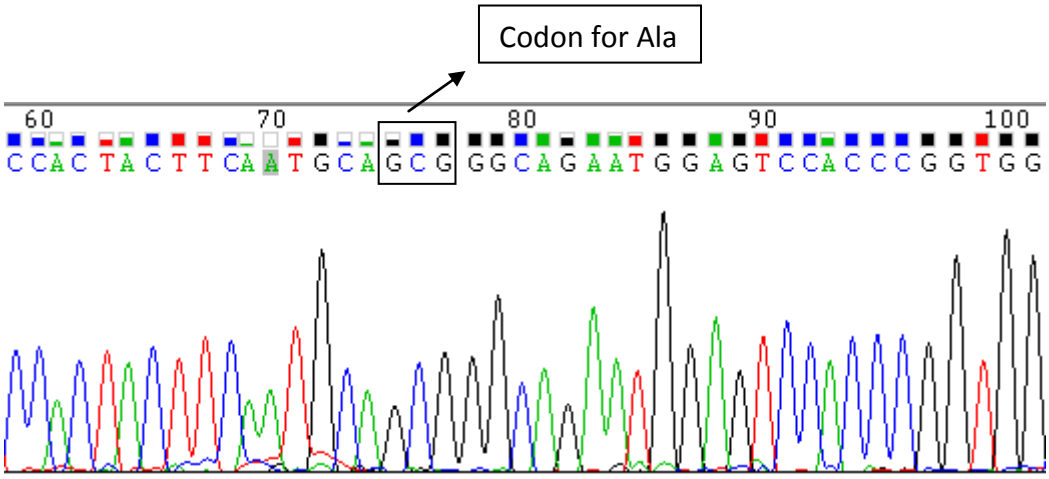


Figure 3.3: R13A mutant pKHA1 plasmid sequencing results. A selected segment of the pKHA1 ORF which encodes for the R13A variant hGSTA1-1 protein. The desired mutation of an Arg codon (CGG) to an Ala codon (GCG) is highlighted. The sequencing results were viewed using the program Finch TV version 1.4.0 (http://www.geospiza.com/FinchTV: Geospiza Inc.).

3.2. Purity of R13A and wild-type hGSTA1-1

3.2.1. Over expression and purification

The heterologous expression system for both mutant and wild-type hGSTA1-1 comprised *Escherichia coli* BL21 (DE3) pLysS cells transformed with pKHA1 plasmid that encoded the respective proteins (section 2.5.2.). Induction studies were performed to optimise the expression of R13A and it was found that the mutant expressed optimally in one tenth (0.1 mM) of the IPTG that was used to express the wild-type protein (Figure 3.4). Both proteins were found to be soluble in the supernatant fraction of the lysate after sonication and centrifugation of the *Escherichia coli* BL21 (DE3) pLysS cells. The proteins were purified by means of cation-exchange chromatography on a CM-Sepharose column and eluted with a salt gradient (section 2.5.2.).The purification of both R13A and wild-type hGSTA1-1 produced a single symmetrical peak at ~140 mM sodium chloride (40% of the 300 mM NaCl buffer), as shown in Figure 3.5

1 L of wild-type expressing culture was found to yield ~35 mg of purified protein whereas 1 L of R13A mutant expressing culture was found to yield only ~1.9 mg of purified protein. Thus despite the optimization of the expression system of the mutant it still had a yield far lower than the wild-type protein. Nonetheless the comparative yields obtained were an improvement of the reported findings of Stenberg *et al*. (1991).

3.2.2. Size and purity determination

SDS-PAGE was used to analyze the molecular mass and purity of wild-type and R13A hGSTA1-1. Both the variant and wild-type proteins were judged to be electrophoretically pure as samples of both proteins showed only a single band (Figure 3.6 A). Additionally as the bands were just above the 25 kDa molecular weight marker (REase BSp) this is a strong indication that the produced proteins are the hGSTA1-1 proteins of interest. The calculated weight of the proteins, accomplished by comparing the distances the GST proteins migrated in comparison to the set of known standards that migrated under the same conditions, was 26.2 kDa which correlates to the expected 27 kDa (Wallace, *et al*. 1998) (Figure 3.6 B). The variant and wild-type proteins were also found to be homogenous using SE-HPLC under native conditions (section 2.7.) (Figure 3.7).

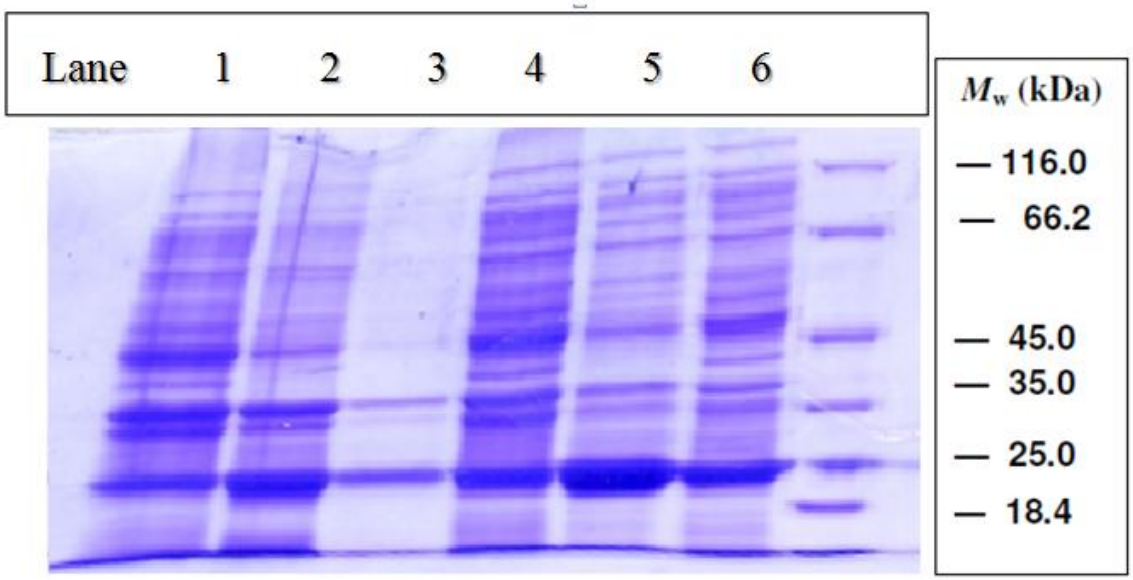


Figure 3.4:A 12% polyacrylamide SDS-PAGE gel showing the results of the induction studies performed to optimize production of R13A hGSTA1-1 protein. *Escherichia coli* BL21 (DE3) pLysS cells transformed with R13A pKHA1 plasmid were induced under various conditions the cells were lysed and their contents evaluated. Lanes 1-3 are the pellet fraction of the lysed cells while lanes 4-6 are the supernatant fractions. Lanes 3 and 6 show the contents of cells induced under the same conditions as the wild-type (1mM IPTG, 37 °C). Lanes 1 and 4 show the contents of cells induced with 1 mM IPTG but incubated at 25 °C. Lanes 2 and 5 show the contents of cells induced with 0.1 mM IPTG and incubated at 37 °C. The thickest protein band is at roughly 25 kDa (corresponding to GST) in lane 5, meaning that the conditions of 0.1 mM IPTG and incubation at 37 °C are optimal for mutant plasmid expression.

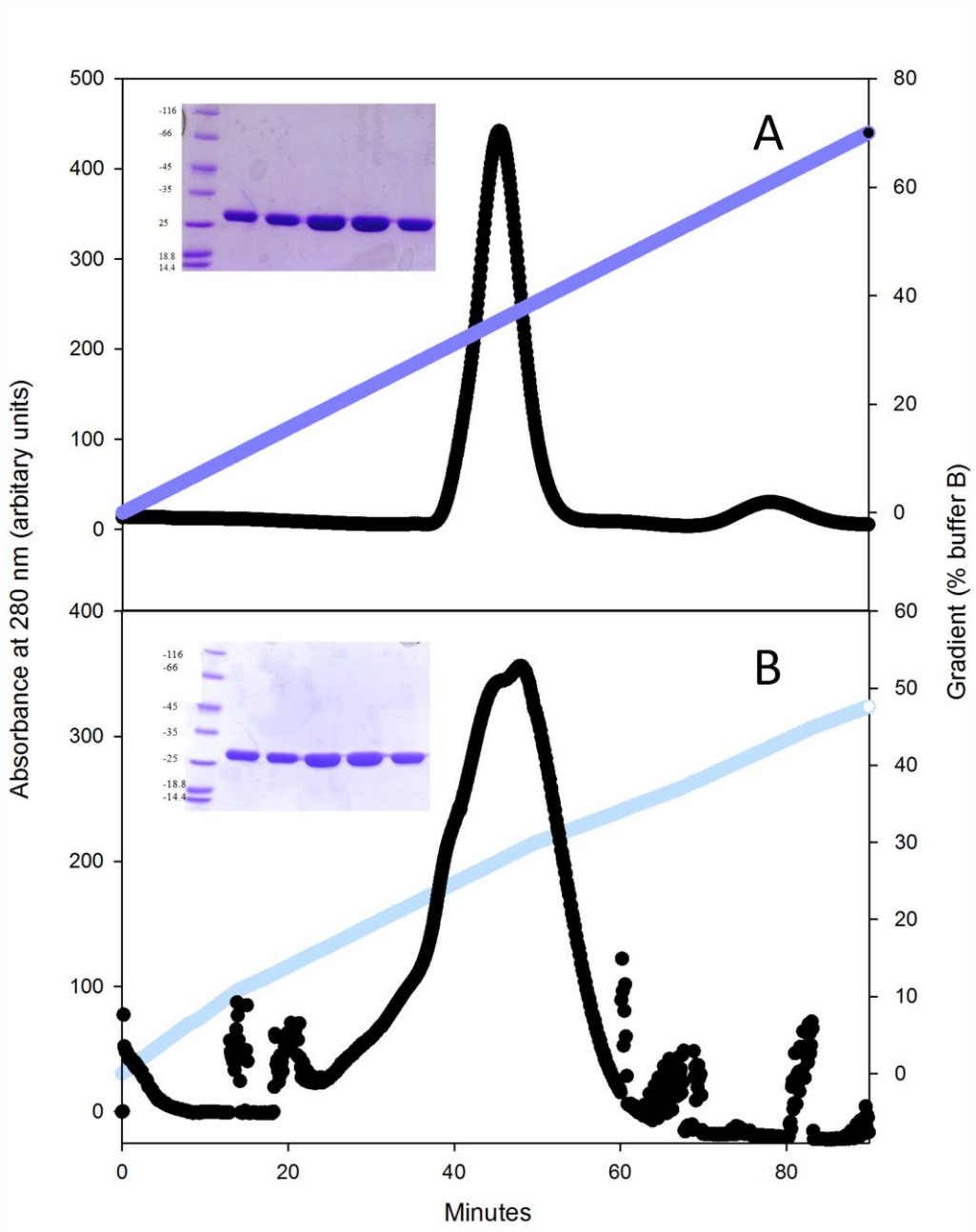


Figure 3.5: Elution profiles of the R13A variant and wild-type hGSTA1-1 proteins from purification. The A280 of effluent (black) and the 0-300 mM NaCl salt gradient (blue) were recorded. (A) Elution of wild-type and (B) R13A hGSTA1-1. The proteins were eluted in 10 mM sodium phosphate buffer, pH 7.45, containing 1 mM EDTA and 0.02% sodium azide. Inserts of SDS-PAGE analysis of selected fractions of the eluted protein taken at regular intervals, proving the proteins are electrophoretically pure are shown.

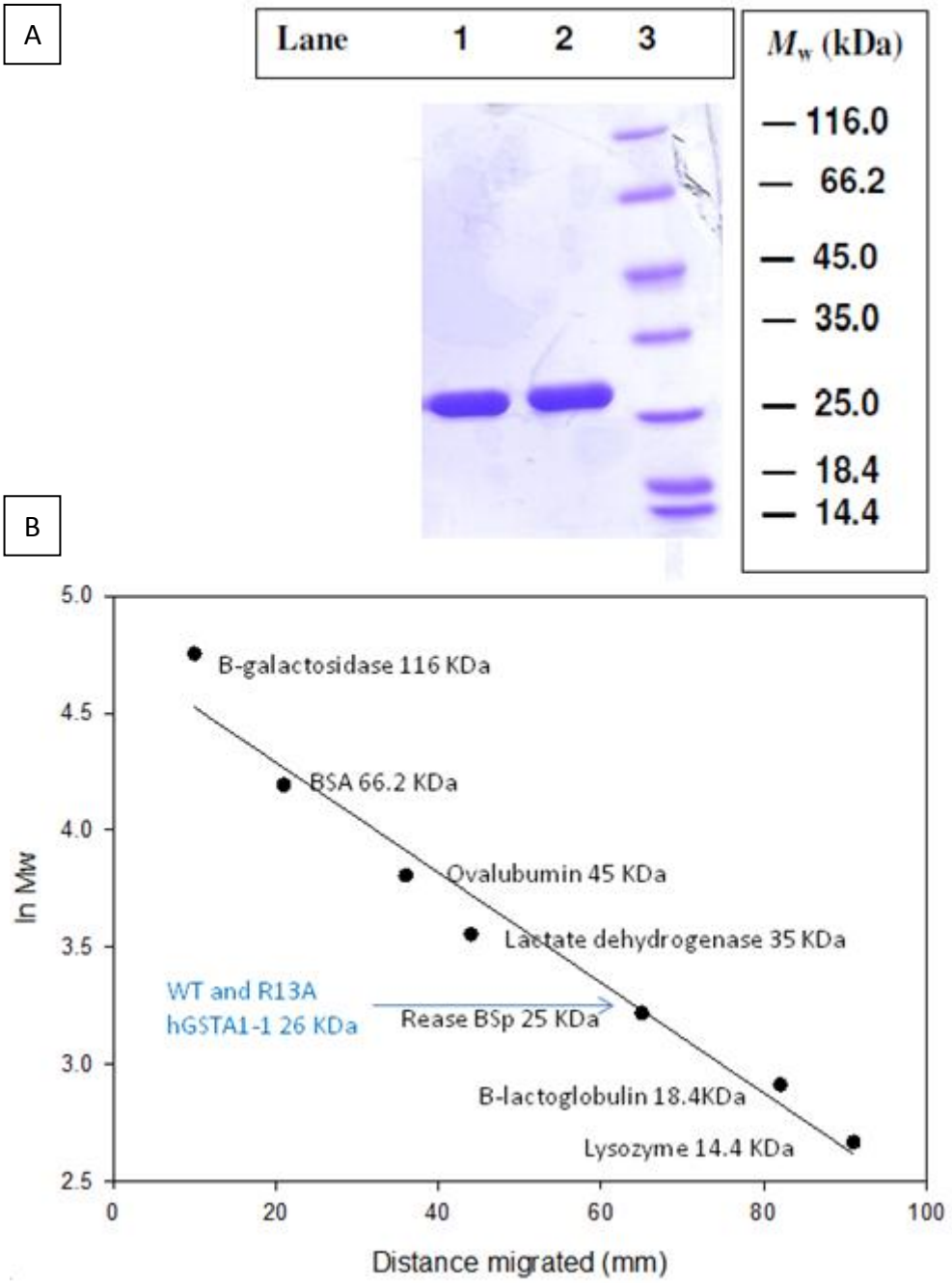


Figure 3.6: SDS-PAGE analysis of the purity and size of the R13A and wild-type hGSTA1-1 subunits. (**A**) SDS-PAGE gel depicting the denatured wild-type (lane 1) and R13A variant (lane 2) hGSTA1-1 subunits run alongside molecular weight markers (lane 3). The molecular weight markers used were β-galactosidase (11 6 kDa), bovine serum albumin (66.2 kDa) ovalbumin (45.0 kDa), lactate deydrogenase (35.0 kDa), REase BSp (25.0 kDa), β -lactoglobulin (18.4 kDa) and lysozyme (14.4 kDa). (**B**) Calibration curve constructed by plotting the natural log of the molecular weight (kDa) of the proteins against their distance migrated (mm). The molecular weight of the wild-type and variant proteins was calculated using linear regression analysis (R^2 = 0.96). The results show that the proteins were pure and had a molecular weight of 26 kDa which is indicative of GST.

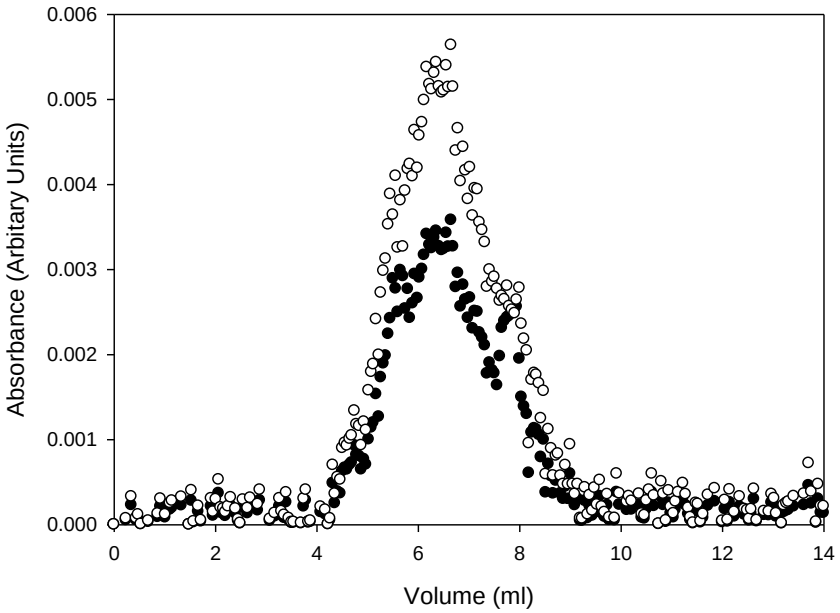


Figure 3.7: SE-HPLC of wild-type and R13A hGSTA1-1. The elution profiles for wild-type (white) and R13A hGSTA1-1 (black) show that both proteins have similar hydrodynamic volumes and that neither protein aggregated. Proteins were eluted at a flow rate of 0.5 ml.min^{-1} at an isocratic pressure of 40 bar. The elution was monitored by absorbance spectroscopy at A_{280}.

3.3. Structural properties of the R13A variant and wild-type proteins

3.3.1. Secondary structural characterisation

The far-UV CD (180 - 250 nm) spectra of folded and chemically denatured (8 M urea) R13A and wild-type hGSTA1-1 are illustrated in Figure 3.8. These spectra are strong indicators of the peptide backbone and overall secondary structure of the proteins (Woody, 1995). The high alpha-helix content of hGSTA1-1, 9 α-helixes (60-65%) (Sinning *et al*,. 1993) and only 4 β-strands, are reflected by the spectra displaying two ellipticity minima at 222 and 208 nm (Woody, 1995). The strong peak at 190 nm is the result of the absorption of circularly polarized light electrons in the carbonyl π-bonding orbital of the peptide-backbone amide bonds (Woody, 1995). There is little change between the spectra of the variant and wild-type protein, thus it may be concluded that the mutation at domain interface of hGSTA1-1 did not induce a gross conformational change in the secondary structure. At 8 M urea both the variant and wild-type CD spectra become featureless losing their distinctive trough and peak. A featureless CD spectrum is characteristic of a fully denatured, unfolded protein that has undergone the loss of all its secondary structural components (i.e. the protein is absent of helices and β-strands and now exists in a random coil conformation).

3.3.2. Tertiary structural characterisation

3.3.2.1. Intrinsic fluorescence

The tertiary structure of R13A and wild-type hGSTA1-1 was analysed using the intrinsic fluorescence spectroscopy of Trp21 and the tyrosine residues of the proteins. Trp21 is found at the domain interface of hGSTA1-1 and is part of a lock-and-key motif that is occluded inside a hydrophobic pocket. This in addition to its proximal location to the H-site and the fluorescence properties of its indole-ring side chain which is influenced by the polarity of its environment make it an excellent probe for monitoring changes in its local environment.

The fluorescence spectra of folded and chemically denatured (8 M urea) R13A and wild-type hGSTA1-1 are illustrated in Figure 3.9. The excitation of the tryptophan and tyrosine residues by irradiation at 280 nm yielded fluorescence emission maximums at 326 nm for the wild-type and 329 nm for the R13A variant hGSTA1-1 proteins.

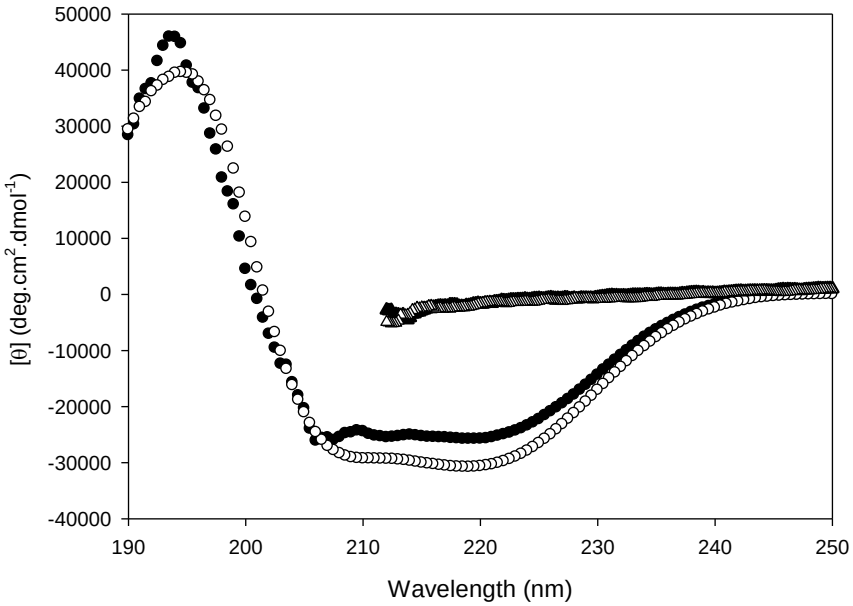


Figure 3.8: Far-UV circular dichroism spectra of R13A and wild-type hGSTA1-1. Spectra are shown for the native (○) and unfolded (▼) (8 M urea) forms of wild-type (white) and R13A (black) hGSTA1-1. Spectra were collected using samples of 2 µM protein in GST storage buffer, pH 7.5. The similar spectra indicate that both proteins have similar secondary structure.

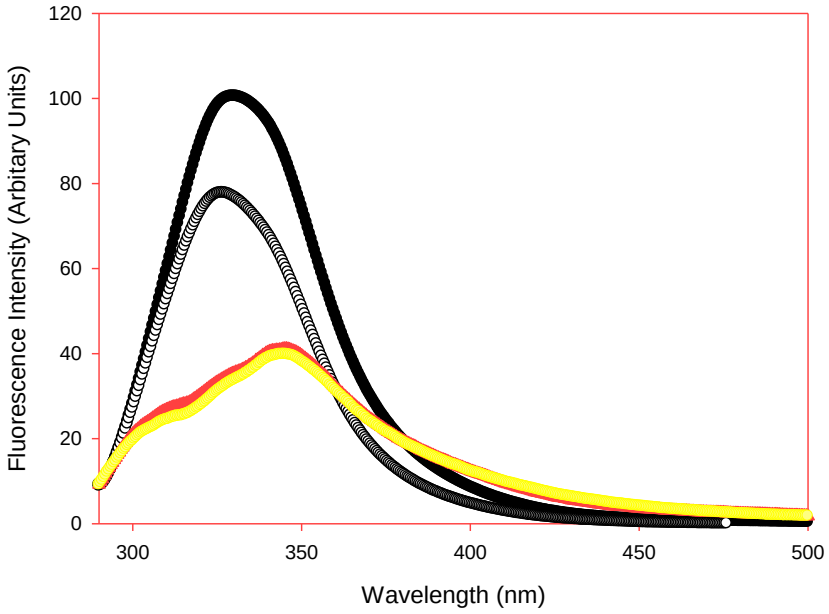


Figure 3.9: Fluorescence emission spectra of wild-type and R13A variant hGSTA1-1. Emission spectra shown for the native forms of wild-type (white) and R13A (black) hGSTA1-1 as well as for the unfolded forms (in 8 M urea) of wild-type (yellow) and R13A (red) hGSTA1-1. Spectra were collected using 2 µM protein in GST storage buffer, pH 7.5. The R13A variant hGSTA1-1 fluoresces with far greater intensity. The fluorescence emission maximum for both the variant and wild-type protein shifted to 347 nm upon the denaturation of the proteins

The different intensities between the emission spectra of the native R13A variant and wild-type proteins indicate a change in the ability of the tertiary environment to quench the fluorescence energy emitted by the excited tryptophan residue. It cannot be due to difference in concentration because under those circumstances the difference in intensity would be mirrored under the denaturing conditions which it is not. Furthermore, the 3 nm red shift in the maximum emission wavelength between the emission spectra of the native R13A variant and wild-type proteins indicates that the tertiary environment of Trp has become slightly more polar or exposed to solvent in the R13A variant.

Together the difference in fluorescent intensity between R13A and wild-type hGSTA1-1 coupled with the red shift of the fluorescence emission maximum of R13A indicates changes to the local tertiary environment of Trp21 in the R13A variant.

3.3.2.2. Extrinsic fluorescence – ANS binding

The binding of ANS to R13A and wild-type hGSTA1-1 was studied in order to probe the effects of the mutation on the H-site of the enzyme. ANS was shown to bind the H-site of hGSTA1-1, the hydrophobic anilino and naphthyl rings occupying the H-site while the polar sulfonate moiety becomes bound at the interface of the G- and H-sites (Dirr *et al.*, 2005).

Fluorescent spectra for 200 μM ANS in the absence and presence of wild-type and R13A hGSTA1-1 is illustrated in Figure 3.10. Free ANS is seen to emit maximally at 518 nm whereas ANS bound to the wild-type and variant protein emitted maximally at 487 and 484 nm respectively. The emission maximum of ANS bound to native wild-type hGSTA1-1 is in close agreement with the 480 nm wavelength maximum reported in the literature (Sluis-Cremer *et al.*, 1996; Dirr and Wallace, 1999; Sayed *et al.*, 2000). Both emission maxima are consistent with the burial of the ANS residue in the highly hydrophobic environment of the H-site. The maximum emission wavelength of bound ANS is sensitive to the hydrophobicity of its binding site (H-site) and the fluorescence of the dye will undergo increasingly stronger blue shifts as the hydrophobic character of the binding site increases (Slavik, 1982). The

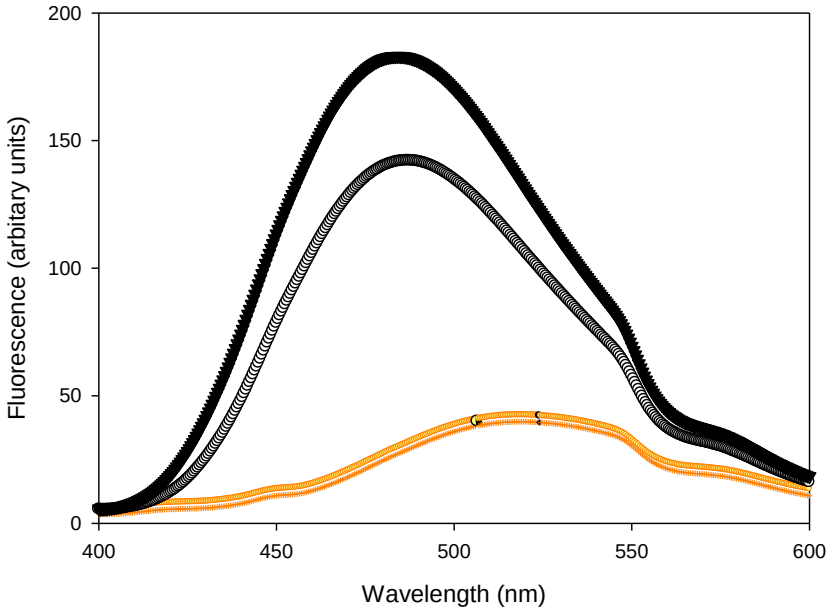


Figure 3.10: Fluorescence emission spectra of free, unbound ANS and ANS bound to the wild-type and R13A variant hGSTA1-1. Spectra of 200 µM ANS bound to 2 µM wild-type (grey) and R13A variant (black) protein as well as free, unbound ANS (orange). ANS was selectively excited at 390 nm and the spectra recorded over the 400-600 nm range using an excitation and emission bandwidth of 5 nm. Each spectrum is the average of 3 accumulations of 3 replicate samples in GST storage buffer. The spectra of protein bound ANS were corrected for the fluorescence contribution from free, unbound ANS.

maxima emission wavelengths for ANS bound to the wild-type and the variant protein therefore indicates firstly that the H-site is partially polar, consistent with previous studies and, secondly, that the H-site of the R13A variant protein is slightly more occluded than the wild-type (Sayed *et al.*, 2002).

Upon binding to the proteins the fluorescence intensity (quantum yield) of ANS is increased and increased especially when it was bound to the R13A variant of hGSTA1-1. The quantum yield of bound ANS is highly sensitive to conformational changes in its binding environment (Dirr and Wallace, 1999) and is prone to quenching by water (Matulis and Lovrien, 1998), indicating that ANS is more strongly shielded from water in the R13A variant or that the H-site has been altered by the R13A mutation. Alternatively, it could mean that the R13A variant has greater affinity for ANS than the wild-type protein and therefore the increase in intensity would be the result of more ANS bound to protein.

3.4. Specific activity: GSH-CDNB conjugation assay

The GSH-CDNB conjugation assay (section 2.10.) (Habig *et al*., 1974) is a sensitive and rapid test that may be used to determine the specific activity of a GST enzyme. The assay was performed at pH 6.5 because at this pH the non-catalysed reaction is minimal.

Immediately upon the addition of the 1 mM CDNB substrate the formation of the yellow product 1-(*S*-glutathionyl)-2,4-dinitrobenzene) began and there was a linear increase in the A_{340} reading for both the wild-type and variant proteins (Figure 3.11). The linear slope indicates that the only limitation on the catalytic activity is the ability of the enzyme to catalyze the reaction and that the assay is therefore suitable to calculate specific activity. The slope of each progress curve provides a precise reading of the reactions initial velocity (v_0) for each specified concentration of the enzyme. The v_0 of both the R13A and wild-type hGSTA1-1 was plotted against their respective concentrations of protein (Figure 3.12). The slope of a straight line plotted through these points will therefore be the specific activity. A systematic error is observed for these plots, namely that the intercepts of the plotted lines do not pass through the origin. The calculated results are nevertheless not affected by the error as the specific activities of the enzymes are calculated from the slopes. It is possible the error is the result of an under correction having been made for the non-enzymatic reaction.

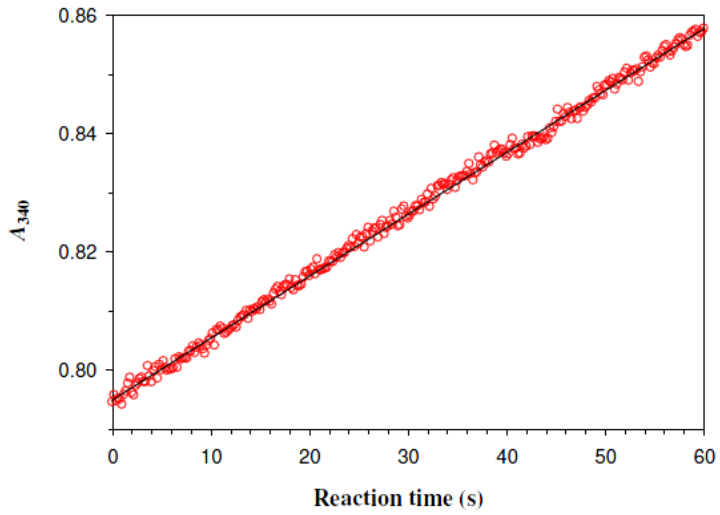


Figure 3.11: Progress curve of the conjugation reaction of GSH with CDNB as catalysed by R13A hGSTA1-1 protein. The monitored production of 1-(*S*-glutathionyl)-2,4-dinitrobenzene, as measured by an increasing A_{340} reading, immediately following initiation of the assay reaction by the addition of 1 mM CDNB to the buffered protein solution for a period of one minute. The change in absorbance at 340 nm over the period of one minute was determined by linear regression and used to calculate the initial velocity (v_0) (R^2 = 0.995)

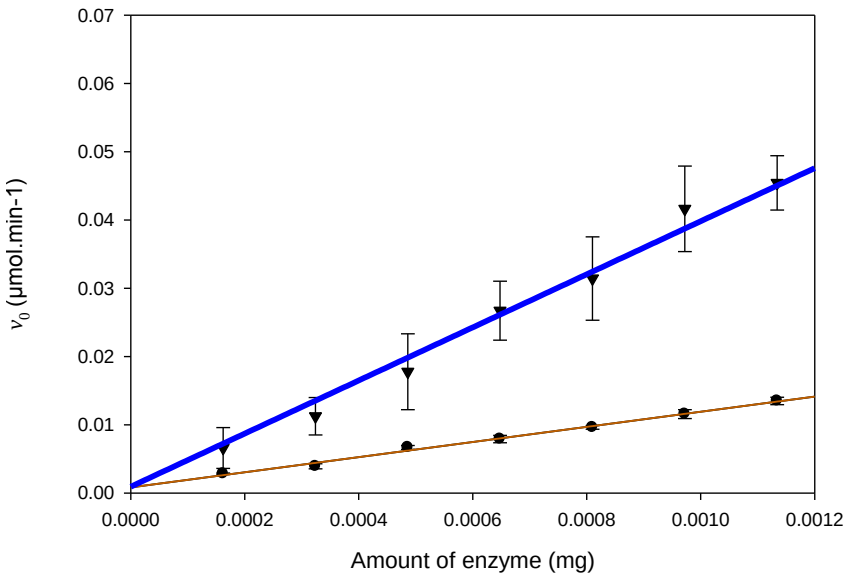


Figure 3.12: Specific activity of wild-type and R13A variant hGSTA1-1 as determined using the GSH-CDNB conjugation assay. The conjugation of CDNB to GSH by enzyme catalysis was monitored by measuring the A_{340} of the formed product 1-(*S*-glutathionyl)-2, 4-dinitrobenzene (ε_{340} = 9600 $M^{-1}.cm^{-1}$). The reaction was performed 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 and 1 mM CDNB. The assay was corrected for the non-catalysed reaction. The specific activity of each enzyme is the slope of the linear regression plots of wild-type (blue). Values are the average of three replicates and error bars showing the Standard deviation are shown.

The specific activity was determined to be 32.32 (± 1.04) μmol.min^{-1}.mg^{-1} for wild-type and 11.08 (± 0.32) μmol.min^{-1}.mg^{-1} for R13A variant (the values in brackets are the calculated std error of the specific activity). This means the variant enzyme loses 70% of its specific activity as compared to the wild-type protein (Figure 3.12). The wild-type value was in agreement with those reported in the literature: 34.9 μmol.min^{-1}.mg^{-1} (Gildenhuys *et al.*, 2010), 39.42 (± 0.21) μmol.min^{-1}.mg^{-1} (Mosebi *et al.*, 2003) and 44.6 (± 2.3) μmol min^{-1}.mg^{-1} (Nathaniel *et al.*, 2003). Likewise the findings agree with those of Stenberg *et al*. (1991) who found that the R13A variant lost 90% of its specific activity as compared to wild-type hGSTA1-1.

The findings indicate that the disruption of the domain interface by breaking the bond between Arg13 and Glu169 does indeed compromise the active site of the enzyme.

3.5. Conformational stability of R13A and wild-type hGSTA1-1

3.5.1. Thermal-induced unfolding

Thermal unfolding was conducted on wild-type and R13A hGSTA1-1. The change in the $[\Theta]_{222}$ values of the proteins were recorded in response to a uniform change in temperature over the range of (20-80 °C). The $[\Theta]_{222}$ is proportional to the α-helical content of a protein and hence a good indicator with which to monitor hGSTA1-1 protein as it is mostly α-helical in nature (Sinning *et al*., 1993).

The proteins were observed to unfold in a monophasic ans highly co-operative sigmoidal manner (Figure 3.13). The thermal-unfolding curves of variant and wild-type proteins do not overlay indicating a change in protein stability. Wild-type hGSTA1-1 unfolded between 58 ºC and 64 ºC while R13A hGSTA1-1 unfolded between 53 ºC and 58 ºC. Normally, thermal unfolding studies are used to calculate the melting temperature (T_m), enthalpy ($\Delta H(T\text{m})$), and heat capacity increment (ΔC_p) of the proteins so that they may be used to determine the protein stability function ($\Delta G_{(T)}$) (Privalov 1979; Pace *et al*., 1989). With the exception of the T_m value however, this is only possible when the unfolding of the protein is thermodynamically reversible. This is often not the case due to the tendency for proteins to aggregate after unfolding at high temperature (Benjwal *et al*., 2005) which is what happens in the case of hGSTA1-1 (Wallace *et al*., 1998).

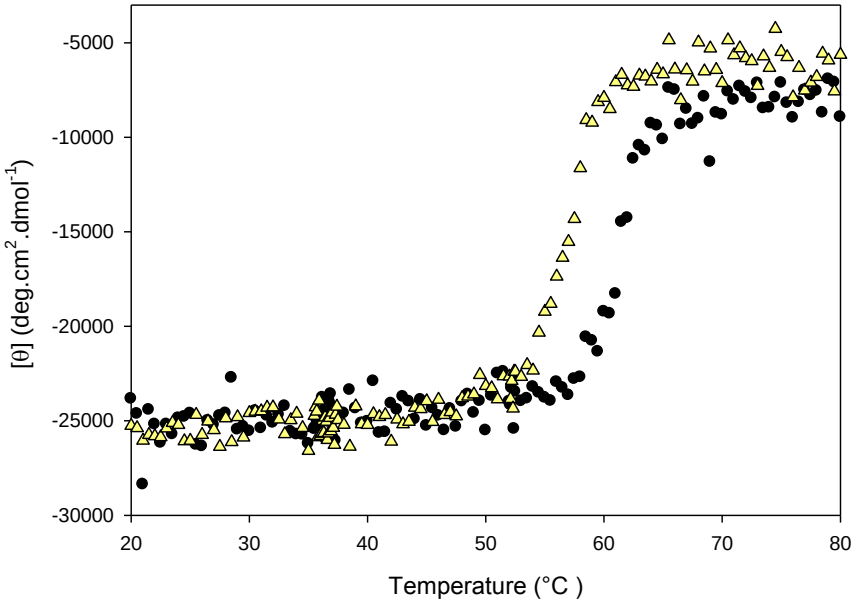


Figure 3.13: Thermal-induced unfolding transitions of wild-type and R13A variant hGSTA1-1. The thermal unfolding curve of wild-type (black) and R13A variant (yellow) hGSTA1-1. Progression of the unfolding transitions was monitored by following the $\mathbf{[\Theta]}_{222}$ which is proportional to the α-helical content of a protein. The readings were made with 2 µM protein using a 1 °C.min^{-1} gradient in 20 mM sodium phosphate buffer, pH 7.5, containing 1 mM EDTA and 0.02% sodium azide.

The T_m value for the variant protein (T_m = 57 °C) was lower than that for the wild-type (T_m = 63 °C). The lower melting temperature observed for mutant protein indicates that the global protein conformation is destabilized because of the loss of the Arg13-Glu169 salt bridge at the domain interface.

3.5.2. Chemical denaturant – induced unfolding

3.5.2.1. Reversibility of urea-induced unfolding

Protein unfolding transitions can only be analysed in terms of their thermodynamic parameters if the unfolding reaction is reversible and takes place under equilibrium conditions (Pace, 1986). It has been established that the unfolding reaction of wild-type hGSTA1-1 is reversible (Wallace *et al.*, 1998) but there is no published data in regards to the unfolding reaction of R13A hGSTA1-1.

The recovery of unfolded R13A hGSTA1-1 to its folded state was established by far-UV CD ($[\Theta]_{222}$) and intrinsic tryptophan fluorescence (Ex_{280} Em_{326}) as secondary and tertiary structure probes respectively (Figure 3.14 and 3.15). The R13A variant was unfolded in the presence of 8 M urea for an hour before the urea was diluted 10-fold with storage buffer and the protein was allowed to refold. Refolded R13A hGSTA1-1 protein showed 90% recovery as determined by both the far-UV CD and intrinsic tryptophan fluorescence probes.

However, recovery does not indicate reversibility. For a true equilibrium to be established the protein cannot possess the property of hysteresis, dependence of a system not only on its current environment but also on its past environment. In other words a protein must be shown to refold along the same pathway it unfolded. This was established by unfolding the variant in the presence of 8 M urea before diluting it with appropriate concentrations of urea and storage buffer such that the final solution had the identical protein and urea concentrations used to create the unfolding curve (section 2.11.2.1.). The refolding pathway of R13A hGSTA1-1 is illustrated in Figure 3.16. The refolding pathway can be seen to overlay the unfolding pathway.

Together, this evidence suggests that both wild-type and R13A variant hGSTA1-1 proteins can establish a state of equilibrium between their folded and unfolded states in the presence

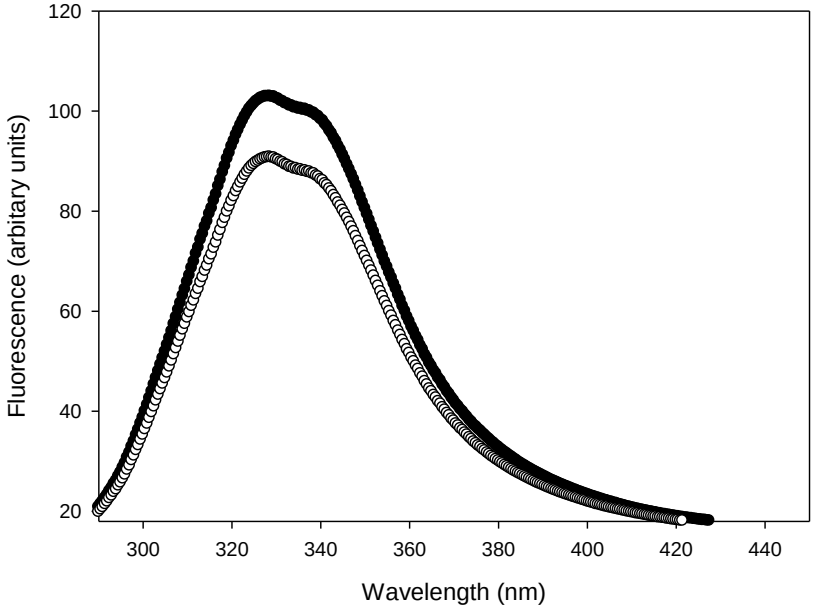


Figure 3.14: Recovery from unfolding of R13A monitored by fluorescence. R13A hGSTA1-1 protein in its native (black) and refolded (white) conformations. Folded R13A had a maximum fluorescence intensity of 103 while refolded R13A had a fluorescence intensity of 93.3. This means refolded R13A has 90% of the fluorescence intensity of the native implying only 90% of the protein successfully refolded.

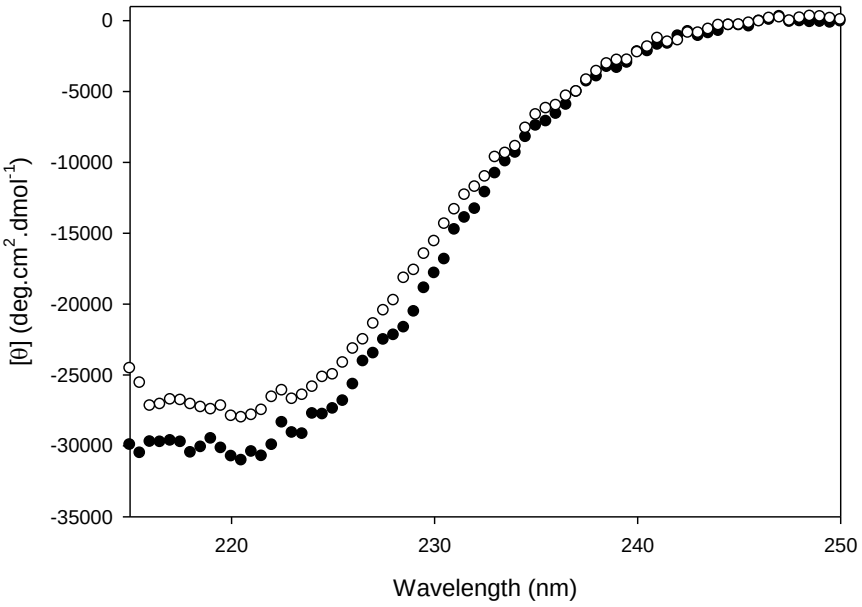


Figure 3.15: Recovery from unfolding of R13A monitored by circular dichroism. R13A hGSTA1-1 protein in its native (black) and refolded (white) conformations. Native R13A had an $[\Theta]_{222}$ of -30743 while refolded R13A had an $\mathbf{[\Theta]}_{222}$ of 27815. This means refolded R13A has only 90.4% of the CD intensity of the native protein implying only 90% of the protein successfully refolded.

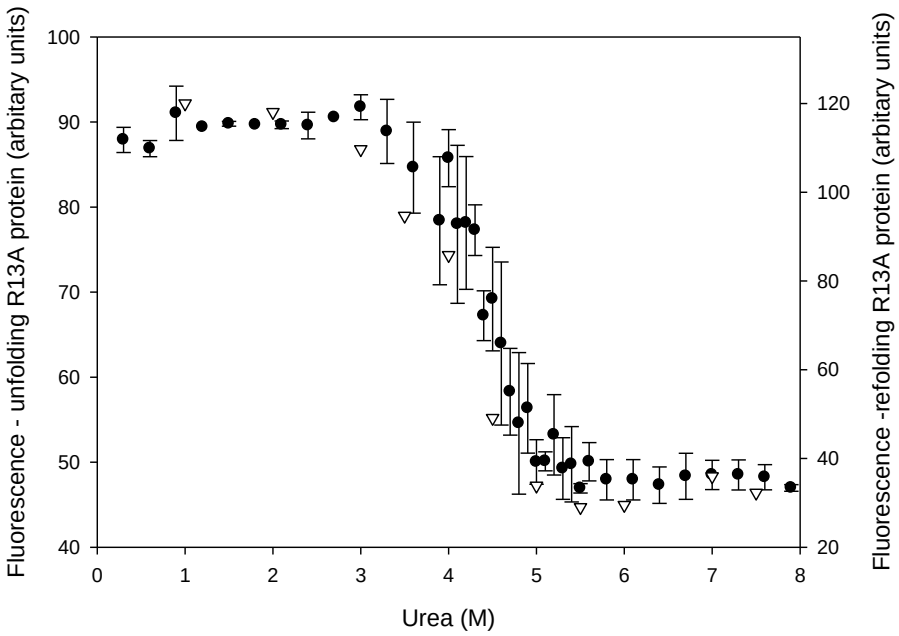


Figure 3.16: Reversibility of unfolding of R13A hGSTA1-1 monitored by fluorescence. Overlay of an unfolding curve (black) with a refolding curve (white) of R13A hGSTA1-1 as monitored by fluorescence spectroscopy. Values are the average of three replicates and error bars showing the std deviation are shown.

of urea. Thus the unfolding transitions of wild-type and R13A hGSTA1-1 were analysed to determine the thermodynamic parameters of unfolding which will indicate the conformational stability of the hGSTA1-1 proteins.

3.5.3. Urea-induced equilibrium unfolding in the presence and absence of ANS

The observed urea-induced unfolding curves for wild-type and R13A hGSTA1-1 in the absence of ANS as determined by both CD and fluorescence probes is shown in Figure 3.17. Additionally in figure 3.17, Rayleigh scatter plots for both proteins depicting only small, unsystematic light scattering between 0 and 8 M urea indicate the absence of aggregates over the unfolding range.

The unfolding curves of both proteins in the absence of ANS are sigmoidal and monophasic, indicative of a highly co-operative and two-state unfolding curve (Privalov, 1992). The unfolding curves for R13A hGSTA1-1 in the absence of ANS as monitored by fluorescence are not co-incident which is indicative that the unfolding pathway is not two-state. Additionally wild-type and R13A hGSTA1-1 were exposed to the hydrophobic dye ANS in the presence of increasing concentrations of urea (0-8 M) (section 2.11.2.3), since ANS is an excellent probe for the detection of intermediates on the unfolding pathway due to its ability to bond to exposed hydrophobic patches of a protein. ANS is doubly useful in this instance because the hGSTA1-1 protein also binds ANS to its H-site. ANS bound to hGSTA1-1 was found to maximally fluoresce at 487 nm. The binding of ANS to wild-type and R13A variant hGSTA1-1 in the presence of urea was monitored spectroscopically by fluorescence (Ex_{390} Em_{465}) as a function of urea concentration. The resulting unfolding curves are illustrated in Figure 3.17. Whereas the unfolding curve of neither protein in the presence of ANS indicates the formation of an unfolding equilibrium intermediate the unfolding curves for both proteins do show a drop in the fluorescence intensity of bound ANS at very low concentrations of urea. This is likely the result of ANS being released from the H-site as the protein unfolds (Alves *et al*., 2006).

A

B

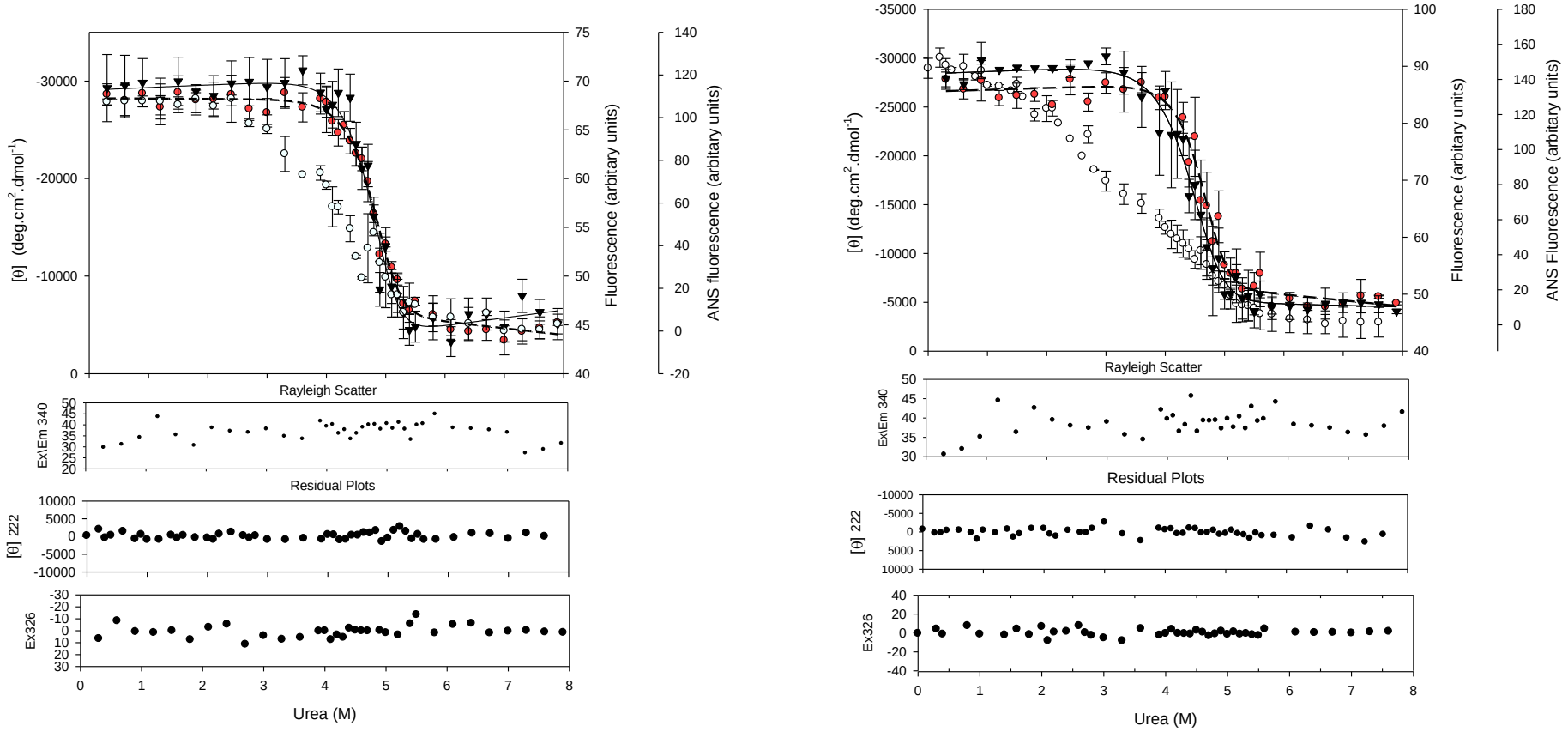


Figure 3.17: Urea-induced equilibrium unfolding of wild-type and R13A hGSTA1-1 in the presence and absence of ANS. Equilibrium unfolding curves of (A) wild-type and (B) R13A hGSTA1-1 as monitored by fluorescence intensity in the absence (black) and presence (red) of ANS and CD (white). Rayleigh light scattering plots (Ex_{340}/Em_{340}) showing there is no aggregation and plots of the residuals of fits for the fluorescence (Ex_{326}) and CD ($[\Theta]_{222}$) monitored unfolding curves are shown in the lower panels.

The variant enzyme loses the ability to bind ANS at the H-site at even lower concentrations of urea than the wild-type which is an indication that it is less stable. For both proteins however, this indicates that protein unfolding is taking place at much lower concentrations of urea than either the fluorescence or CD probes indicate in the absence of the ANS ligand. Indicating that neither the wild-type nor R13A hGSTA1-1 unfold via a two-state process. This finding agrees with a growing number of papers (Wallace *et al*., 1998) that indicate that the unfolding of wild-type is in fact three-state and that the unfolding curves in the absence of ANS as monitored by CD and fluorescence are indicative of only the second transition of the three-state unfolding mechanism, the unfolding of a dimeric intermediate (N_2^*) to unfolded monomer (2U) (Wallace *et al*., 1998).

Therefore the unfolding curves as monitored by the probes of CD and fluorescence were fitted using a two state model:

$$N_2 \rightarrow 2U$$

even though the evidence is suggestive of a three-state mechanism. The free energy change in the absence of denaturant ($\Delta G_{(H2O)}$), the dependence of free energy on denaturant concentration (*m*-value) and the midpoint of the unfolding transition (C_m) were calculated through use of Sigma Plot V11.0 (Systat Software Inc; Chicago, IL, USA)) (section 2.11.2.3.). The thermodynamic parameters are summarised in Table 4. Attempts were made to calculate thermodynamic parameters using three-state models but no appreciable fit could be obtained from the data ($R^2 < 0.5$).

The calculated parameters indicate that the R13A variant protein is indeed less stable than the wild-type. The mutation has caused a decrease in the conformational stability of the protein as well as shifted the midpoint of the unfolding transition (C_m) to a lower concentration of urea. There is little change in the slopes of the unfolding transitions (*m*-value) of R13A variant and wild-type protein according to the far-UV CD probe but there is a significant decrease as determined by the intrinsic fluorescence probe. This indicates a decrease in the dependence of the free energy change of unfolding upon denaturant concentration (Pace *et al*., 1989) which only the fluorescence probe is sensitive to.

Table 4: Thermodynamic parameters of two-state equilibrium unfolding for both wild-type and R13A hGSTA1-1
The values obtained are from fitting the equilibrium unfolding data to a two–state dimeric model ($N_2 \rightarrow 2U$). All values were calculated from three replicates and the numbers in parentheses represent the standard error.

	Far-UV CD		**Intrinsic Fluorescence**		***Average***	
	Wild - Type	**R13A**	**Wild - Type**	**R13A**	**Wild - Type**	**R13A**
$\Delta G(H_2O)$ ($kJ.mol^{-1)}$	100.42 (± 6.27)	84.52 (± 8.37)	106.36 (± 1.26)	93.14 (± 8.37)	103.38	88.82
***m*-value ($kJ.mol^{-1}.M^{-1}$ urea)**	14.77 (± 1.26)	13.72 (± 1.76)	16.07 (± 0.21)	12.34 (± 1.67)	15.4	13.01
C_m (M urea)	5	4.7	4.9	4.5	4.95	4.6
R^2	0.9833	0.959	0.946	0.946	N/A	N/A

3.6. R13A hGSTA1-1 Crystallization

All attempts to crystallize the R13A variant protein failed. In all circumstances, barring one, the protein aggregated before crystallizing. Only under the conditions of 0.1 M HEPES pH7.0, 30% v/v Jeffamine ED-2001 pH 7.0 (buffer 39, Hampton index HR2-144) did crystals form but they were not suitable for X-ray diffraction. Nucleation was promoted over crystal growth and many small crystallites were formed rather than one large crystal.

CHAPTER 4. Discussion and conclusion

4.1. A conserved interdomain salt bridge in the Y-GSTs

This study shows that the Arg13 residue of hGSTA1-1 is conserved throughout examples of the Alpha, Mu, Pi, *Plasmodium falciparum* and Sigma GSTs (Figure 3.1). Additionally, in each case the role of the Arg residue is the same, to form a salt bridge interaction between domain 1 and 2 of a GST subunit (Figure 3.2). This finding is not surprising given that the Alpha, Mu, Pi and Sigma classes of GST have been shown to be amongst the most recently evolved classes of GST and that they are all closely related (Pemble and Taylor, 1992; Tomarev *et al*., 1993). In fact it has been proposed by Atkinson and Babbitt (2009) that these four GST classes form a major subgroup of the GSTs which they termed the 'Y-GSTs'. In addition to higher structural and sequence similarity amongst each other, compared to the other GSTs, they all use a highly conserved tyrosine residue (Tyr9) to interact with the sulfhydryl group of GSH to promote catalysis (section 1.5.1.). That the Arg residue is also conserved in the *Plasmodium falciparum* GST class is not surprising as there is evidence that it too should be considered a Y-GST (Fritz-Wolf *et al*., 2003). The role of the conserved Arg residue in *Plasmodium falciparum* is nonetheless unique to this class. Fitting, since *Plasmodium falciparum* is an entirely unique class of GST that exists as an inactive tetramer or active dimer depending on the environmental conditions of the protein (Fritz-Wolf *et al*., 2003; Tripathi *et al*., 2007; Liebau *et al*., 2009). In *Plasmodium falciparum* GSTs the Arg13 residue of the enzyme may form either an interdomain salt bridge or an intersubunit interaction depending on whether the enzyme is in its tetramer or dimer configuration (Fritz-Wolf *et al*., 2003).

While the Arg residue and its role were highly conserved in the classes Alpha, Mu, Pi, *Plasmodium falciparum* and Sigma GST, the residue on domain 2 with which they form the salt bridge is not. In some cases it was a Glu and in others an Asp residue. Moreover, the location**s** of either residue had far greater topological freedom in domain 2 (Figure 3.2). This too is not surprising since it is known that the N-terminal domain (the thioredoxin containing domain 1) is more conserved (Armstrong, 1997; Ladner *et al*., 2004). This is proposed to be the result of two factors (1) the greater need for specificity for the binding of GSH by the N-terminal domain and (2) the evolutionary pathway for GSTs has been proposed to have been

the addition of a second domain (the C-terminal domain) to a domain ancestor of the N-terminal domain which had a thioredoxin fold and whose role it was to recognize and bind GSH (Ladner *et al*., 2004).

It is also interesting to note that in the case of the other GSTs, the C/S-GSTs, it is their vital to catalysis Ser or Cys residues which were found in the place of the Arg residue (Figure 3.1). This observation is in agreement with the finding that the position of the active site has shifted slightly between the Y-GSTs and the C/S-GSTs major subgroups (Atkinson and Babbitt, 2009).

In all observed cases of Y-GSTs the Arg residue is conserved and forms a salt bridge interaction, providing evidence that the Arg residue and its salt bridge interaction are important for the proper catalytic function in the Y-GSTs. It is likely that the salt bridge interaction is important for maintaining the proper conformation of the active site loop and thus the correct positioning of the catalytic Tyr residue. Thus it appears to perform a very similar role to the equally well conserved hGSTA1-1 Arg15 residue of the active site loop (Gildenhuys *et al*., 2010; Dobreva, MSc Dissertation, 2005 (http://wiredspace.wits.ac.za/handle/10539/1519)). Unlike the Arg15 residue, the Arg13 residue does not seem to carry out a direct role in the enzymes catalysis (Gildenhuys *et al*., 2010; Dobreva, MSc Dissertation, 2005 (http://wiredspace.wits.ac.za/handle/10539/1519)).

The conserved Arg residue in the Y-GST major subgroup of the GST superfamily of proteins located at the domain interface was therefore mutated to an Ala in hGSTA1-1. The choice of residue eliminated the bulk of the side chain of Arg as well as its positive charge that interacts with Glu169 in domain 2 of hGSTA1-1 thus breaking a domain-domain salt bridge interaction. The resulting R13A hGSTA1-1 variant makes an excellent model to study the importance of the domain interface of the Y-GSTs. In particular it helps to elucidate the role of the conserved Arg13-Glu169 interdomain salt bridge interaction in hGSTA1-1 and provides insight into the role of the equivalent interactions in the other classes of the Y-GSTs

4.2. The R13A mutation does not affect the global structure of the enzyme

The mutation does not cause any major structural alterations in the native state of the protein. Far-UV circular dichroism spectroscopy showed that the secondary structure of hGSTA1-1 was not affected (section 3.3.1.). Further, the R13A hGSTA1-1 variant was shown to be enzymatically active, though its specific activity was lower than the wild-type, indicating that proper dimerisation did indeed take place and hence for much of the protein the tertiary structure is intact as well (3.4.). This may be inferred because cytosolic GSTs are only active as dimers (Dirr *et al*., 2004; Fabrini *et al*., 2009). Analysis by SE-HPLC also confirmed that dimerisation took place (section 3.2.2.). This means the reduction in specific activity observed for R13A hGSTA1-1 cannot be attributed to large-scale structural rearrangements in the protein. Instead, evidence (section 4.3.) points to localized structural rearrangements at and around the H-site as well as the important domain interface Trp21 residue (section 3.3.2.). It is likely that it is these small localized structural rearrangements which are responsible for the changes in the catalytic activity of the variant protein.

4.3. Loss of the Arg13-Glu169 salt bridge compromises the active site of hGSTA1-1

Steady-state kinetic data revealed a substantial decrease of 70-90% in the specific activity of the variant enzyme compared to that of the wild-type (section 3.4.; Stenberg *et al*., 1991). Further, it was found that the K_M of the variant for GSH was eight fold higher while the K_M for CDNB was 1.4 times higher than that of the wild-type (Stenberg *et al.*, 1991). Additionally inhibition studies by Stenberg *et al.* (1991) found that the variant protein had a lower IC_{50} value for the active-site inhibitors *S*-hexyl glutathione and BSP (Kolobe *et al*., 2004) compared to the wild-type. A decrease in the IC_{50} value indicates that a smaller concentration of inhibitor is required to suppress activity and therefore the affinity of the binding site for inhibitor has increased.

The findings of this study argue against the postulate that the role of Arg13 may have been the neutralization of the negative charge of the carboxylate groups of GSH (section 1.5.8.) (Stenberg *et al*., 1991). Rather, it suggests that the R13A mutation and the resultant disruption of the domain-domain interface alters the structure and hence function of the

active-site of hGSTA1-1. Specifically by looking at all the experimental results together it seems clear that it is the H-site, which the Arg13 residue is near, and not the G-site that is most affected.

The increased affinity of the R13A variant for *S*-hexyl glutathione (Stenberg *et al*., 1991), the binding of which is dominated by the hexyl substituent and not by the peptide moiety of GSH, compared to the wild-type is the first indication that possible structural changes at the H-site are dominant over possible changes at the G-site. This finding is reinforced by the R13A variant GSTs preference for BSP another H-site inhibitor (Stenberg *et al*, 1991; Kolobe *et al*., 2004). Note, while strictly speaking BSP is a non-competitive inhibitor, normally meaning the inhibitor does not bind the active site, and that it has two binding sites in hGSTA1-1 BSP (Kolobe *et al*., 2004) it is only its second, low affinity binding site at the H-site of the enzyme which influences enzyme activity. Furthermore, though it binds to the H-site it is classed as a non-competitive inhibitor because it does not stop CDNB binding to the H-site as well. The large and hydrophobic H-site is able to accommodate both BSP and substrate simultaneously. In fact the H-site of hGSTA1-1 is able to accommodate three molecules of BSP at once (Kolobe *et al*., 2004).

Also, it can be seen from ligand binding studies using ANS (section 3.3.2.2.), a ligand known to bind at the H-site (Dirr *et al.,* 2005), that ANS shows increased fluorescence when bound to the R13A variant compared to that of the wild-type. The quantum yield of bound ANS fluorescence has been shown to be both susceptible to quenching by water (Matulis and Lovrien, 1998) as well as sensitive to conformational changes in its binding environment (Dirr and Wallace, 1999). Hence it is not possible to know whether the increased quantum yield is due to an increased affinity between R13A hGSTA1-1 and ANS or if it is due to structural changes at the H-site. However, either possibility is indicative of a change taking place at the H-site of the protein. A good future experiment would be to calculate, using ITC, the K_d of ANS at the H-site to clear up the ambiguity of the results.

More evidence indicating that structural changes are taking place at the H-site and not the G-site was revealed by using Trp21 as a local probe of its surrounding tertiary structure (section 3.3.2.). The fluorescence spectra of Trp21 showed an increase in the quantum yield of Trp fluorescence for the variant protein as well as a slight red shift. The shift indicates that the Trp residue is more solvent exposed in the variant protein. The increase in quantum yield

found for the variant is more difficult to explain but as the Arg13 residue is located near the Trp21 it may be just that its replacement with Ala in the variant leads to less quenching of Trp fluorescence. The positive amino groups of the Arg side chain are able to quench the florescence of the indole ring of Trp via electron transfer (Muino and Callis, 2008; Osysko and Muino, 2011) or even proton transfer (Yu *et al*., 1992). Since Trp21 is located closely to the H-site of the enzyme and it is vital in forming a domain-domain interaction as part of a highly conserved lock-and-key motif any change in its environment can have important consequences. Specifically, flexibility around Trp21 has been indicated in altering the substrate promiscuity and catalytic function of the H-site in hGSTA1-1 (Hou *et al.,* 2007). The finding that the R13A alters the Trp21 lock-and-key motif is not surprising given that the Glu169 residue which forms a salt bridge with Arg13 is on helix 6 and that several of the residues which form the hydrophobic pocket into which Trp21 is inserted are also present on helix 6 (Figure 1.5).

While it may seem contradictory that the active site of the variant R13A hGSTA1-1 has compromised productive binding of CDNB and GSH but maintains the ability to bind ANS, BSP and S-hexyl glutathione, and with stronger affinity, this observation correlates with the promiscuous nature of the H-site and an altered environment around Trp21(Hou *et al.,* 2007). A similar observation has been seen before; Balchin *et al*. (2010) discovered that a W21A variant hGSTA1-1 did not have compromised functionality at the G-site but the ability of the H-site to bind CDNB was compromised.

To explain the eight fold increase in the K_M of the variant for GSH as compared to the wild-type protein it is necessary to discuss the dynamics of helix 9 and its impact on the catalysis of GST enzymes (section 1.5.5.). The evidence (section 4.4.) points not to a change in the binding capacity of GSH but rather to a decrease in the ability of the enzyme to convert the GSH substrate to product. The K_M value is determined by both the physical step of substrate binding and the chemical conversion of substrate to product (k_{cat}). Indeed it can be seen that the k_{cat}/K_M (catalytic efficiency) value for GSH as found for R13A hGSTA1-1 was over twenty times lower than that of the wild-type (Stenberg *et al*., 1991).

4.4. The R13A mutation influences helix 9 dynamics

The unfolding transition curve of R13A hGSTA1-1 in the presence of ANS (section 3.5.3.) showed that the variant lost the ability to bind ANS at the H-site at lower concentrations of urea compared to the wild-type (Figures 3.17). A similar effect, which was attributed to greater helix 9 flexibility, with respect to ANS binding was observed in the M51A and M51A/F52S variants studied by Alves *et al*. (2006). This suggests that the dynamics of helix 9 is affected by the R13A mutation.

In support of this, it has been previously shown that helix 9 is a critical component of the H-site (Mosebi *et al.*, 2003; Dirr *et al.*, 2005; Nilsson, *et al.*, 2002; Nielanik *et al.*, 2001) which this study has shown has been structurally altered. Moreover, ANS has been shown to be a sensitive probe of C-terminal and helix 9 dynamics (Dirr and Wallace, 1999). That the mutation affects helix 9 is plausible in light of the position of Arg13 at the domain interface. The guanidinium group of Arg13 projects from the active-site loop that is formed between β-strand 1 and helix 1 of domain 1 towards Glu169 on helix 6 to form a salt bridge. Elimination of this interaction would likely affect the conformation of the loop region. Phe10 which interacts with the Ser212, Glu214, Ala216, Ile 219 and Phe220 residues of helix 9 (Dirr and Wallace, 1999) is a part of this loop and the disruption of these interactions have been seen to increase the flexibility of helix 9 (Ibarra *et al*., 2001). Additionally since it has been shown that the Trp21 lock-and-key motif is altered it means that the position of helix 8, which is part of the Trp21 lock-and-key motif, must also be affected and since helix 8 is directly connected to helix 9 this would lead to a change in its conformation flexibility (Figure 1.5). This was found to be the case in the W21A variant hGSTA1-1 (Baclhin *et al*, 2010). New conformational flexibility at the Trp21 lock-and-key motif in the R13A variant could therefore easily induce a conformational change in the C-terminal helix.

Though none of the residues in helix 9 contribute directly to the chemical mechanism of catalysis in hGSTA1-1 (Alves *et al.*, 2006) changes at the C-terminus have been shown to influence catalysis indirectly in a number of ways that maintain the favourable pK_a of Tyr9 at the active site (section 1.5.5.). Additionally, experiments have shown the apparent on-rate for glutathione binding to GST A1-1 is 450 $mM^{-1}.s^{-1}$ (Gustafson *et al.*, 1999). This binding rate is lower than the diffusion controlled value that could be expected for an enzyme substrate association. When hGSTA1-1 is mutated such that helix 9 becomes less stable or truncated

from the protein GSH has been shown to have much greater access to the G-site and the on-rate of binding for GSH increased (Gustafson *et al.*, 1999). Nonetheless it has been seen for several such mutations (F220A, F220T, M208K and M208E, α9del) that while the on rate of GSH binding is increased, the K_d of GSH binding to enzyme remains relatively unaltered and in each case the specific activity of the enzyme also become compromised (Dirr and Wallace, 1999; Gustafson *et al.*, 1999).

It is possible that the increased conformational flexibility of helix 9 is responsible for the lower specific activity of the R13A variant and is also the reason why the K_M of GSH is increased for the variant. The affinity between GSH and hGSTA1-1 may not have been compromised at all; the compromised ability of Tyr9 to stabilize the GS^- anion and promote catalysis might be responsible for the decreased catalytic activity of the enzyme. This is not the first time such a finding has been reported, studies on the R15L mutation (Gildenhuys *et al.*, 2010) showed similar findings.

4.5. Role of Arg13 in the stability of hGSTA1-1

Over expression of R13A and wild-type HGSTA1-1 encoding pKHA1 resulted in a lower yield of variant protein as compared to the wild-type as expected (section 3.2.1.; Stenberg *et al*., 1991). This may be the first indication of decreased protein stability but without further examination it is impossible to rule out decreased stability or translatability of the mRNA of the variant as the cause. Given that both the wild-type and variant protein were found to be soluble in the supernatant fraction of the lysate and there was little change to the global structure of the proteins it is reasonable to conclude that protein folding was unaffected.

The stability of the wild-type and R13A proteins was assessed by temperature-denaturation (section 3.5.1.) and urea-induced equilibrium unfolding studies in the presence and absence of ANS (section 3.5.3.). Temperature-induced denaturation of the hGSTA1-1 proteins were irreversible as shown for other GSTs (Kaplan *et al*., 1997; Dragani *et al*., 1998; Wallace and Dirr., 1999) and hence thermodynamic parameters could not be calculated. Nevertheless, a comparison of the two T_m values for the proteins (temperature at which the protein is half unfolded) are a clear indicator that the R13A variant is less stable than the wild-type enzyme.

A decrease of 6 ºC was observed for the T_m of R13A variant hGSTA1-1 compared to that of the wild-type (Figure 3.13).

Equilibrium unfolding curves for both R13A and wild-type GSTA1-1 indicated that the ligandin function of the R13A hGSTA1-1, the ability to bind ANS at the H-site, was altered as compared to the wild-type. Moreover, as the equilibrium unfolding curves for both proteins in the presence of ANS were not co-incident with their recorded spectroscopic unfolding transition curves in the absence of ANS, it indicates that neither protein unfolds via a two-state mechanism. A three-state fit nevertheless could not be fitted to the data suggesting that the fluorescence and CD probes are not sensitive enough to detect the intermediate.

This agrees with a growing amount of evidence suggesting that the unfolding of hGSTA1-1 is three-state (Wallace *et al*., 1998; Wallace *et al*., 2000; Mosebi *et al*., 2003; Alves *et al*., 2006). The unfolding process of R13A hGSTA1-1 is therefore perhaps better represented by the formula (Wallace *et al*., 1998):

$$N_2 \rightarrow N_2^* \rightarrow 2U$$

where N_2^* is a transiently formed dimeric intermediate along the path of unfolding. The dimeric intermediate most likely represents the same species detected using stopped flow unfolding kinetics for wild-type protein (Wallace *et al*, 1998). The transient dimeric intermediate N_2^* for wild-type was suggested to be the result of structural changes at the domain interface and definitely around Trp21, since the probe of Trp fluorescence is most sensitive to it (Wallace *et al*, 1998; Wallace *et al*., 2000). It is also thought to be the result of the destabilization or unfolding of helix 9 as shown by equilibrium unfolding curves using catalytic activity as a probe (Dirr and Wallace, 1999; Wallace *et al*., 2001; Mosebi *et al*., 2003). This matches up with what has already been seen of the R13A variant, namely that it disrupts and causes local structural changes at the domain interface and definitely at and/or around Trp21 and that it seems to increase the conformation flexibility of helix 9. It is worth noting, despite the flexibility of helix 9, evidence has shown that helix 9 does not affect the overall conformational stability of hGSTA1-1 (Dirr and Wallace, 1999). The removal of helix 9 has been shown to hardly affect the stability of the protein. Hence the $N_2^* \rightarrow 2U$ transition

is essentially representative of the global unfolding of the proteins and is what the unfolding curves of R13A and wild-type hGSTA1-1 as monitored by CD and fluorescence in the absence of ANS represent (Wallace *et al*., 1998; Dirr *et al*., 1999; Mosebi *et al*., 2003; Alves *et al*.,2006). Therefore the urea-induced equilibrium unfolding curves of R13A and wild-type hGSTA1-1 monitored by fluorescence and CD indicate the conformational stability of the variant in the absence of denaturant was ~17 kJ.mol^{-1} lower than that of the wild-type.

The determined *m*-value for the equilibrium unfolding curve of the R13A variant monitored by fluorescence was significantly lower than the *m*-value determined from any of the other unfolding curves. According to the linear energy model (Pace, 1975) as proteins unfold in the presence of a chemical denaturant there is a linear relationship between its change in stability and urea concentration. This linear relationship is represented by the *m*-value or slope of the unfolding transition (Myers *et al.*, 1995). The decreased *m*-value for the R13A variant indicates that it is unfolded over a broader range of urea concentrations (Figure 3.17) and may be explained through three possible reasons (Wallace *et al.*, 1998):

I. There has been a change in the exposed surface area of the native and/or unfolded state of the R13A variant protein.
II. An intermediate has become significantly populated during the equilibrium unfolding of the R13A variant, meaning that the protein no longer displays two state unfolding .
III. The mutation has altered the interaction of the denaturant with the protein in the denatured state.

The first reason is unlikely as such a change would show a co-incidental change in the far-UV CD and fluorescence unfolding curves, as previously mentioned was not the case. It also cannot be the third reason since both chemical denaturants urea and guanidinium chloride have been shown to unfold GSTs completely and both denaturants yield similar $\Delta G_{(H2O)}$ values for unfolding suggesting that there are no specific binding sites for either denaturant on the protein for the folded or unfolded conformation (Dirr and Reinemer, 1991). This leaves the second reason, the accumulation at equilibrium of a likely partially unfolded dimeric intermediate of the R13A variant during unfolding. It must be a dimeric intermediate because it has long been established for wild-type hGSTA1-1 that the subunits of the protein cannot be separated into distinct folding units (Wallace *et al*., 1998; Fabrini *et al*., 2009). The lower *m*-value determined for the fluorescence equilibrium unfolding curve of R13A

hGSTA1-1 suggest that the fluorescence probe is more sensitive to the dimeric intermediate than CD and that the first phase of unfolding has changed for the R13A variant protein. That the fluorescence probe was most sensitive to the dimeric intermediate adds additional support to the finding that the intermediate is the result of structural changes at the domain interface and definitely around Trp21 (Wallace *et al*., 1998, Wallace *et al*., 2000).

X-ray crystallography of the R13A variant would be a powerful tool that would help elucidate the exact structural changes that have resulted and to explain the loss in protein stability. Unfortunately all attempts to crystallize the R13A variant failed and X-ray diffraction studies could not be performed (section 3.6.).

4.6. Conclusion

This study found that the highly conserved domain interface Arg13 residue, and its salt bridge interaction with Glu169, is far more important to the function of the hGSTA1-1 enzyme than it is to either the global structure or stability of the protein. The loss of the Arg13-Glu169 salt bridge interaction does not disrupt proper protein dimerisation and the loss in stability is below what is expected for the loss of a salt bridge interaction. The evidence shows that Arg13-Glu169 salt bridge is important in maintaining the correct conformation of the active site loop, the structure of the H-site and the proper conformation of the C-terminal helix 9, all of which have a role in maintaining the proper catalytic function of the enzyme. In particular the loss of the Arg13-Glu169 interdomain salt bridge compromises the ability of the Tyr9 residue to lower the pK_a of the thiol group of cysteine in GSH and hence prime it for nucleophilic attack on the substrate at the H-site. Given that the Alpha, Mu, Sigma, Pi and *Plasmodium falciparum* classes of GST all share the highly conserved domain interface residue Arg13 which forms a salt bridge contact with a Glu or Asp residue on domain 2, and that these classes all use a highly conserved Tyr residue as a central part of their catalytic mechanisms, it cannot be coincidence. Therefore, it is reasonable to postulate that the highly conserved Arg residue is of vital importance to the catalytic function of these classes of GST. Like it does in hGSTA1-1, the Arg residue probably maintains the proper conformation of the active site loop and hence the proper orientation between the critical to catalysis Tyr residue and the GSH substrate in these classes of GST. However, the C-terminal helix 9 that is so critical in explaining the effects of the

R13A mutation on hGSTA1-1 is unique to class Alpha GSTs it is difficult to speculate if the equivalent mutation in the other GST classes would affect the catalysis of the enzymes as strongly.

CHAPTER 5. References

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. (1999) Regulation of JNK signaling by GSTp. *EMBO J.* **18**, 1321–1334.

Adman, E. T., Le Trong, I., Stenkamp, R. E., Nieslanik, B. S., Dietze, E. C., Tai, G., Ibarra, C. and Atkins, W. M. (2001) Localization of the C-terminus of rat glutathione *S*-transferase A1-1: crystal structure of mutants W21F and W21F/F220Y. *Proteins* **42**, 192-200.

Altschul, S., Gish, W., Miller, W., Myers, E. and Lipman, D. (1990). Basic local alignment search tool. J. *Mol. Biol.* **215**, 403–410.

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 Met^{51} and Phe^{52} in function and dimer stability. *Biochem. J.* **393**, 523-528.

Andújar-Sánchez, M., Smith, A. W., Andújar-Sánchez J.M., Rodriguez-Vico, F., Heras-Vazquez, F.J., Jara-Pérez, V. and Cámara-Artigas, A. (2005) Crystallographic and Thermodynamic Analysis of the Binding of *S*-Octylglutathione to the Tyr 7 to Phe mutant of Glutathione *S*-Transferase from *Schistosoma* ***Biochemistry***, *44*, 1174–1183.

Anfinsen, C. B. (1973) Principles that Govern the Folding of Protein Chains. *Science* 181, 223-230.

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

Atkinson, H.J. and Babbitt, P.C. (2009) Glutathione Transferases Are Structural and Functional Outliers in the Thioredoxin Fold. *Biochemistry **48**, 11108-11116.*

Balchin, D., Dirr, H.W. and Sayed, Y. (2011) Energetics of ligand binding to human glutathione transferase A1-1: Tyr-9 associated localisation of the C-terminal helix is ligand-dependent. *Biophys Chem* **156**, 153–158.

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

Dragani, B., Iannarelli, V., Allocati, N., Masulli, M., Cicconetti, M. and Aceto, A. (1998) Irreversible thermal denaturation of glutathione transferase P1-1. Evidence for varying structural stability of different domains. *Int. J. Biochem. Cell B.* ***30***, 1155–1163.

Bennet, M.J., Choe, S. and Eisenberg, D. (1994) Domain swapping: Entangling alliances between proteins. *Proc. Natl. Acad. Sci.* **91**, 3127-3131.

Benson, A.M., Talalay, P., Keen, J.H. and Jokoby. W.B. (1977) Relationship between the soluble glutathione-dependent delta 5-3-ketosteroid isomerise and the glutathione S-transferases of the liver. *Proc. Natl. Acad. Sci. USA*. **74**, 158-162.

Birdsall, B., King, R.W., Wheeler, M.R., Lewis, C.A., Godde, S.R., Dunlap, R.B. and Roberts, G.C.K. (1983) Correction for light absorption in fluorescence studies of protein ligand interactions. *Anal. Biochem*. **132**, 353-361.

Birnboim, H.C. and Doly, J. (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucl. Acid Res.* 7, 1513–1523.

Board, P. G., Baker, R. T., Chelvanayagam, G. and Jermiin, L. S. (1997) Zeta, a novel calss of glutathione transferases in a range of species from plants to humans. *Biochem. J.* 328, 929-935.

Board, P.G. and Mannervik, B. (1991) The contribution of the C-terminal sequence to the catalytic activity of GST2, a human Alpha-class glutathione transferase. *Biochem. J.* **275**, 171-174.

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**, 24798-24806.

Boyland, E. and Chasseaud, L.F. (1969) The role of glutathione and glutathione *S*transferases in mercapturic acid biosynthesis. *Adv. Enzymol.* **32**, 173-219.

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

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**, 717-727.

Chayen. N.E. (2004). Methods for separating nucleation and growth in protein crystallisation. *Prog Biophys Mol Biol*, **88**, 329-337

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. (2001)Glutathione S-transferase Mu modulates the stress-activated signals by suppressing apoptosis signal-regulating kinase 1. *J. Biol. Chem*. **276**, 12749-12755

Chung, C.T., Niemela, S.L. and Miller, R.H. (1989) One-step preparation of competent *Escherichia coli*: Transformation and storage of bacterial cells in the same solution. *Proc. Natl. Acad. Sci. U.S.A* **86**, 2172-2175.

Copley, S.D. and Dhillon, J.K. (2002). Lateral gene transfer and parallel evolution in the history of glutathione biosynthesis genes. *Genome Biol*. **3**, 0025.1-0025.16

Cornish-Bowden, A. (1995). *Fundamentals of Enzyme Kinetics* 3rd ed, pp 71, Portland Press, London.

Da Fonseca, R. R., Johnson, W. E., J O'Brien, S., Vasconcelos, V. and Antunes, A. (2010) Molecular evolution and the role of oxidative stress in the expansion and functional diversification of cytosolic glutathione transferases. *BMC Evol. Biol.* 10, 281-292.

Danger, D.P., Baldwin, W.S. and LeBlanc, G.A. (1992) Photoaffinity labelling of steroidhormone- binding glutathione S-transferases with [3H] methyltrienolone. Inhibition of steroid-binding activity by the anticarcinogen indole-3-carbinol. *Biochem. J.* **288**, 361-367.

Dill, A. K. (1990). Dominant forces in protein folding. *Biochemistry* **29**, 7133-7151.

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

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**, 15631-15640.

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

Dirr, H., Reinemer, P. and Huber, R. (1994) X-ray crystal structure of cytosolic glutathione S-transferases. Implications for protein architecture, substrate recognition and catalytic function. *Eur. J. Biochem.* **220**, 645-661.

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,** 19480-19487.

Dobreva, M, A: Thermodynamics of Complexation: Variant R15L hGST A1-1 binding to ANS and $GSO3^-$**.** *MSc dissertation*. University of the Witwatersrand, Molecular and Cell Biology; 2005.

Dragani, B., Stenberg, G., Melino, S., Petruzzelli, R., Mannervik, B. and Aceto, A. (1997) The conserved N-capping box in the hydrophobic core of glutathione S-transferase P1-1 is

essential for refolding. Identification of a buried and conserved hydrogen bond important for protein stability. *J Biol Chem* **272**, 25518-25523.

Ellman G. L. (1959) Tissue sulfhydryl groups. *Arch. Biochem. Biophys.* **82**, 70–77.

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.

Erhardt, J. and Dirr, H. (1995) Native dimer stabilizes the subunit tertiary structure of porcine class Pi glutathione S-transferase. *Eur. J. Biochem.* **230**, 614-620.

Fabrini, R., De Lica, 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***, 10473–10482.

Fersht, A. (1972). equilibria in a- and d- chymotrypsin. The energetics and importance of the salt bridge. *J. Mol. Biol*. **64**, 497-509.

Fersht, A. (1999) *Structure and Mechanism in Protein Science. A Guide to Enzyme Catalysis and Protein Folding*, pp 132- 168, W. H. Freeman & Co., New York.

Fritz-Wolf. K., Becker, A., Rahlfs, A., Harwaldt, P., Schirmer, R. H., Kabsch, W., and Becker, K. (2003) X-ray structure of glutathione S-transferase from the malarial parasite Plasmodium falciparum. *PNAS* **100**, 13821-13826.

Garcia-Saez, I., Parraga, A., Phillips, M. F., Mantle, T. J. and Coll, M. (1994) Molecular structure at 1.8 A of mouse liver class Pi glutathione *S*-transferase complexed with S-(pnitrobenzyl) glutathione and other inhibitors. *J. Mol. Biol.* **237**, 298-314.

Geerlof, A., Coutard, B.B., Egloff, M.-P., Enguita, F.J., Fogg, M.J., Gilbert, R.J.C., Groves, M.R., Haouz, A., Nettleship, J.E., Nordlund, P., Owens, R.J., Ruff, M., Sainsbury, S., Svergun, D.I. and Wilmanns, M. (2006). The impact of protein characterization in structural proteomics. *ACTA Crystallogr D*, 62, 1125-1136

Gildenhuys, S. Wallace, L.A., Dirr, H.W. (2008) Stability of Reduced *Escherichia coli* Glutaredoxin 2: A Monomeric Structural Homologue of the Glutathione Transferase Family. *Biochemistry*. **47**, 10801-10808.

Graminski, G. F., Kubo, Y. and Armstrong, R. N. (1989) Spectroscopic and kinetic evidence for the thiolate anion of glutathione at the active site of glutathione *S*-transferase. *Biochemistry* **28**, 3562-3568.

Guex, N. and Peitsch, M.C. (1997) SWISS-MODEL and the Swiss-Pdb viewer: an environment for comparative protein modelling. *Electrophoresis* **18**, 2714-2723.

Gustafsson, A., Etahadieh, M., Jemth, P. and Mannervik, B. (1999) The C-terminal region of human glutathione transferase A1-1 affects the rate of glutathione binding and the ionization of the active-site Tyr9. *Biochemistry* **38**, 16268-16275.

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

Habig, W.H., Pabst, M.J. and Jakoby, W.B. (1974) Glutathione S-transferases. The first enzymatic step in mercapturic acid formation. *J. Biol. Chem.* **249**, 7130-7139.

Habig, W.H., Pabst, M.J., Fleischner, G., Gatmaitan, Z., Arias, I.M. and Jakoby, W.B. (1974a) The identity of glutathione S-transferase B with ligandin, a major binding protein of liver. *Proc. Natl. Acad. Sci. U. S. A.* **71**, 3879-3882.

Han, J., Batey, S., Nickson, A. A., Teichmann, S. A. and Clarke, J. (2007) The folding and evolution of multidomain proteins. *Nature* **8**, 319-328.

Harding, S.E. & Chowdhry, B.Z. (2001). (eds). *Protein-Ligand interactions: structure and spectroscopy*, Oxford university press, Oxford.

Hoesch, R.M. and Boyer, T.D. (1989) Localization of a portion of the active site of two rat liver glutathione S-transferases using a photoaffinity label. *J. Biol. Chem.* **264**, 17712-17717.

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**, 12336-12344.

Hornby, T. A. J., Luo, J-K., Stevens, M. J., Wallace, A. S., Kaplan, W., Armstrong, N. R., and Dirr, W. H. (2000) Equilibrium folding of dimeric class μ glutathione transferases involves a stable monomeric intermediate. *Biochemistry* **39**, 12336-12344.

Hou, L., Honaker, M.T., Shireman, L.M., Balogh, L.M., Roberts, A.G., Ng, K., Nath, A and Atkins, W.M. (2007) Functional promiscuity correlates with conformational heterogeneity in A-class glutathione *S*-transferases. *J. Biol. Chem.* **282**, 23264-23274.

Huang, J., Koide, A., Makabe, K. and Koide, S. (2008) Design of protein function leaps by directed domain interface evolution. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 6578-6583.

Ibarra C., Nieslanik B. S. and Atkins W. M. (2001) Contribution of aromatic– aromatic interactions to the anomolous pKa of Tyr-9 and the C-terminal dynamics of glutathione S-transferase A1-1. *Biochemistry* **40**, 10614–10624.

Ish-Horowicz, D. and Burke, J.F. (1981) Rapid and efficient cosmid cloning. *Nucl. Acid Res.* **9**, 2989–2998**.**

Jaenicke, R. (1999) Stability and folding of domain proteins. *Progr. Biophys. Mol. Biol.* **71**, 155-241.

Johansson, A.S. and Mannervik, B. (2001) Interindividual variability of glutathione transferase expression in man, in Interindividual variability in drug metabolism in man (Pacifici, G.M. and Pelkonen, O. eds) Taylor and Francis, London, in press.

Jones, S. and Thornton, M.J. (1995). Protein-protein interfaces: a review of protein dimmer structures. *Prog. Biophys. Molec. Biol.* **63**, 31-65.

Jones, S. and Thornton, M.J. (1996) Principles of protein-protein interactions. *Proc. Natl. Acad. Sci. U.S.A.* **93**. 13-20.

Jones, S., Marin, A. and Thornton, J. M. (2000) Protein domain interfaces: characterisation and comparison with oligomeric protein interfaces. *Protein Eng* **13**, 77-82.

Kaplan, W., Husler, P., Erhardt, J., Sluis-Cremer, N. and Dirr, W. H. (1997) Conformational stability of pGEX-expressed *Schistosoma japonicium* glutathione Stransferase. *Protein Sci.* **6,** 399-406.

Keskin, O., Ma, B., and Nussinov, R. (2005) Hot-regions in protein-protein interactions: the organization and contribution of structurally conserved hot spot residues. *J. Mol. Biol.* **345**, 1281-1294.

Kolobe, D., Sayed, Y., and Dirr, H.W. (2004) Characterization of bromosulphophthalein binding to human glutathione S-transferase A1-1: thermodynamics and inhibition kinetics. *Biochem. J*. **382**, 703–709.

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 crystalographyc and Calorinmetric Study. *J. Mol. Biol.* **349**, 825-838.

Kumar, S and Nussinov, R. (1999) Salt bridge stability in monomeric proteins. *J. Mol. Biol.* **293**, 1241-55.

Ladner, J. E., Parsons, F. J., Rife, L. C., Gilliland, L. G., and Armstring, N. R. (2004). Parallel evolutionary pathways for glutathione transferases: structure and mechanism of the mitochondrial class kappa enzyme rGSTK1-1. *Biochemistry* **43**, 352-361.

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

Lakowicz, J.R and Weber, G. (1973) Quenching of fluorescence by oxygen: Detection of structural fluctuations in proteins on the nanosecond time scale. *Biochemistry* **12**, 4171-4179.

Lakowicz, J.R. (1999) *Principles of fluorescence spectroscopy*. Plenum Press, New York, USA.

Larsen, T.A., Olson, A.J., and Goodsell, D.S. 1998. Morphology of protein-protein interfaces. *Structure.* **6:** 421–427.

Lawrence, R.A. and Burke, R.F. (1997). Glutathione peroxidase activity in selenium-deficient rat live. *Biochem. Biophys. Res. Commun*. **71**, 952-958.

Le Trong, I.L., Stenkamp, R.E., Ibarra, C., Atkins, W.M. and Edman, E.T. (2002) 1.3 angstrom resolution structure of human glutathione transferase with *S*-hexylglutathione bound reveals possible extended ligandin binding site. *Proteins: Struct. Funct. Bioinf.* **48**, 618-627.

Levinthal, C. (1968). Are there pathways for protein folding? *J. Chem. Phys.* **65**, 44-45.

Lian, L.Y. (1998) NMR structural studies of glutathione *S*-transferase. *Cell. Mol. Life Sci.* **54,** 359-362.

Liang, F., Alssadi, R., Morehead, P., Awasthi, Y.C., Godley, B.F. (2005). Enhanced expression of glutathione S-transferase A1-1 protects against oxidative stress in human retinal pigment epithelial cells. *Exp. Eye. Res.* **80**, 113-119.

Liebau, E., Dawood, K.F., Fabrini, R., Fischer-Riepe, L., Perbandt, M., Stella, L., . Pedersen, J.Z., Bocedi, A., Petrarca, P., Federici, G. and Ricci, G. (2009) Tetramerization and cooperativity in *Plasmodium falciparum* glutathione *S*-transferase are mediated by atypic loop 113–119. *J. Biol. Chem.* **284,** 22133–22139.

Lins, L. and Brasseur, R. (1995). The hydrophobic effect in protein folding. *FASEB J.* **9**, 535-540.

Liou, J., Huang, T., Chang, G. (2000). Inhibition of octopus glutathione transferase by Meisenheimer complex analog S-(2,4,6-trinitrophenyl) glutathione. *J. Protein. Chem.* **19**, 615-620.

Listowsky, I., Abramovitz, M.m Homma, M. and Niitsu, Y. (1988). Intracellular Binding and Transport of Hormones and Xenobiotics by Glutathiones-Transferases. *Drug. Metab. Rev*. **19,** 305-318.

Luo, J-K., Hornby, A. T. J., Wallace, A. W., Chien, J., Armstrong, N. R. and Dirr, W.H. (2002). Impact of domain interchange on conformational stability and equilibrium folding of chimeric class μ gluthione transferases. *Protein Sci.* **11,** 2208-2217.

Mannervik, B. (1985) The isoenzymes of glutathione transferase. *Adv. Enzymol. Relat Areas Mol. Biol.* **57**, 357-417.

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

Mannervik, B. and Guthenberg, C. (1981) Glutathione transferase (human placenta). *Methods Enzymol.* **77**, 231-235.

Matulis, D. and Lovrien, R. (1998) 1-Anilino-8-naphthalene sulfonate anion-protein binding depends primarily on ion pair formation. *Biophys. J.* **74**, 422-429.

Maurer, H.R. (1971) Basic principles of polyacrylamide gel electrophoresis and some recent advances of the technique. *Ann. Biol. Clin.* **29**, 205-210.

McPherson, A. (1990). Current approaches to macromolecular crystallization. *Eur. J. Biochem*. 189, 1-23

Meng, E. C., Pettersen, E. F., Couch, G. S., Huang, C. C., and Ferrin, T. E. (2006) Tools for integrated sequence-structure analysis with UCSF Chimera. *BMC Bioinf.* **7**, 339.

Meyer, D.J., Coles, B., Pemble, S.E., Gilmore, K.S., Fraser, G.M. and Ketterer, B. (1991) Theta, a new class of glutathione transferases purified from rat and man. *Biochem. J.* **274**, 409-414.

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**, 15326-15332.

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**, 536-540.

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.

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, 241–246.

Nieslanik, B., Ibarra, C. and Atkins, W.M. (2001) The C-terminus of glutathione Stransferase A1-1 is required for entropically-driven ligand binding. *Biochemistry* **40**, 3536-3543.

Nieslanik, B.S. and Atkins, W.M. (2000) The catalytic Tyr-9 of glutathione *S*-transferase A1-1 controls the dynamics of the C terminus. *J. Biol. Chem.* **275**, 17447-17451.

Nilsson, L.O., Edalat, M., Pettersson, P.L. and 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* **1597,** 157-163.

Orengo, C.A., Michie, A.D., Jones, S., Jones, D.T., Swindells, M.B. and Thornton, J.M. (1997) CATH – a hierarchic classification of protein domain structures. *Structure* **5**, 1093-1108.

Pace, C. N. (1986) Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods Enzymol.* **131**, 266-280.

Pace, C.N., Shirley, B.A. and Thomson, J.A. (1989) in *Protein Structure*: a practical approach (Creighton, T.E. ed) 2nd edn, pp 311-330, IRL Press, Oxford University Press, Oxford.

Papworth, C., Bauer, J.C., Braman, J. and Wright, D.A. (1996) Site-directed mutagenesis using double-stranded plasmid DNA templates. *Strategies* **9**, 3-4.

Parbhoo, N: The role of a conserved interdomain interaction in Escherichia coli glutaredoxin-2. *MSc dissertation*. University of the Witwatersrand, Molecular and Cell Biology; 2010.

Patrick, G.L (2005) *An introduction to Medicinal Chemistry* 3rd ed, Oxford university press inc, New York, USA.

Pemble, S. E., Wardle, A. F. and Taylor, J. B. (1996) Glutathione S-transferase class Kappa: characterization by the cloning of rat mitochondrial GST and identification of human homologue. *Biochem. J.* **319**, 749-754.

Pemble, S.E. and Taylor, J.B. (1992) An evolutionary perspective on glutathione transferases inferred from class-theta glutathione transferase cDNA sequences. *Biochem. J.* **287**, 957-963.

Perito, B., Allocati, A., Casalone, E., Masulli, M., Dragani, B., Polsinelli, M., Aceto, A. and Ilio, C.D. (1996) Molecular cloning and overexpression of a glutathione transferase gene from *Proteus mirabilis*. *Biochem. J.* **318**, 157-162.

Perkins, S.J. (1986) Protein volumes and hydration effects. *Eur. J. Biochem.* **157,** 169-180.

Privalov, P. L. (1979) Stability of proteins: small globular proteins. *Adv Protein Chem* **33**, 167-241.

Privalov, P.L. (1996) Intermediate states in protein folding. *J. Mol. Biol.* **258**, 707-725.

Privalov, P.L. and Gill, S.J. (1988) Stability of protein structure and hydrophobic interaction. *Adv. Protein Chem.* **39**, 191-234.

Ptitsyn, O. B., Bychkova, V. E. and Uversky, V. N. (1995). Kinetic and equilibrium folding intermediates. *Philos Trans R Soc Lond B Biol Sci* **348**, 35-41.

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

Ranson, H., Rossiter, L., Ortelli, F., Jensen, B., Wang, X., RothC. W., Collins, F. H. and Hemingway, J. (2001) Identification of a novel class of insect glutathione S-transferases involved in resistance to DDT in the malaria vector *Anopheles gambiae*. *Biochem*. J. (2001) **359**, 295-304.

Reinemer, P., Dirr, H.W., Ladenstein, R., Huber, R., Lo, B.M., Federici, G. and Parker, M.W. (1992) Three-dimensional structure of class Pi glutathione S-transferase from human placenta in complex with S-hexylglutathione at 2.8 A resolution. *J. Mol. Biol.* **227**, 214-226.

Richardson, J.S (1981) The anatomy and taxonomy of protein structure. *Adv. Prot.Chem*. **34**, 246-253.

Riddles, P.W., Blakely, R.L. and Zerner, B. (1979) Ellman's reagent: 5, 5′-dithiobis (2-nitrobenzoic acid) a reexamination. *Anal. Biochem*. **94**, 75-81

Rupp, B. and Wang, J. (2004). Predicitve models for protein crystallization. *Methods*, 34, 390-407

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

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**, 169-172.

Schmid, F. X., in *Protein Folding*, T. Creighton, Editor. (1992), W. H. Freeman & Co.: New York. p. 199-203.

Semisotnov, G.V., Rodionova, N.A., Razgulyaev, O.I., Uversky, Y.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* **3**, 119-128.

Shank, E.A., Cecconi, C., Dill, J.W., Marqusee. S and Bustamante, C. (2010) The folding cooperativity of a protein is controlled by its chain topology. *Nature*, **465**, 637-641.

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

Sining, I., Kleywegt, G. J., Cowan, S. W., Reinemer, P., Dirr, H. W., Huber, R., Gilliland, G. I., Armstrong, R. N., Ji, X., Board, P. G., Olin, B., Mannevrik, B., and Jones, T. A. (1993). Structure Determination and Refinement of Human Alpha Class Glutathione Transferase A1-1, and a comparison with the Mu and Pi Class Enzymes. *J.Mol.Biol.* **232**, 192-212.

Slavik, J., Horák, J., Říhová, L. and Kotyl, A. (1982) Anilinonaphthalene sulfonate fluorescence and amino acid transport in yeast. *J. Membrane. Biol*. **64**, 175-179.

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**, 484-488.

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**, 80-84.

Stenberg, G., Board, P.G., Carlberg, I. and Mannervik, B. (1991) Effects of directed mutagenesis on conserved arginine residues in a human class Alpha glutathione transferase. *J. Biochem*. **274**, 549-555.

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**, 10421-104218.

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**, 193-197.

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**, 15534-15541.

Stickle, F.D., Presta, G.L., Dill, A.K., and Rose, D.G. (1992) Hydrogen bonding in globular proteins. *J. Mol. Biol.* **226**, 1143-1159.

Stites, E.W. (1997) Protein-protein interactions: interface structure, binding thermodynamics, and mutational analysis. *Chem. Rev.* **97**, 1223-1250.

Stoychev, S.H., Nathaniel, C., Fanucchi. S., Brock, M., Li, S., Asmus, K., Woods, V.L. and Dirr, H.W. (2009) Structural Dynamics of Soluble Chloride Intracellular Channel Protein CLIC1 Examined by Amide Hydrogen-Deuterium Exchange Mass Spectrometry. *Biochemistry*. **48**, 8413-8421.

Stoychev, S: The role of the domain interface in the stability, folding and function of CLIC1**.** *PhD thesis***.** University of the Witwatersrand, Molecular and Cell Biology; 2008.

Stryer, L. (1965) The interaction of a naphthalene dye with apomyoglobin and apohemoglobin. A fluorescent probe of non-polar binding sites. *J. Mol. Biol.* **13**, 482–495

Stryer, L. (1986). Fluorescence spectroscopy of proteins. Science 132, 526-533.

Studier, F.W., Moffat, B.A. (1986) Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J.Mol.Biol.* **189**, 113-130.

Thompson J.D., Higgins, D.G. and Gibson, T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673-4680.

Tomarev, S.I., Zinovieva, R.D., Guo, K. and Piatigorsky, J. (1993) Squid glutathione Stransferase. Relationships with other glutathione S-transferases and S-crystallins of cephalopods. *J. Biol. Chem.* **268**, 4534-4542.

Tripathi , T., Rahlfs, S., Katja, B. and Bhakuni, V. (2007) Glutathione mediated regulation of oligomeric structure and functional activity of *Plasmodium falciparum* glutathione *S*-transferase. *BMC Struc. Biol.* **7**, 67.

Wallace, L. A. and Dirr, H. W. (1999) Folding and assembly of dimeric human glutathione transferase A1-1. *Biochemistry* **38**, 16686-16694.

Wallace, L. A., Blatch, G. L. and Dirr, H. W. (1998). A topologically conserved aliphatic residue in alpha-helix 6 stabilises the hydrophobic core in domain II of glutathione transferases and is a structural determinant for the unfolding pathway. *Biochem J* **336**, 413-418.

Wallace, L. A., Burke, J., and Dirr, W. H. (2000) Domain-domain interface packing at conserved Trp-20 in class α glutathione transferase impacts on protein stability. *Biochim. Biophys. Acta* **1478**, 325-332.

Wallace, L. A., Sluis-Cremer, N. and Dirr, H. W. (1998) Equilibrium and kinetic unfolding properties of dimeric human glutathione transferase A1-1. *Biochemistry* **37**, 5320-5328.

Wang, R.W., Newton, D.J., Huskey, S.E., McKeever, B.M., Pickett, C.B. and Lu, A.Y. (1992) Site-directed mutagenesis of glutathione S-transferase YaYa. Important roles of tyrosine 9 and aspartic acid 101 in catalysis. *J. Biol. Chem.* **267**, 19866-19871.

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**, 1415-1423.

Wetlaufer., D.B. (1973) Nucleation, RaPid folding, and Globular Intrachain regions in Proteins. *Proc. Nat. Acad. Sci. U.S.A* **70**, 697-701.

Widersten, M., Bjornestedt, R. and Mannervik, B. (1996) Involvement of the carboxyl groups of glutathione in the catalytic mechanism of human glutathione transferese A1-1. *Biochemistry* **35**, 7731-7742.

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

Wodak, S. J. and Janin, J. (1981). Location of structural domains in protein. *Biochemistry* **20**, 6544-52.

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

Appendix

Table A: **Showing the buffers used from the Hampton index (Index HR2-144) in attempts to crystallize the R13A hGSTA1-1 enzyme**.

Index number	**Buffer**
5	0.1 M HEPES pH 7.5, 2.0 M Ammonium sulfate
11	0.1 M HEPES pH 7.5, 3.0 M Sodium chloride
15	0.1 M HEPES pH 7.5, 0.5 M Magnesium formate dihydrate
20	0.1 M HEPES pH 7.5, 1.4 M Sodium citrate tribasic dihydrate
21	1.8 M Ammonium citrate tribasic pH 7.0
22	0.8 M Succinic acid pH 7.0
23	2.1 M DL-Malic acid pH 7.0
24	2.8 M Sodium acetate trihydrate pH 7.0
25	3.5 M Sodium formate pH 7.0
26	1.1 M Ammonium tartrate dibasic pH 7.0
27	2.4 M Sodium malonate pH 7.0
29	60% v/v Tacsimate pH 7.0
33	1.1 M Sodium malonate pH 7.0, 0.1 M HEPES pH 7.0, 0.5% v/v Jeffamine ® ED-2001 pH 7.0
34	1.0 M Succinic acid pH 7.0, 0.1 M HEPES pH 7.0, 1% w/v Polyethylene glycol monomethyl ether 2,000
35	1.0 M Ammonium sulfate, 0.1 M HEPES pH 7.0, 0.5% w/v Polyethylene glycol 8,000
36	15% v/v Tacsimate pH 7.0, 0.1 M HEPES pH 7.0, 2% w/v Polyethylene glycol 3,350
38	0.1 M HEPES pH 7.0, 30% v/v Jeffamine ® M-600 ® pH 7.0

39	0.1 M HEPES pH 7.0, 30% v/v Jeffamine ® ED-2001 pH 7.0
44	0.1 M HEPES pH 7.5, 25% w/v Polyethylene glycol 3,350
52	0.2 M Ammonium acetate, 0.1 M HEPES pH 7.5, 45% v/v (+/-)-2-Methyl-2,4-pentanediol
55	0.05 M Magnesium chloride hexahydrate, 0.1 M HEPES pH 7.5, 30% v/v Polyethylene glycol monomethyl ether 550
56	0.2 M Potassium chloride, 0.05 M HEPES pH 7.5, 35% v/v Pentaerythritol propoxylate (5/4 PO/OH)
59	0.02 M Magnesium chloride hexahydrate, 0.1 M HEPES pH 7.5, 22% w/v Poly(acrylic acid sodium salt) 5,100
61	0.2 M L-Proline, 0.1 M HEPES pH 7.5, 10% w/v Polyethylene glycol 3,350
63	5% v/v TacsimateTM pH 7.0, 0.1 M HEPES pH 7.0, 10% w/v Polyethylene glycol monomethyl ether 5,000
68	0.2 M Ammonium sulfate, 0.1 M HEPES pH 7.5, 25% w/v Polyethylene glycol 3,350
72	0.2 M Sodium chloride, 0.1 M HEPES pH 7.5, 25% w/v Polyethylene glycol 3,350
76	0.2 M Lithium sulfate monohydrate, 0.1 M HEPES pH 7.5, 25% w/v Polyethylene glycol 3,350
80	0.2 M Ammonium acetate, 0.1 M HEPES pH 7.5, 25% w/v Polyethylene glycol 3,350
84	0.2 M Magnesium chloride hexahydrate, 0.1 M HEPES pH 7.5, 25% w/v Polyethylene glycol 3,350
87	0.2 M Sodium malonate pH 7.0, 20% w/v

	Polyethylene glycol 3,350
88	0.2 M Ammonium citrate tribasic pH 7.0, 20% w/v Polyethylene glycol 3,350
89	0.1 M Succinic acid pH 7.0, 15% w/v Polyethylene glycol 3,350
90	0.2 M Sodium formate, 20% w/v Polyethylene glycol 3,350
91	0.15 M DL-Malic acid pH 7.0, 20% w/v Polyethylene glycol 3,350
92	0.1 M Magnesium formate dihydrate, 15% w/v Polyethylene glycol 3,350
93	0.05 M Zinc acetate dihydrate, 20% w/v Polyethylene glycol 3,350
94	0.2 M Sodium citrate tribasic dihydrate, 20% w/v Polyethylene glycol 3,350
95	0.1 M Potassium thiocyanate, 30% w/v Polyethylene glycol monomethyl ether 2,000
96	0.15 M Potassium bromide, 30% w/v Polyethylene glycol monomethyl ether 2,000