

https://dspace.org/

A mixed-charge cluster facilities glutathione transferase dimerisation

Walters, John Clive

2006-11-14T06:55:29Z
http://hdl.handle.net/10539/1656

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

A MIXED-CHARGE CLUSTER FACILITATES GLUTATHIONE TRANSFERASE DIMERISATION

John Clive Walters

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, 2005

Declaration

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

_____

_____ day of _____ 2005

Abstract

Cytosolic glutathione transferases (GSTs) are obligate stable homo- and heterodimers comprising two GST subunits. Interactions across the subunit interface play an important role in stabilising the subunit tertiary structure and maintain the dimeric structure required for activity. The crystal structure of a rat Mu class GST consisting of two type one subunits (rGST M1-1) reveals a lock-and-key motif and a mixed-charge cluster at the subunit interface. Previous investigations revealed the lock-and-key motif was not essential for dimerisation. It was therefore postulated that the mixed-charge cluster at the dimer interface is primarily responsible for subunit association. Statistical analyses of individual rGST M1-1 chains did not predict the presence of any charge clusters. This suggests that the mixed-charge cluster forms only upon dimerisation and reinforces the probability that quaternary structure stabilisation is a major role of the mixed-charge cluster. Arginine 81 (Arg-81), a structurally conserved residue in the GST family involved in the mixed-charge cluster, was mutated to alanine. Phenylalanine 56 (Phe-56), the 'key' residue in the lock-and-key motif, was mutated to serine. These changes were engineered to disrupt the mixed-charge cluster and the lock-and-key motif situated at the dimer interface of rGST M1-1. Sizing by gel filtration chromatography of the mutant GST identified that these engineered amino acids resulted in a stable monomeric protein (F56S/R81A rGST M1). The F56S/R81A rGST M1 displayed almost no catalytic activity, suggesting perturbations of the active site or substrate binding sites. Structural investigations of the monomer by far- and near-UV circular dichroism revealed a similar secondary structural content to the wild-type. However, the tryptophan fluorescence properties suggested the tryptophans were situated in more hydrophilic environments than in the wild-type. ANS binding studies indicated a large increase in the accessible hydrophobic surface area of the monomer. Urea-induced equilibrium unfolding of F56S/R81A rGST M1 follows a cooperative two-state unfolding model. The unfolding data indicates decreased conformational stability and a large increase in the solvent exposed surface area of the monomer. In conclusion, the mixed-charge cluster at the dimer interface of rGST M1-1 is essential

for monomeric association, which subsequently contributes to catalytic activity of the dimer and the stabilities of individual rGST M1-1 subunits.

For my late grandfather

George Marx,

who supported me through sickness and health

and for my grandmother

Madeline Marx,

who supported both of us through sickness and health.

Acknowledgements

Professor Heini Dirr for all his support and guidance.

Members of the Protein Structure-Function Research Programme, University of the Witwatersrand, for all their support and friendship.

The National Research Foundation, South Africa and the University of the Witwatersrand for financial support.

Table of Contents

Declaration ii
Abstract iii
Acknowledgements vi
Table of Contents vii
List of Figures ix
List of Tables xi
List of Abbreviations xii
CHAPTER 1 1
INTRODUCTION 1
1.1. Forces and evolutionary pathways guiding oligomerisation 1
1.2. Definition and classification of oligomeric assemblies 2
1.3. Molecular complementarity of protein-protein interfaces 3
1.3.1. Geometric complementarity 4
1.3.2. Physicochemical complementarity 7
1.3.2.1. Amino acid composition 7
1.3.2.2. The hydrophobic effect 8
1.3.2.3. Electrostatic complementarity 9
1.4. Charge Clusters 10
1.4.1. Locality and proposed functionality 10
1.4.2. Amino acid composition of charge clusters 12
1.5. Glutathione Transferases 13
1.5.1. Classification and catalytic function 13
1.5.2. Divergent evolution 14
1.5.3. Three-dimensional structure of the cytosolic GSTs 15
1.5.3.1. Subunit conformation 15
1.5.3.2. Dimerisation and intersubunit interactions at the subunit interface of cytosolic GSTs 17
1.5.3.3. The mixed-charge cluster at the dimer interface of rGST M1-1 19
1.5.3.4. Conformational stability and folding of the cytosolic GSTs 22
1.6. Objectives 23
CHAPTER 2 25
MATERIALS AND METHODS 25
2.1. Materials 25
2.2. Generation of the F56S/R81A mutant plasmid 25
2.2.1. Oligonucleotide primer construction 25
2.2.2. Site-directed mutagenesis 26
2.3. Protein over-expression and purification 28
2.3.1. Protein concentration determination 29
2.4. Characterisation of the F56S/R81A protein 29
2.4.1. Sodium dodecyl sulphate polyacrylamide gel electrophoresis 29
2.4.2. Non-denaturing size exclusion high-performance liquid chromatography 30
2.4.3. Specific enzyme activity 31

2.5. Spectroscopic studies of the F56S/R81A 31
2.5.1. Circular dichroism 31
2.5.4. Fluorescence spectroscopy 32
2.5.4.1. Selective tryptophan fluorescence 32
2.5.4.2. Extrinsic fluorescence spectroscopy - ANS binding 32
2.6. Urea-induced equilibrium unfolding of wild-type rGST M1-1 and F56S/R81A rGST M1 33
2.6.1. Reversibility of unfolding 34
2.6.2. Data fitting of the equilibrium unfolding transitions 34
2.6.2.1. Monomeric two-state model (I ↔ U) 34
2.6.2.2. Three-state model (N_2 ↔ 2I ↔ 2U) 36
2.6.3. Data analysis and molecular graphics 37
CHAPTER 3 38
RESULTS AND DISCUSSION 38
3.1. Generation of wild-type rGST M1-1 and F56S/R81A rGST M1 expression plasmids 38
3.2. Purification of wild-type rGST M1-1 and F56S/R81A rGST M1 43
3.3. Physiochemical properties of wild-type rGST M1-1 and F56S/R81A rGST M1 49
3.3.1. Oligomeric status 49
3.3.2. Specific activity 53
3.4. Spectral properties of the wild-type rGST M1-1 and F56S/R81A rGST M1 58
3.4.1. Far-UV circular dichroism 58
3.4.2. Near-UV circular dichroism 61
3.4.3. Fluorescence spectroscopy 63
3.4.3.1. Non-substrate ligand binding properties 66
3.5. Thermal inactivation of wild-type rGST M1-1 and F56S/R81A rGST M1-1 69
3.6. Urea-induced equilibrium unfolding of the wild-type rGST M1-1 and F56S/R81A rGST M1 71
3.6.1. Techniques monitoring unfolding transitions 71
3.6.2. Reversibility of unfolding 71
3.6.3. Urea induced equilibrium unfolding transitions 73
3.6.4. Analysis of the equilibrium unfolding transitions according to a two-state unfolding pathway (I ↔ U) 82
Conclusion 86
References 87

List of Figures

Figure 1. Complementarity at the dimer interface of rGST M1-1 viewed perpendicular to the two-fold axis. 6

Figure 2. Proposed evolutionary pathways of GSTs.. 15

Figure 3. Ribbon representation of homodimeric rGST M1-1 viewed down the two-fold axis. 16

Figure 4. Ribbon representation of the mixed-charge cluster seen at the dimer interface of rGST M1-1 (6GST). 20

Figure 5. Structure-based sequence alignment of structurally related GSTs. 21

Figure 6. Restriction digest map of pET-20b(+) rGST M1 expression plasmid containing the wild-type rGST M1 encoding cDNA insert. 40

Figure 7. Separation of restriction enzyme digest fragments of the wild-type rGST M1-1 plasmid, F56S/R81A rGST M1 mutant plasmid and F56S rGST M1 plasmid on 1 % agarose gel. 41

Figure 8. Section of the nucleotide sequence coding region of rGST M1 wild-type with corresponding amino acid sequence and the nucleotide sequence coding region of F56S/R81A rGST M1. 42

Figure 9. SDS-PAGE separation of the F56S/R81A rGST M1 over-expression induction study..... 44

Figure 10. Solubility determination of the F56S/R81A rGST M1. 46

Figure 11. Elution profiles of the chromatography purification steps of F56S/R81A rGST M1..... 47

Figure 12. SDS-PAGE separation showing the progressive F56S/R81A rGST M1 purification. 48

Figure 13. SEC-HPLC elution profiles of wild-type rGST M1-1 and F56S/R81A rGST M1..... 50

Figure 14. SEC-HPLC elution profiles of wild-type rGST M1-1 spiked with F56S/R81A rGST M1. 51

Figure 15. A schematic representation of the GST catalysed conjugation of GSH to CDNB..... 54

Figure 16. Specific activity of (A) F56S/R81A rGST M1 and (B) wild-type rGST M1-1 proteins. 55

Figure 17. Far-UV circular dichroism spectra of wild-type rGST M1-1 and F56S/R81A rGST M1. 60

Figure 18. Near-UV circular dichroism spectra of wild-type rGST M1-1 and F56S/R81A rGST M1. 62

Figure 19. Fluorescence emission spectra of rGST M1-1 and F56S/R81A rGST M1. 65

Figure 20. Fluorescence emission spectra of ANS bound to wild-type rGST M1-1 and the F56S/R81A rGST M1. 68

Figure 21. Thermal denaturation curves of F56S/R81A rGST M1 and wild-type rGST M1-1 from 20 - 70°C monitored by circular dichroism spectroscopy at 222 nm. 70

Figure 22. Fluorescence emission spectra of native and refolded wild-type rGST M1-1 and F56S/R81A rGST M1..... 72

Figure 23. Urea-induced equilibrium unfolding of 4 μM F56S/R81A rGST M1 and 2 μM wild-type rGST M1-1. 75

Figure 24. Urea-induced equilibrium unfolding of 4 μM F56S/R81A rGST M1 and 2 μM wild-type rGST M1-1. 76

Figure 25. Urea-induced unfolding of 4 μM F56S/R81A rGST M1 and 2 μM wild-type rGST M1-1 monitored by tryptophan fluorescence. 78

Figure 26. Rayleigh scattering of 4 μM F56S/R81A rGST M1, at different concentrations of urea. 81

List of Tables

Table 1. Thermal cycling conditions utilised to generate the F56S/R81A rGST M1 expression plasmid. 27

Table 2: Thermodynamic parameters obtained from urea equilibrium denaturation of wild-type rGST M1-1 and F56S/R81A rGST M1 monitored by fluorescence and circular dichroism. 79

List of Abbreviations

A	Absorbance
A_{280}	Absorbance at 280 nm
Å	Angstrom
A1-1	Alpha class homodimeric glutathione S-transferase
amp	Ampicillin
ANS	8-Anilino-1-naphthalenesulphonate
ΔASA	Change in solvent accessible surface area
bp	Base pairs
BSA	Bovine serum albumin
camp	Chloramphenicol
CD	Circular dichroism
cDNA	Complementary DNA
CDNB	1-Chloro-2,4-dinitrobenzene
C_m	Chemically-induced midpoint of unfolding
ΔC_p	Change in heat capacity between native/unfolded states and transition state
DHODA	Dihydroorotate dehydrogenase A
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide triphosphate
DTT	Dithiothreitol
E. coli	*Escherichia coli*
EDTA	Ethylene diamine tetraacetic acid
EtBr	Ethidium bromide
F	Fluorescence intensity
F56S	Replacement of wildtype phenylalanine (F) with serine (S) at position 56 of rGST M1-1
ΔG	Change in Gibbs free energy
ΔG_f	Change in Gibbs free energy of unfolding
$\Delta G(H_2O)$	Change in Gibbs free energy in the absence of denaturant

GRASS	Graphical Representation and Analysis of Structures Server
GSH	Reduced glutathione
G-site	Glutathione binding site
GST	Glutathione S-transferase
hGST	Glutathione S-transferase from human
H-site	Hydrophobic electrophile binding site
I	Monomeric intermediate
I_2	Dimeric intermediate
IPTG	Isopropyl β-D-thiogalactoside
kb	Kilobase
kDa	Kilodaltons
K_{eq}	Equilibrium constant
λ_{em}	Emission wavelength
λ_{ex}	Excitation wavelength
λ_{max}	Emission wavelength maximum
LB	Luria Bertani
L-site	Non-substrate ligand binding site
M	Molar
M1-1	Mu class homodimeric glutathione S-transferase with two type 1 subunits
MAPEG	Membrane-associated proteins in eicosanoid and glutathione metabolism
θ	Mean residue ellipticity
MOPS	4-Morpholinepropanesulphonic acid
ε	Molecular extinction coefficient
M_r	Relative Molecular Mass
m-value	Susceptibility of protein to denaturant, i.e., cooperativity
N_2	Native dimer
OD	Optical density
P1-1	Homodimeric Pi class glutahione S-transferase with two type 1 subunits

PCR	Polymerase chain reaction
PDB	Protein Data Bank
pI	Isoelectric point
pK_a	Acid ionisation constant
R	Universal gas constant
R81A	Replacement of wildtype arginine (R) with alanine (A) at position 81 of rGST M1-1
rGST	Glutathione S-transferase from rat
rGST M1	Mu class rat glutathione S-transferase monomer
rGST M1-1	Homodimeric Mu class glutathione S-transferase, subunit type 1 from rat
rpm	Revolutions per minute
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
SEC-HPLC	Size-exclusion high performance liquid chromatography
*Sj*26GST	26 kDa GST from *Schistosoma japonicum*
T	Temperature in Kelvin
TEA	Tris-EDTA Acetate
TEMED	N,N,N'N'-tetramethylethylenediamine
T_m	Temperature-induced midpoint of unfolding
Tris	Tris(hydroxymethyl)-aminomethane
U	Unfolded monomer
UV	Ultra-violet

The IUPAC-IUBMB three letter and one letter codes are used for amino acids.

CHAPTER 1

INTRODUCTION

1.1. Forces and evolutionary pathways guiding oligomerisation

Protein structure consists of four distinct levels: the primary, secondary, tertiary and quaternary level. The quaternary level refers to the complex arrangement and interactions of protein subunits with each other. These subunits utilise varying subunit interfaces to associate. Protein-protein interactions mediate multimeric protein assembly (Jones and Thornton, 1995). Aspects such as reduced hydrophobic surface area, increased stability and novel function creation (due to intersubunit contact) are the driving forces behind the existence of multimeric proteins (Goodsell and Olson, 1993). Association of the correct molecular partners to form multimers is a critical aspect of highly specific biochemical processes such as signal transduction, detoxification and the expression of genetic information (McCammon, 1998). For example, collagen, a fibrous multimeric protein assembly, is required for both structural and functional roles within the cell. Oligomerisation enables the cell to form structural frameworks used for scaffolding (e.g. tubulin) and mechanical transduction (e.g. in muscle contraction). The defining components of protein-protein complexes are frequently the crucial determinants to protein function.

Extensive studies identified various evolutionary pathways that guide protein-protein association. One study divided 32 homodimeric proteins into three groups, depending on which evolutionary pathway they followed (Xu *et al.*, 1998). The first pathway involved direct dimer formation by a two-state model where no stable intermediates were found. The surfaces of the individual monomers share the fundamental characteristic of hydrophobicity with protein cores (Tsai and Nussinov, 1997). This characteristic is thought to drive protein-protein association. In the second

mechanism, the individual subunits first fold to stable monomeric structures which then associate to form the functional oligomer. The individual subunits in this three-state model are compact relative to their interfaces, which are extensively hydrated (Tsai and Nussinov, 1997). The rearrangement of domains in a multidomain protein, referred to as domain swapping, is a two-state model and describes the third pathway. In this proposed mechanism the interdomain interactions are replaced by inter-monomer interactions (Bennett *et al.*, 1994).

1.2. Definition and classification of oligomeric assemblies

Multimeric complexes occur between identical or non-identical subunits, distinguished as either homo- or heterocomplexes, respectively. Subunit organisation of either type of complex can be isologous or heterologous, with structural symmetry (Monod *et al.*, 1965; Nooren and Thornton, 2003). If the surfaces on two interacting subunits are identical the interface is isologous. These interfaces have a two-fold axis of symmetry and only occur in dimers. Heterologous interfaces are described by interactions between different surfaces on two subunits, and without closed symmetry these can lead to aggregation (Nooren and Thornton, 2003).

Multimeric complexes can further be distinguished according to whether a complex is obligate or non-obligate (Nooren and Thornton, 2003). The stability and functionality of an obligate complex depends on the formation of the multimeric state. A non-obligate complex is one composed of units that are each independently stable and show independent functionality.

The association of subunits within protein complexes can be transient or permanent. Transient complexes are those that form temporarily. These complexes are common in a broad range of biological processes including signal transduction, hormone-receptor binding, nuclease inhibition and enzyme allostery, and are a type of non-obligate interaction (Jones and Thornton, 1996; Nooren and Thornton, 2003).

Obligate complexes are by definition permanent (Jones and Thornton, 1996; Nooren and Thornton, 2003). Proteins whose stability or functions depend on subunit association display permanent assemblies. These proteins are found in all aspects of life, from those in extensive multimeric constructs (e.g. muscle fibres), to smaller precise ones (e.g. the GroEL–GroES chaperonin complex). Non-obligate permanent complexes do exist, but association is usually irreversible (e.g. an enzyme-inhibitor complex).

1.3. Molecular complementarity of protein-protein interfaces

Topologically, protein-protein interfaces vary from being relatively flat to an assortment of convoluted protein surfaces. Additionally, extremely complex interfaces contain extended arms, deep cavities or possibly even loops which completely surround the adjoining subunit chain. The *arc* and *trp* repressors and interleukins 5 and 10 all possess extensively interdigitated interfaces (Larsen *et al.*, 1998). These geometric interface arrangements are essential for specific complementarity and correct molecular association. Thus, specific complementary is required for accurate molecular recognition (Chothia and Janin, 1975; Duquerroy *et al.*, 1991) and stabilisation of newly formed constructs.

Protein-protein interfaces exhibit intermediate hydrophobicity between the compact protein core and the protein surface (Jones and Thornton, 1996). Protein-protein interfaces are composed of hydrophobic patches inter-dispersed with intersubunit hydrogen bonds and water molecules along the entire interface (Larsen *et al.*, 1998). Interactions within these hydrophobic patches stabilise protein-protein interfaces (Argos, 1988; Chothia and Janin, 1975; Janin *et al.*, 1988; Janin and Chothia, 1990; Jones and Thornton, 1995; Jones and Thornton, 1996). Although exceptions exist that demonstrate no hydrophobic interactions occur at the interface (Tsai *et al.*, 1997), the hydrophobic effect is the prevailing driving force for oligomerisation.

In a theoretical study, the removal of apolar side-chains from the aqueous environment was shown to be the thermodynamic driving force behind protein-protein associations (McCammon, 1998). Consequently during association, complementary molecular surfaces have the potential to significantly reduce their solvent exposed surface area, and accordingly gain stability for the formation of dimeric complexes from intra-monomeric hydrophobic effects. This facilitates correct subunit positioning for the formation of complementary hydrogen bonds and other electrostatic and van der Waals interactions, without unfavourable steric hindrance (Meyer *et al.*, 1996).

1.3.1. Geometric complementarity

Molecular complementarity refers to the geometric interlocking, charge reciprocation and side chain packing at protein-protein interfaces (Jones and Thornton, 1996). Suggestions have been made that although the hydrophobic effect drives protein-protein association, it is hydrogen bonds and salt bridges that confer specificity to the association (Fersht, 1984). Architectural interface complementarity is an additional requirement for specificity and thus association. Although side-chain packing at interfaces is generally not optimal, overall good geometric complementarity is retained (Hubbard and Argos, 1994). This geometric complementarity can be characterised according to the size of buried surface area upon association, paucity of buried water molecules, packing density of atoms at the interface (Chothia and Janin, 1975) and contour correlation between subunits (Lawrence and Colman, 1993). The size of protein-protein interfaces can be defined in terms of the change in solvent accessible surface area (ΔASA) upon association. The area of accessible surface that becomes inaccessible to the solvent due to protein-protein association is referred to as ΔASA.

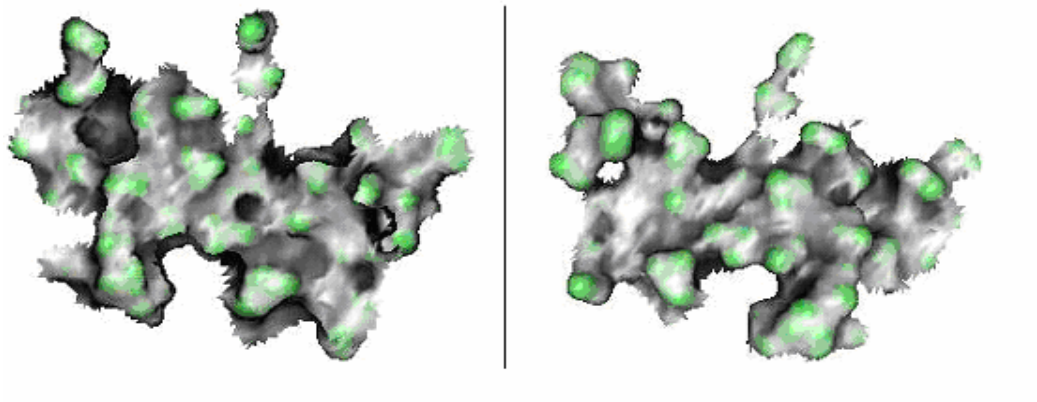
Interface topology can vary from a flat to highly convoluted structures. In Figure 1A, the dimer interface topology of rat Mu class GST comprising of two type one subunits (rGST M1-1) is shown with respect to each separate chain. In this case, the

protein-protein interface is relatively flat with slight convolutions and no interdigitation. In contrast, the *met*, *mnt*, *arc* and *trp* repressors represent proteins with extremely interdigitated interfaces (Larsen *et al.*, 1998). A third intermediate group of interfaces demonstrate interdigitation to a lesser extent. In this group, the majority of the interface forms between two globular subunits when a short terminal chain or flexible loop on one subunit forms a tight grip with the adjoining subunit, thereby locking them together (Larsen *et al.*, 1998).

Molecular recognition depends on shape complementarity. It involves the physical fit between two surfaces which can be relatively flat or interdigitated (Jones and Thornton, 1995; Jones and Thornton, 1996; Larsen *et al.*, 1998). Argos (1988) considered that an overall flatness would be expected when comparing the symmetry involved in the associations. Although symmetry does not imply a flat interface, it does require complementing pockets for any protrusions (Jones and Thornton, 1995).

Shape complementarity within an interface has been quantified through many different methods. Jones and Thornton (1996) devised the gap volume index which measures the volume of cavities between interacting surfaces as a function of the interface accessible surface area. Most interfaces analysed by the gap volume index were found to be relatively flat compared with the rest of the protein surface. This was particularly the case for homodimers and permanent heterocomplexes (Jones and Thornton, 1996). Lawrence and Colman (1993) devised the shape correlation index based on distance and the angle of the normal vectors to the molecular surface. When both methods were incorporated to quantify shape complementarity, oligomers and inhibitor complexes showed better shape complementarity than antibody-antigen complexes (Lo *et al.*, 1999). This combined method also revealed packing density at the centre of interfaces resembles that of the protein interior (Lo *et al.*, 1999).

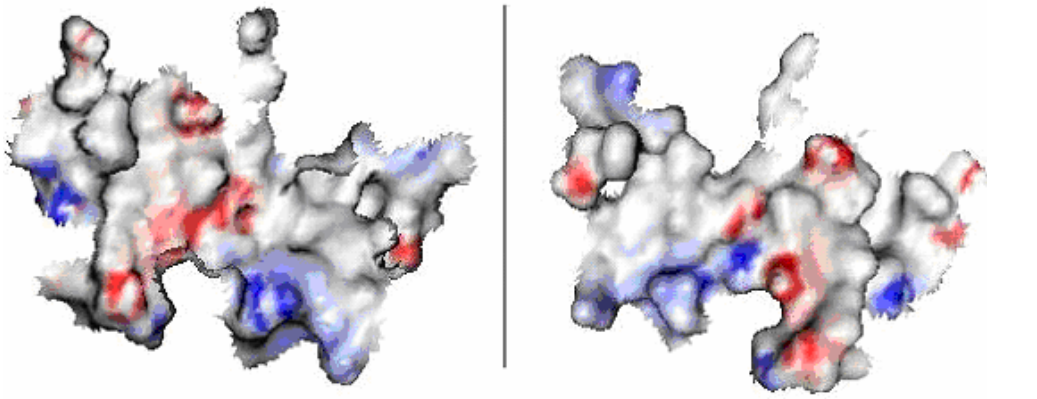
A.



Dimer axis

Chain B Chain A

B.



Dimer axis

Chain B Chain A

Figure 1. Complementarity at the dimer interface of rGST M1-1 viewed perpendicular to the two-fold axis.

The homodimeric structure of rGST M1-1 generated using the Graphical Representation and Analysis of Structures Server (GRASS), http://trantor.bioc.columbia.edu/GRASS/surfserv_enter.cgi, (Nayal *et al*., 1999) showing (A) The concave and convex geometric curvature of the interface represented by varying shades of grey and green respectively, and (B) The negative and positive electrostatic charges represented by the varying shades of red and blue, respectively. Increased colour intensity corresponds to increased curvature and charge strength.

1.3.2. Physicochemical complementarity

The ability of biological molecules to interact with each other in a specific manner is controlled by fundamental molecular interactions. Many criteria have been used to identify and characterise these interactions thereby enhancing our understanding of the biological function of proteins. Among them are those that consider physicochemical properties of protein-protein interfaces such as electrostatics (Sheinerman *et al.*, 2000), hydrophobicity (Tsai *et al.*, 1997; Tsai and Nussinov, 1997) and amino acid composition (Janin *et al.*, 1988). Recently, approaches based on scoring the evolutionary conservation of variable residues have identified crucial regions facilitating complementarity (Bogan and Thorn, 1998).

In order to comprehend molecular complementarity, the physicochemical properties of protein-protein interfaces need to be fully understood. Energetic contributions to protein-protein interactions by residues located at protein interfaces vary considerably (Bogan and Thorn, 1998). A few crucial residues at protein-protein interfaces could contribute dominantly to the binding free energy. Bogan and Thorn (1998) compiled a database of alanine-scanning mutations at protein interfaces. Residues mutated to alanine that have a large effect ($\Delta\Delta G \geq 2$ kcal.mol^{-1}) on the binding free energy of the protein complex were referred to as 'hot spots'. There is a good correlation between structurally conserved residues and 'hot spots' (Keskin *et al.*, 2005). Thus, critical residues involved in complementarity can be identified based on their energetic contribution and structural composition.

1.3.2.1. Amino acid composition

Protein-protein interface characteristics differ from general protein surfaces. Oligomers utilise residues that protrude from one subunit into another to secure the subunits together (Jones and Thornton, 1995). Shown in Figure 3 is the Phe-56 lock-

and-key motif of rGST M1-1, which is a good example of a protein utilising protruding residues to lock individual subunits together.

Intersubunit protein-protein interfaces are less polar than protein surfaces and typically contain charged residues. The proportion of non-polar and polar interactions at oligomer interfaces varies considerably (Janin *et al.*, 1988). Studies show that approximately 33 % of the interface area, 38 % of the protein buried surface area and 13 % of the protein accessible surface area consist of the non-polar residues isoleucine, leucine, phenylalanine, valine, cysteine, and methionine. Of the charged residues, aspartic acid, glutamic acid and leucine contribute 33 % to accessible surface area of the whole protein, 14 % to interface area and 15 % to buried surface area of the whole protein. Arginine and leusine are the highest contributing residues to protein interfaces. Despite the apolar nature of protein-protein interfaces, they are the only areas where concentrated numbers of arginine residues are found. Arginine contributes four times the amount of hydrogen bonds across interfaces when compared to lysine (Janin *et al.*, 1988) and it is typically involved in multiple salt bridges across interfaces (Tsai *et al.*, 1997). Thus, the abundance and arrangement of arginine residues in amongst a more apolar environment at the interface surface facilitates the formation of hydrogen bonds and salt bridges.

1.3.2.2. The hydrophobic effect

Stabilisation of protein-protein associations, as well as protein folding, depends crucially on the 'hydrophobic effect' (Dill, 1990; Tsai and Nussinov, 1997). Studies exploring this concept have revealed that the contribution of the 'hydrophobic effect' to protein-protein associations is not as strong as it is for protein folding (Tsai and Nussinov, 1997).

Jones and Thornton (1996) discovered the average hydrophobicity of the protein core is positive, while that of protein surfaces is negative. However, recent studies indicated subunit interfaces in homodimeric proteins were hydrophobic and contain

twice the buried surface area as protein-protein complexes (Bahadur *et al.*, 2003). Hydrophobic groups in protein-protein interfaces are scattered over the entire surface, forming patches interspersed with charged and polar residues. Protein interface composition is more similar to that of protein surfaces than protein cores.

1.3.2.3. Electrostatic complementarity

Complementarity of multimeric interfaces is resultant not only from the shape of the interacting surfaces, but also the relative positioning of the charged side-chain residues at each surface (Jones and Thornton, 1995). Figure 1B represents the protein-protein interface charge distribution of rGST M1-1. This diagram demonstrates the electrostatic complementarity that is essential for dimerisation. Charged interactions at protein-protein interfaces together with geometric shape (Larsen *et al.*, 1998) play an important role in generating unique structures and conferring conformational specificity (Fersht, 1984). Interactions between the charged chemical groups of ionisable residues (histidine, arginine, lysine, N-terminal amide, C-terminal carboxyl groups, aspartic acid and glutamic acid) characterise electrostatic interactions. These interactions can occur between residues that are sequentially distant from each other along the polypeptide backbone thereby contributing to tertiary and quaternary conformation. The contribution of electrostatic interactions to protein structure may be attractive or repulsive. They can stabilise or destabilise protein-protein interactions. Hydrogen bonds and salt bridges have been shown to be essential in determining binding specificity of interfaces (Fersht, 1984). The energy cost of burying an ion pair within the interface is not as high as the equivalent burial in protein cores, which explains the greater contribution of ion pairs and charge clusters to protein-protein interface stabilisation.

From a survey of 355 protein dimer interfaces (Jones and Thornton, 1995), the total calculated number of intersubunit salt bridges was found to be a maximum of five for each individual dimer. Of these protein interfaces 56 % showed no intersubunit salt bridges. A further investigation demonstrated that dimeric dihydroorotate

dehydrogenase A (DHODA) contains two intersubunit salt bridges that form between Glu-206 of subunit one, and Lys-296 of the other subunit. These conserved residues were found not to be essential for enzyme activity, but perturbation of the salt bridges resulted in inactive monomers (Ottosen *et al.*, 2002). Upon exposure to 0.15 M salt DHODA became stabilised rather than destabilised. This implies that hydrophobic interactions at the dimer interface are more important to this protein-protein association than electrostatic forces, which are expected to weaken in the presence of salt (Ottosen *et al.*, 2002). This research demonstrated that these two salt bridges hold the subunits of dimeric DHODA together only partially, and that salt bridges are not essential for dimerisation.

In a mathematical analysis of 20 dimers and tetramers (Janin *et al.*, 1988), acceptor and donor groups within 3.5 Å of each other all displaying acceptable angular geometry were assumed to be hydrogen bonded. Within this dataset 264 potential hydrogen bonds were identified. Certain proteins like mellitin and the subtilisin inhibitor showed no hydrogen bonds at the interface. Some extremely small interfaces that cover less than ~ 1000 $Å^2$ per subunit only have one or two hydrogen bonds. For example, uteroglobin has one very weak hydrogen bond for ~ 1500 $Å^2$ of interface area. Interfacial hydrogen bonds formed between protein subunits involve charged donor or acceptor groups. Thus, subunit interfaces form relatively few intermolecular hydrogen bonds in proportion to their molecular recognition surfaces, thereby emphasising the specificity required for their formation.

1.4. Charge Clusters

1.4.1. Locality and proposed functionality

A charge cluster describes the spatial distribution of charged residues within a protein with extreme concentration of charge relative to the rest of the protein's charge distribution (Zhu and Karlin, 1996). In addition, charge clusters make more contacts

with other residues within the same cluster than with surrounding residues (Heringa and Argos, 1991). Charge clusters are classified as positive, negative or mixed.

Identifying important cluster residues in protein dimerisation has been achieved by graph-spectral analysis (Brinda *et al.*, 2002). Properties such as amino acid composition, solvent accessibility and residue conservation, were used to identify side chain clusters occurring at the protein interface. Generally, these clusters were found near the center of protein-protein interfaces, which excludes them from solvent.

A study examining approximately 40 000 non-redundent linear protein sequences revealed that 20-25 % of the eukaryotic and 6-8 % of the *Escherichia coli* proteins display at least one significant linear charge cluster (Zhu and Karlin, 1996). In the same analysis 186 three-dimensional protein structures were investigated in which 10.2 % contained at least one charge cluster. Of these proteins 4.3 % exhibited charge clusters connecting two or more chains. The greatest proportion were mixed-charge clusters which involve approximately equal amounts of anionic and cationic residues. Furthermore, side-chain solvent accessibility of residues involved in charge clusters is greater than 30 % suggesting that most of the charge clusters examined are found on the protein surfaces (Zhu and Karlin, 1996). Conversely, charge clusters seen at the protein-protein interface are inaccessible to solvent (Zhu and Karlin, 1996). Thus charge clusters linking two chains are scarce but have a crucial function.

Generally, charged clusters function to create and stabilise protein conformation (Zhu and Karlin, 1996) and are important in processes such as protein transport, localisation and regulatory function. It has been shown that charge clusters force preformed secondary structural elements to associate. This occurs by either inward collapse of the structural elements with the inducing cluster anchor sites, or by direct formation of a central interaction core made up partly of cluster residues (Heringa and Argos, 1991). In addition, mixed-charge clusters are involved in the formation of multi-domain complexes, whereas same sign clusters contribute electrostatic repulsion to ensure separation between protein constructs (Zhu and Karlin, 1996).

Large clusters can be found in loops (this is not predominant) which suggest that the cluster functions in the packing of secondary structures and/or in linking separate molecules. Charge clusters function in all aspects of protein function and structure. They therefore pose an exciting angle for further investigation, particularly in protein-protein associations since they are potential mediators of highly specific interactions.

1.4.2. Amino acid composition of charge clusters

Side chain clustering can facilitate long-range interactions making it possible to connect structural elements far from each other in the primary sequence (Heringa and Argos, 1991). A study has shown that dense clusters (all cluster types including charge clusters) are biased towards larger residues such as tryptophan, histidine, arginine, tyrosine, glutamic acid, glutamine and phenylalanine (Heringa and Argos, 1991). These residues are favoured because substantial contact is possible amongst these groups, even through pairwise association alone. This correlates with the residue accessible surface area, molecular weight and residue size. The clusters examined prefer polar to hydrophobic residues. The highly hydrophobic phenylalanine and tryptophan residues are unexpectedly still favoured over the other preferred residues. This implies that clusters are not limited only to the hydrophobic core of proteins but are also seen near to the protein surface. Additionally, the analysis showed that glutamic acid, ionised tyrosine, aspartic acid, histidine, arginine and occasionally lysine are the preferred residues within charged clusters. The pK_a of tyrosine is high therefore its hydroxyl group prefers to be close to nitrogen atoms found in histidine and lysine residues. When residues of equal polarity are compared, it is usually the amino acids that contain larger numbers of methyl groups that are favoured. Thus, arginine is favoured over lysine, glutamic acid over aspartic acid and glutamine over asparagine. This facilitates maximum surface contact for correct fundamental recognition and interaction.

1.5. Glutathione Transferases

1.5.1. Classification and catalytic function

Glutathione transferases (GSTs) (EC 2.5.1.18) are enzymes that function in the phase II detoxification of xenobiotic mutagens, carcinogens and other toxic substances which are constantly present within the enviroment of organisms (Sheehan *et al.*, 2001). In general, the GSTs catalyse the nucleophilic addition of reduced glutathione thiol groups to the electrophilic centres in various organic compounds. Ultimately, a more water-soluble glutathione conjugate is formed thereby aiding detoxification and excretion. The general reaction catalysed by GSTs is described by:

$$R\text{-}X + GSH \rightarrow GS\text{-}R + X\text{-}H \qquad (1)$$

where R is the hydrophilic substrate with a leaving group X.

GSTs are representatives of two highly complex and very different superfamilies (Hayes and Strange, 2000). In a review paper (Sheehan *et al.*, 2001), a probable evolutionary divergence pattern shown in Figure 2 has been established based on sequence alignments. From these comparisons the soluble mammalian GSTs have been separated into eight classes, namely Alpha, Mu, Pi, Sigma, Theta, Zeta, Omega and Kappa, (Board *et al.*, 1997; Board *et al.*, 2000; Mannervik, 2003; Meyer *et al.*, 1991; Meyer and Thomas, 1995; Pemble *et al.*, 1996a). Bacteria, insects and plants contain a further four classes, Beta, Delta, Phi and Tau (Hayes and McLellan, 1999). Collectively these GSTs make up the first of the two superfamilies. The second superfamily is composed of the microsomal GSTs and these have been given the title of membrane-associated proteins in eicosanoid and glutathione metabolism (MAPEG) (Jakobsson *et al.*, 1994).

1.5.2. Divergent evolution

Several models describing the possible evolutionary pathway adopted by GSTs have been proposed (Armstrong, 1997; Ladner *et al.*, 2004; Sheehan *et al.*, 2001). Initial models stated that class Kappa GST was the ancestor of Theta and cytosolic GSTs and that these proteins began via a divergent evolutionary process (Armstrong, 1997; Pemble *et al.*, 1996b). Sheehan *et al*. (2001) proposed a contrasting view in which the Kappa enzyme is not an ancestor of the canonical fold but rather an early divergence from an ancestral gene.

Figure 2 is a schematic of the proposed GST evolution adapted from Sheehan *et al*. (2001). The multiple GST classes are shown to evolve from a common ancestor via divergent evolution. Two major points of divergence are known. The first is evolution of the Kappa and cytosolic GST superfamilies from the electron transfer proteins thioredoxin and glutaredoxin (Ladner *et al.*, 2004). The second is a divergence of class Alpha/Mu/Pi/Sigma from a supposed Theta precursor protein (see Figure 2). Divergence of classes Alpha/Mu/Pi GSTs seem to have occurred by the acquisition of a curved dimer interface and an important lock-and-key hydrophobic interaction, whereas class Sigma GST retained the flat hydrophilic dimer interface characteristic of the older GSTs.

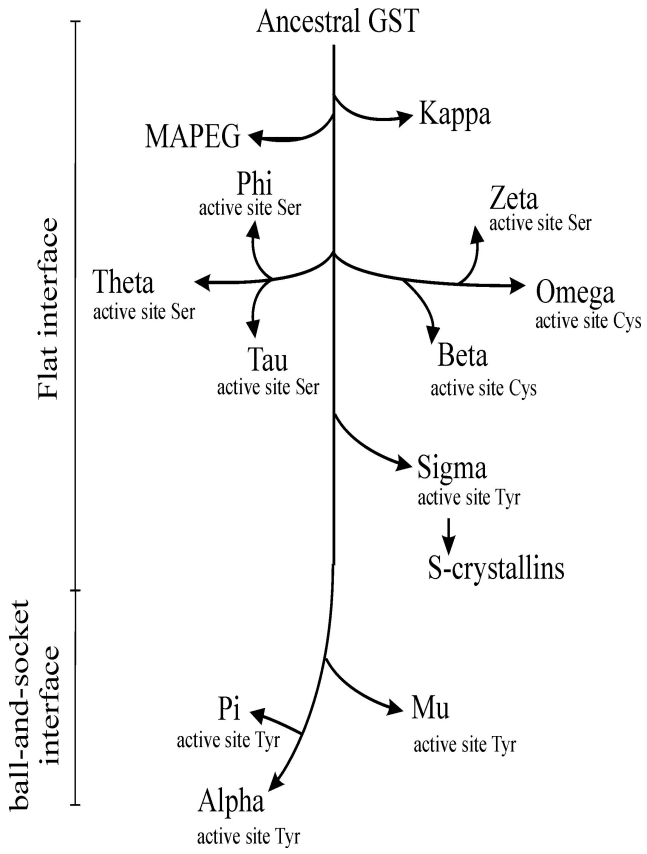


Figure 2. Proposed evolutionary pathways of GSTs taken from Sheehan *et al.* (2001).
The cytosolic and class Kappa GSTs are proposed to have diverged from a common thioredoxin/glutaredoxin ancestor via parallel mechanisms. Divergence from the supposed Theta precursor protein occurred with the introduction of the lock-and-key motif by the class Alpha/Mu/Pi GSTs.

1.5.3. Three-dimensional structure of the cytosolic GSTs

1.5.3.1. Subunit conformation

GSTs are obligate dimers and exist as stable homo- or heterodimers that posses a conserved archetypical fold (Dirr *et al.*, 1994). The crystal structure of M1-1 (Ji *et al.*, 1992) is shown in Figure 3. The canonical GST fold is composed of two domains in each subunit separated by a short linker region. The N-terminal domain (domain 1) resembles a thioredoxin-like fold composed of an antiparallel β-sheet between three

α-helices arranged in a *βαβαββα* motif (Wilce and Parker, 1994). The C-terminal domain (domain 2) consists of five tightly packed amphipathic α-helices forming a conserved hydrophobic core.

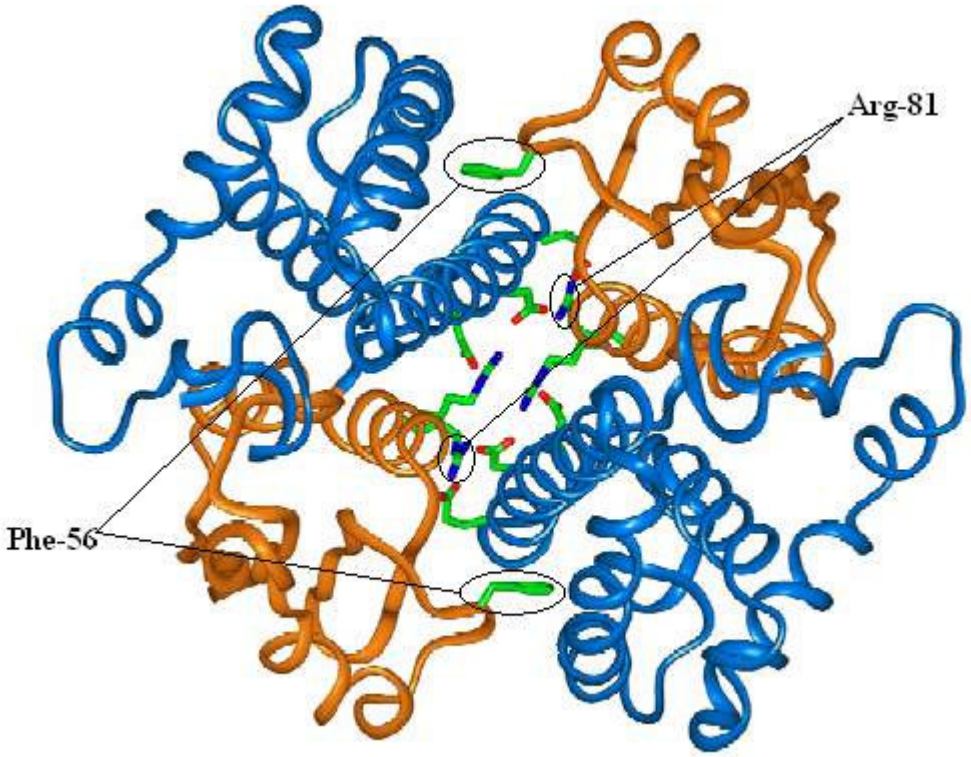


Figure 3. Ribbon representation of homodimeric rGST M1-1 viewed down the two-fold axis.

Illustrated in stick representation is the hydrophobic lock-and-key motif around Phe-56 of rGST M1-1 (6gst), and the microenvironment around Arg-81, adapted from Hornby *et al.* (2002). Blue and brown ribbons represent domains 2 and 1, respectively. The charged cluster around Arg-81 is displayed in stick representation (residues unlabeled).

Domain 2 is less conserved within and between classes than domain 1. The Mu class subunit displays an extensive domain interface, with the surface area that becomes buried upon domain association being comparable to the buried surface area upon dimerisation. Interactions between domains in the GST subunit occur primarily between α-helix 1 and 3 in domain 1, and α-helix 4 and 6 in domain 2.

1.5.3.2. Dimerisation and intersubunit interactions at the subunit interface of cytosolic GSTs

Dimerisation of GSTs is achieved by specific interactions between domain 1 of subunit 1 and domain 2 of subunit 2. Dimer stabilisation in class Alpha (Sayed *et al.*, 2000), class Pi (Stenberg *et al.*, 2000) and class Mu (Hornby *et al.*, 2002) is dependent on these interactions. In addition, correct molecular folding is also influenced by these interactions (Dirr, 2001). These interactions result in a V-shaped cleft being formed between subunits. However, class Beta lacks this V-shaped cleft due to a much denser set of interactions between domain 2 helices, therefore creating a more extensive subunit interface (Rossjohn *et al.*, 1998).

GST intersubunit interactions contribute to the stability of the individual subunit tertiary structure, and dimerisation is essential for the correct functional conformation of the active sites (Dirr, 2001; Sayed *et al.*, 2000; Stenberg *et al.*, 2000) These active sites are situated at each end of the V-shaped cleft on both sides of the dimer interface. Each active site comprises two regions. The first binds reduced glutathione at the G-site and the second binds hydrophobic electrophilic compounds at the H-site. A comparison of Alpha/Mu/Pi to the prehistoric GST classes shows a conserved hydrophobic lock-and-key interaction at the dimer interface found within the loop between α-helix 2 and β-strand 3 in domain 1. In this interaction, a conserved phenylalanine residue (Phe-51 in Alpha, Phe-56 in Mu and Phe-47 in Pi) projects from subunit one into a hydrophobic compartment created by α-helix 4 and α-helix 5 within domain 2 of the neighbouring subunit. Figure 3 shows the lock-and-key motif situated around Phe-56 in rGST M1-1. It was believed that the ancestral class sigma

and theta GSTs lack this interaction but substitute the interaction with increased electrostatic interactions between their subunits. Recently, however, the motif has been shown to be present in sigma GST (Agianian *et al.*, 2003).

GST dimer interfaces are separated into two groups, namely Alpha/Mu/Pi/sj26 containing a curved interface with the hydrophobic lock-and-key motif, and the ancestral members Sigma/Theta with a flatter, more hydrophilic interface lacking the hydrophobic lock-and-key motif (Armstrong, 1997). Examining rGST M1-1 as a representative of the curved interface group, it is seen that the interactions at either end of the interface are hydrophobic while more polar contacts are seen at the dimer's two-fold axis (Ji *et al.*, 1992). This is highlighted by the existence of the conserved hydrophobic lock-and-key motif, and the central polar mixed-charge cluster which is shown in Figure 3.

When dimerisation occurs in class Alpha, Mu and Pi GSTs there is approximately 14 % burial of the initial solvent accessible surface area of the monomeric units (Dirr *et al.*, 1994). The Omega and Theta classes demonstrate an open dimer interface arrangement (Reinemer *et al.*, 1996), hence the buried solvent accessible surface area due to dimerisation is smaller when compared to Alpha, Mu and Pi classes. Inter-chain salt bridges are thought to be essential for dimerisation in GSTs. This is demonstrated by rGST M1-1 which displays significant salt bridges between Arg-81 from chain A, and Glu-90 and Asp-97 from chain B, and similarly, Arg-81 of chain B with Glu-90 and Asp-97 from chain A. These interactions are relevant to dimerisation as they contribute to the mixed-charge cluster (Figure 4) seen at the interface which is thought to be essential for dimer formation (Zhu and Karlin, 1996). Interdomain salt bridges exist between Arg-77 of domain 1 and Asp-97 and E-100 of domain 2 of the same subunit (Ji *et al.*, 1992). The two-guanidino groups of each subunit's Arg-77 are stacked on top of each other at the interface. In addition, these symmetrical guanidino groups are structurally conserved. Structure based sequence alignment (Figure 5) has identified Arg-70 (Pi), Arg-68 (Sigma), Arg-77 (Mu) and Arg-72 (*S. japonicum*) as highly conserved and significant residues in GST intersubunit interactions.

1.5.3.3. The mixed-charge cluster at the dimer interface of rGST M1-1

The rGST M1-1 charge cluster shown in Figure 4 involves two glutamic acids, one aspartic acid and three arginine residues from chain A, and two glutamic acids, one aspartic acid and four arginine residues from chain B. These residues interact with each other via the formation of electrostatic links (Ji *et al.*, 1992). The presence of Arg-81, Arg-77, Phe-154, Glu-90, E-100 and Asp-97 is a good example of the amino acid preference shown in clusters (Heringa and Argos, 1991). In rGST M1-1, the two uncharged residues Phe-154 (chain A) and Phe-154 (chain B) are connected via hydrogen bonds to E-100 (chain A) and E-100 (chain B), respectively. Arg-81 (chain A) forms salt bridges with Glu-90 and Asp-97 (chain B) and the interactions are reciprocated. These hydrogen bonds and salt bridges are speculated to be involved in the quaternary stabilisation of the enzyme (Zhu and Karlin, 1996). Statistical analysis shows no charged clusters when individual rGST M1-1 subunits are examined. This indicates that the charge cluster forms only upon dimerisation and thereby reinforces the possibility that quaternary structure stabilisation is one of the major roles of a charge cluster (Zhu and Karlin, 1996).

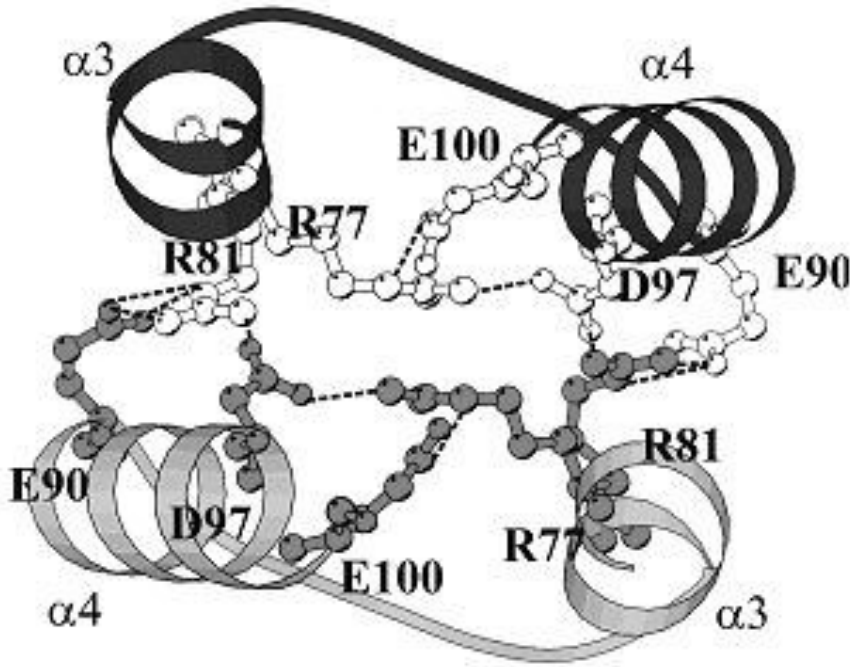


Figure 4. Ribbon representation of the mixed-charge cluster seen at the dimer interface of rGST M1-1 (6GST).
Dark and light ribbons are used to distinguish between the backbone conformations of the two subunits taken from Luo *et al.* (2002).

Zhu and Karlin (1996) speculate that the rGST M1-1 charge cluster is primarily responsible for monomeric association. This is probable as these charge interactions were still intact when the lock-and-key motif was removed (Hornby *et al.*, 2002). Therefore, as seen in class Alpha (Sayed *et al.*, 2000) the lock-and-key motif is not essential for dimerisation.

Arg-81 is part of the mixed-charge cluster seen at the dimer interface of rGST M1-1. From Figure 5, it can be seen that Arg-81 is conserved in the two ancestral GSTs Sigma and Pi. Arg-81 is substituted with Glu-81 in the Omega, Tau and Zeta classes. Glutamic acid and arginine are both polar however they have opposite charges. Ser-76 and Cys-75 are the positional equivalents to Arg-81 from rGST M1-1 in class

Alpha and class Theta GSTs, respectively. Although serine is not a charged residue it has the capacity to form polar interactions, unlike the hydrophobic cysteine which is not able to interact electrostatically. Due to the positional conservation of Arg-81, its replacement is expected to result in the charge cluster's function becoming compromised, as well as alter the cluster conformation.

PDB Code	Residue sequence number	Sequence
6GST	(80)	ARKH-----*H*LCGETEEERIRADIVENQVM-DNRMQLI--MLCYNP----
1K3Y	(76)	ASKY-----*N*LY*G*KDIKERALIDMYIEGIA-DLGEMIL--LLPVCPPE--
1AQW	(73)	GRTL-----*G*LY*G*KDQQEAALVDMVNDGVE-DLRCKYI--SLIYT-----
2GSQ	(71)	AREF-----*G*LD*G*KTSLEKYRVDEITETLQ-DIFNDVV--KIKFAPEA--
1EEM	(94)	DEAYP--*G*KKLLPDDPYEKACQKMILELFS-KVPSLVG--SFIRSQN---
1GWC	(77)	DEVFAST*G*PSLLPADPYERAIARFWVAYVDDKLVAPWR--QWLRGKT---
1LJR	(75)	SCKY*Q*-TPDHWYPSDLQARARVHEYLGWHADCIR*G*TFGIPLWVQVLGPLI
1FW1	(80)	EETRPT--PRLLPQDPKKRASVRMISDLIAGGIQPLQNLSVLKQV*G*----

Figure 5. Structure-based sequence alignment of structurally related GSTs.
A structure-based sequence comparison representing closely related GST structural neighbours is shown. Assigned to each structural neighbour are the PDB codes 1k3y (Alpha), 1aqw (Pi), 2gsq (Sigma), 1eem (Omega), 1gwc (Tau), 1ljr (Theta) and 1fw1 (Zeta). Positional equivalent residues to Arg-81 in rGST M1-1 (6gst) are represented in orange. The sequence alignments were performed using the COMPARER on-line server at http://www.cryst.bioc.cam.ac.uk/~robert/cpgs/COMPARER/comparer.html (Sali and Blundell, 1990), which uses the DiCE structural alignment program (unpublished).

In a study to assess the individual domain contributions to stability and catalytic activity of rGST M1-1 and rGST M2-2, two domain exchanged chimeric mutants were compared to the wild-type enzymes (Luo *et al.*, 2002). The chimeric mutants were created by swapping domain 1 of rGST M1-1 with that of rGST M2-2 to form M(21)-(21), and *vice versa* to form M(12)-(12). Major structural differences were found in the charge cluster at the dimer interface. One such difference is the orientation of Arg-77. In rGST M1-1 Arg-77 is situated in a comparatively open microenvironment, surrounded by five water molecules within a 4 Å radius. At least two of these water molecules are able to hydrogen bond with the side chain amide groups. Salt bridges are formed between Arg-77 and Asp-97, and Arg-77 and E-100

of the same subunit. A more rigid and convoluted subunit interface around Arg-77 is seen within the chimera (Luo *et al.*, 2002). Only one water molecule is found within 4 Å of Arg-77. The altered orientation in M(12)-(12) results in the loss of the salt bridge between Arg-77 and Asp-97 in the same subunit; alternatively Arg-77 forms salt bridges with E-100 and Asp-97 in the other subunit.

Dimeric stability of rGST M(12)-(12) is affected by the increase in both polar and non-polar intersubunit interactions around Arg-77. A comparison of the crystal structures for rGST M1-1 and M(12)-(12) indicate that the orientation of Arg-77 in the charge cluster region appears to play an important role in the stability of the native dimer as well as the monomeric intermediate. By removing Arg-81 from the charge cluster a positional rearrangement of Arg-77 is likely, thereby causing a structural and interactive change within the cluster which will affect its anticipated function.

1.5.3.4. Conformational stability and folding of the cytosolic GSTs

Class Alpha (Wallace *et al.*, 1998a), Pi (Dirr and Reinemer, 1991; Erhardt and Dirr, 1995) and *Sj*26GST (Kaplan *et al.*, 1997) enzymes demonstrate a co-operative two-state unfolding mechanism. Class Mu (Hornby *et al.*, 2000), Sigma (Stevens *et al.*, 1998) and Beta (Sacchetta *et al.*, 1993) GSTs, however, display a multi-state equilibrium pathway involving stable monomeric intermediates.

GSTs are obligate dimers. It is not known if a stable, functional monomeric GST subunit can exist. Abdalla *et al*. (2002) generated a stable monomeric GST subunit by introducing ten site-specific mutations in the dimer interface of hGST P1-1. Although this highly mutated GST P1 monomer retained affinity for a number of electrophilic compounds, it displayed no catalytic activity. There are two main areas of interaction at the inter-subunit interface of hGST A1-1, namely: (1) the hydrophobic lock-and-key motif consisting of the Phe-52 ‘key’ residue, and (2) the Arg/Glu region consisting of Arg-69 and Glu-97 from both subunits (Vargo *et al.*, 2004). The

investigators observed a single mutation at either Phe-52 or Arg-69 in hGST A1-1 resulted in the formation of a monomeric hGST A1-1. It was concluded that Phe-52 and Arg-69 are major determinants of dimer formation and that a single mutation at either position substantially hinders dimerisation (Vargo *et al.*, 2004).

The composition of domain 2 seems to be a major determinant for quaternary structural stability (Luo *et al.*, 2002). This was demonstrated by proteolytic cleavage of human Pi class isoenzyme which resulted in the removal of the first 47 residues of the enzyme, leaving a fragment which still formed a structured dimer due to the intact domain 2 (Aceto *et al.*, 1995; Martini *et al.*, 1993). Studies have also shown that complete removal of domain 1 in the Pi class enzyme destabilised the tertiary structure so much that recognition by anti-(domain 2) antibodies was not possible (Gulick *et al.*, 1992). This indicates that structural integrity of domain 2 appears to depend on interdomain interactions with domain 1.

The dimeric GST structure is required to maintain catalytically functional conformations of the individual subunits, as well as the non-substrate ligand-binding site at the dimer interface (Dirr, 2001). Dependence of subunit stabilisation on quaternary interactions is less significant for classes Sigma and Mu but crucial for subunit stability in class Alpha and Pi GSTs (Dirr, 2001).

1.6. Objectives

From the available theoretical and structural data, rGST M1-1 is a good dimeric model for analysing protein-protein interactions, in particular the inter-subunit mixed-charge cluster at the dimer interface. An R81A mutant was generated to disrupt the mixed-charge cluster at the dimer interface. To ensure that the Phe-56 hydrophobic lock-and-key motif did not interfere with the investigation, a double mutant rGST M1 F56S/R81A was generated to remove this interaction. These perturbations allowed a comparative study between the wild-type and the variant protein. If the mutations

resulted in stable monomeric GST M1 species, structural and functional investigations of the individual subunits and the processes whereby they associate will be undertaken. Questions surrounding the mixed-charge cluster at the dimer interface and the role it plays in complementarity, dimerisation, dimer stability and molecular recognition will be investigated, and the relevant nature of the contributing forces within the cluster analysed.

CHAPTER 2

MATERIALS AND METHODS

2.1. Materials

The restriction enzymes *blp1*, *Sac II,* the plasmid preparation FlexiPrep™ Kit and SDS-PAGE molecular mass markers were purchased from Amersham Biosciences (Buckinghamshire, UK). The Quikchange™ Site-directed Mutagenesis kit was purchased from Stratagene (La Jolla, CA, USA). The wild-type rGST M1-1 expression plasmid was kindly provided by Prof. R. N. Armstrong (Vanderbilt University School of Medicine, Nashville, Tennessee). The F56S rGST M1 expression plasmid was kindly provided by Dr. J. A. T. Hornby (University of the Witwatersrand). 8-Anilino-1-naphthalene sulphonate (ANS) and 1-chloro-2,4-dinitrobenzene (CDNB) were sourced from Sigma (St. Louis, MO, USA). Ultrapure urea was from Merck (Darmstadt, Germany). All other reagents were of analytical grade.

2.2. Generation of the F56S/R81A mutant plasmid

2.2.1. Oligonucleotide primer construction

Oligonucleotide primers used to introduce the R81A mutation were designed for use with the commercially available Quikchange™ Site-Directed Mutagenesis kit (Stratagene). The method was based upon that of Braman *et al*. (1996). The computer software package Gene Runner, v3.04, was used to analyse the primers for dimers, bulges and loops. Primer design was in accordance with the published nucleotide sequence encoding rGST M1-1 (Ji *et al.*, 1992). Note that, however, the F56S rGST M1 expression plasmid served as the template plasmid so that the R81A mutation

was a second codon change substitution. The oligonucleotide primers used to create the F56S/R81A rGST M1 expression plasmid had the following sequences:

F56S/R81A rGST M1 Forward primer:
5'-GC TAC CTT GCC ***GCT*** AAG CAC CAC CTG TGT GG-3'

F56S/R81A rGST M1 Reverse primer:
5'-CC ACA CAG GTG GTG CTT ***AGC*** GGC AAG GTA GC-3'

The replacement of Arg-81 codon with alanine is represented by the bold italicised triplet codons in the F56S/R81A rGST M1 primers. The underlined nucleotides represent a translationally silent mutation that incorporates the *blp1* restriction site.

2.2.2. Site-directed mutagenesis

Site directed mutagenesis was performed using the Quikchange™ Site-directed Mutagenesis kit (Papworth *et al*., 1996). The mutagenesis reaction consisted of 125 ng of each mutagenic primer, 26 ng of double-stranded DNA template (F56S rGST M1), 10 mM dNTP mix and 10 x reaction buffer (100 mM potassium chloride, 100 mM ammonium sulphate, 200 mM Tris-HCl, pH 8.8, 20 mM magnesium sulphate, 1 % Triton® X-100 and 1 mg/ml nuclease-free bovine serum albumin (BSA); supplied with kit) with a final volume of 50 μl. *PfuTurbo*® DNA polymerase (2.5 U/μl) was added finally to the reaction mixture. The thermal cycling conditions used to generate the F56S/R81A rGST M1 expression plasmid are summarised in Table 1.

Table 1. Thermal cycling conditions utilised to generate the F56S/R81A rGST M1 expression plasmid.

PCR Step	Cycles *	Time (seconds)	Temperature (^{o}C)
Denaturing	16	30	95
Annealing	16	60	56
Extension	16	300	68

*One cycle is defined as involving all three steps beginning with the denaturing step and completing with the extension step.

The methylated parental DNA template was removed from the reaction mixture by a one hour, 37 ^{o}C digestion with the restriction enzyme *Dpn*I (10 U/μl; Stratagene, USA). The *Dpn*I-treated DNA was then used to transform *Escherichia coli* XL1-blue supercompetent cells.

Five randomly selected colonies possibly containing the mutant plasmid were grown at 37 ºC in 3 ml Luria-Bertani (LB) media (10 g tryptone, 10 g yeast extract and 5 g NaCl per litre) containing ampicillin at a final concentration of 200 μg/ml. Plasmid DNA was extracted using the plasmid preparation FlexiPrep™ Kit (Amersham Biosciences Buckinghamshire, UK) as per the manufacturers instructions. Extracted plasmid DNA and template F56S rGST M1 were digested overnight with *blp1* at 37 ºC. Wild-type rGST M1-1 expression plasmid was digested overnight with *Sac II* at 37 ºC. The restriction fragments were separated on a 1 % (w/v) agarose gel in TEA buffer containing 0.5 μg/ml ethidium bromide.

To confirm the presence of the mutated R81A codon the cDNA encoding the F56S/R81A double mutation was sequenced by Inqaba Biotechnical Industries (Pty) Ltd (Pretoria, South Africa). Sequencing was achieved using the BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems) and a SCE 2410 Genetic Analyser from Spectrumedix (State College, PA, USA). After confirming the presence of the R81A mutation, the plasmid DNA was transformed into *Escherichia coli* BL21 (DE3) pLysS cells for protein over-expression.

2.3. Protein over-expression and purification

Both wild-type rGST M1-1 and F56S/R81A rGST M1 were purified similarly. Ampicillin resistance is conferred by the Bla gene of the pET-20b(+) parental expression vector and chloramphenicol resistance stems from the pLysS plasmid which encodes the production an acetyl transferase and lysozyme. Lysozyme is an inhibitor of T7 RNA polymerase which keeps the target gene expression suppressed prior to IPTG induction.This prevents toxicity and cell death during the early growth phase of the cells. Upon a freeze-thaw cycle of the cells the expressed lysozyme aids cell lysis. A liquid culture of *Escherichia coli* BL21 (DE3) pLysS containing either the wild-type rGST M1-1 or F56S/R81A rGST M1 expression vectors was grown for 16 hours in LB broth at 37 °C in the presence of 200 μg/ml ampicillin and 35 μg/ml chloramphenicol (referred to as LBamp/camp). A growth study was performed in order to measure the time of growth to mid-log phase (OD_{600} = 1). The stationary phase culture was diluted 10-fold into fresh LBamp/camp broth and grown at 37 °C to mid-log phase (OD_{600} = 1). After reaching mid-log phase, protein expression was induced with the addition of isopropyl-β-D-thiogalactopyranoside (IPTG) to a final concentration of 0.5 mM. Protein expression was allowed to continue for 16 hours after which the cells were harvested by centrifugation at 6 500 x *g* for 30 minutes. All pellets were resuspended in 20 mM MOPS, 0.02 % NaN_3, pH 7.0 solution and then frozen overnight at -20 °C. Frozen cells were allowed to thaw and then lysed by pulse sonication three times for 30 seconds at 4 °C using a Heat Systems Sonicator (Ultrasonics Inc, USA). To remove cellular debris the lysate was centrifuged at 12 000 x *g* at 4 °C. Supernatant containing protein was aspirated and loaded onto a SP-Sepharose™ cation exchange column equilibrated with cell resuspension buffer. The column was washed with 10 column volumes of the same buffer to remove unbound proteins. Bound proteins eluted on a 0-300 mM NaCl gradient and column effluent absorbtion was monitored spectroscopically at 280 nm. Target protein samples were pooled and concentrated under pressure in an ultrafiltration unit using a PM10

membrane with a molecular weight cut-off of 10 kDa. The concentrated protein sample was then loaded on a Sephadex™ G-75 column equilibrated with 20 mM sodium phosphate, 100 mM NaCl, 0.02 % NaN_3, pH 6.5 assay buffer. This buffer was used for all experiments described hereafter. The purity of the protein and native oligomeric state was assessed using 15 % acrylamide SDS-PAGE (Laemmli, 1970) and SEC-HPLC, respectively.

2.3.1. Protein concentration determination

Protein concentration was determined using the molecular extinction coefficient (ε) at 280 nm which was calculated according to the method described by Perkins (1986):

$$\varepsilon(M^{-1}cm^{-1}) = 5\,550\ \Sigma\text{Trp residues} + 1\,340\ \Sigma\text{Tyr residues} + 150\ \Sigma\text{Cys residues} \quad (2)$$

where 5 550, 1 340 and 150 are the molar extinction coefficients at 280 nm ($M^{-1}cm^{-1}$) of tryptophan, tyrosine and cysteine, respectively**.** The extinction coefficients of the dimeric rGST M1-1 and monomeric F56S/R81A were calculated at 81 480 $M^{-1}cm^{-1}$ and 40 740 $M^{-1}cm^{-1}$, respectively, and were then utilised to determine protein concentration spectrophotometrically using the Beer-Lambert Law:

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

where A is the absorbance, ε is the molar extinction coefficient ($M^{-1}cm^{-1}$) at a given wavelength λ, c is the molar concentration and l is the path length in centimetres.

2.4. Characterisation of the F56S/R81A protein

2.4.1. Sodium dodecyl sulphate polyacrylamide gel electrophoresis

Protein over-expression and purity was assessed using discontinuous SDS-PAGE according to the method of Laemmli (1970). Separating gels consisted of 15 % acrylamide: 0.4 % bisacrylamide and stacking gels of 5 % acrylamide:0.1 % bisacrylamide. Protein samples were prepared at a 1:2 dilution with loading buffer (0.5 mM Tris-HCl, pH 6.8, 20 % (v/v) glycerol, 10 % (w/v) SDS, 100 mM β-mercaptoethanol, 0.05 % (w/v) bromophenol blue) and boiled for 5 minutes before loading onto gels. All gels were run at 140 V for 2 hours. Gels were stained with 0.25 % Coomassie Brilliant Blue R250 in 45 % methanol and 10 % acetic acid in water for 2 hours, and then destained in 15 % (v/v) acetic acid and 10 % (v/v) methanol. Molecular mass markers from Amersham Biosciences (Buckinghamshire, UK) contained: phosphorylase b (97 kDa), albumin (66 kDa), ovalbumin (45 kDa), carbonic anhydrase (30 kDa), trypsin inhibitor (20.1 kDa) and α-lactalbumin (14.4 kDa).

2.4.2. Non-denaturing size exclusion high-performance liquid chromatography

The oligomeric status of both wild-type rGST M1-1 and F56S/R81A rGST M1 were determined using non-denaturing size exclusion high-performance liquid chromatography (SEC-HPLC). SEC-HPLC was performed using a LKB 2150 pump (Pharmacia) at a flow rate of 0.5 ml/min. The column used was a TSK G2000 SW_{XL} size exclusion column with resolution of 5–150 kDa (Tosohaas, Japan). The column was connected to a Jasco FP-2020 Plus intelligent fluorescence detector. Final protein concentrations loaded on to the column ranged from 0.5 μM-100 μM. All samples loaded had a final volume of 20 μl. The column effluent was monitored by fluorescence at a flow rate of 0.5 ml/min and recorded using a chart recorder with a chart speed of 30 cm/hr. Protein samples were excited at a wavelength of 295 nm. Depending on the protein analysed, the emission wavelength was either 340 nm or 345 nm.

The standard curve was constructed using a Spectroseries UV100 absorbance detector from SP Thermoseparation products and an HPLC buffer of 0.1 M Na_2HPO_4; 0.1 M Na_2SO_4 and 0.05 % NaN_3, pH 6.7. Proteins used for column calibration were thyroglobulin (670 kDa), γ-globulin (158 kDa), ovalbumin (44 kDa) and myoglobin (17 kDa).

2.4.3. Specific enzyme activity

Enzyme activity was measured spectrophotometrically at 340 nm by monitoring the formation of *S*-2,4-dinitrophenyl glutathione in 0.1 M sodium phosphate, 1 mM EDTA, pH 6.5 (Habig and Jakoby, 1981). The reaction was initiated by the addition of GST enzyme to a 3 ml final volume solution with a final concentration of 1 mM reduced glutathione (GSH) and 1 mM CDNB. The reaction was then followed for 60 seconds and the rates determined using the software on the Jasco spectrophotometer. The protein concentration in the assay mix was chosen to ensure linear progress curves. All reactions were corrected for non-enzymatic controls, which were less than 10 % of the enzymatic rates. The specific activity of the enzyme was determined using an extinction coefficient of 9 600 $M^{-1}cm^{-1}$ for *S*-2,4-dinitrophenyl glutathione.

2.5. Spectroscopic studies of the F56S/R81A

2.5.1. Circular dichroism

Far-UV (200-260 nm) and near-UV (250-310 nm) CD spectra were obtained for native wild-type rGST M1-1 and F56S/R81A rGST M1 on a spectropolarimeter (Jasco, Japan) using a 2 mm path length. Mean residue ellipticity [θ] was calculated using:

$$[\theta] = 100(S)/C.n.l \qquad (4)$$

where S denotes the CD signal averaged over 20 accumulations in millidegrees after subtraction of the solvent baseline. C, n and *l* represent the protein concentration (in mM), number of residues and path length (in cm), respectively.

The spectra were smoothed using the negative exponential smoothing technique (SigmaPlot® v8.0). This methodology smoothes the data using polynomial regression and weights computed from the Gaussian density function with a sampling proportion of 0.1.

2.5.4. Fluorescence spectroscopy

2.5.4.1. Selective tryptophan fluorescence

Intrinsic tryptophan fluorescence emission spectra of native and unfolded (in 8 M urea) wild-type rGST M1-1 and F56S/R81A rGST M1 were obtained by selectively exciting the tryptophan residues at 295 nm. All experiments were conducted at 20 ^{o}C using a Luminescence Spectrometer (Perkin Elmer, USA). The software, FLwinlab v4.00.0 was used for analysis. A quartz cuvette with a pathlength of 1 mm was used in all experiments. A scan rate of 300 nm per minute was used to collect emission spectra and both the excitation and emission bandwidths were set at 5 nm. Spectra were corrected for the Raman peak and plotted using Sigmaplot® software, v8.0. Spectra were smoothed using the negative exponential smoothing technique (SigmaPlot® v8.0). This methodology smoothes the data using polynomial regression and weights computed from the Gaussian density function with a sampling proportion of 0.1.

2.5.4.2. Extrinsic fluorescence spectroscopy - ANS binding

ANS stock was made to a concentration of 2 mM. The concentration was confirmed using its extinction coefficient (λ_{350} = 4 950 $M^{-1}cm^{-1}$). ANS stock was prepared in 50 mM Na_2HPO_4 buffer containing 0.02 % NaN_3.

A final concentration of 200 μM ANS in 20 mM sodium phosphate buffer, pH 6.5, was added to protein concentrations ranging from 2-4 μM. ANS was excited at 400 nm and the emission spectrum recorded from 400 nm to 600 nm. Equilibrium unfolding in the presence of ANS was performed in the following way: dilutions of the protein with urea were set up as in Section 2.6 and incubated at 20 ºC for 1 hour to reach equilibrium. ANS was then added to a final concentration of 200 μM and the protein dilutions left for 1 hour to allow binding of ANS to any exposed hydrophobic patches. Each sample was excited at 390 nm at 20 ºC using a 1 mm cuvette and the emission spectra were recorded over a 390-600 nm wavelength range. The spectra were averaged over 3 accumulations at a scan speed of 300 nm/min. Both excitation and emission slit widths were kept at 5 nm. Spectra were corrected for free ANS and the emission was plotted as a function of urea concentration using SigmaPlot® v8.0.

2.6. Urea-induced equilibrium unfolding of wild-type rGST M1-1 and F56S/R81A rGST M1

Equilibrium unfolding studies were performed at 20 ^{o}C in 20 mM sodium phosphate, 100 mM NaCl, 0.02 % NaN_3, pH 6.5. Fresh 10 M urea stock solutions were prepared weekly and stored at 4 ^{o}C. The molarity of the urea stock solutions was checked by refractometry before use (Pace, 1986).

Protein samples (2-4 μM) were incubated with urea (0-8 M) for at least one hour in order to achieve equilibrium. The sample reactions were subsequently analysed using far-UV CD, tryptophan fluorescence and ANS binding (see section 2.5). Rayleigh scattering, due to aggregation throughout the unfolding process, was monitored by setting the excitation and emission wavelengths at 295 nm.

2.6.1. Reversibility of unfolding

Refolding of 2 μM wild-type rGST M1-1 and 4 μM F56S/R81A rGST M1 incubated in 8 M urea was achieved by the ten-fold dilution of each sample reaction with buffer (20 mM sodium phosphate, 100 mM NaCl, 0.02 % NaN_3, pH 6.5). Fluorescence emission spectra for the refolded proteins (0.1 μM protein in 0.8 M urea) were compared to control emission spectra in which equivalent concentrations of protein and denaturant were used (i.e. 0.1 μM protein in 0.8 M urea).

2.6.2. Data fitting of the equilibrium unfolding transitions

2.6.2.1. Monomeric two-state model (I ↔ U)

The unfolding transition curve of F56S/R81A rGST M1 is characterised by single sigmoidal transitions, suggesting the absence of any thermodynamically stable intermediates. The F56S/R81A rGST M1 unfolding curve was therefore analysed according to a two-state assumption (Pace, 1986) in which the folded (N) and unfolded (U) conformational states are assumed to exist in significant concentrations within the equilibrium unfolding transition region. Therefore for a two-state mechanism:

$$f_U + f_f = 1 \qquad (5)$$

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

$$y = y_f f_f + y_U f_U \qquad (6)$$

where y is the signal obtained for the respective spectroscopic probe, f_f represents the fraction of folded protein, f_U represents fraction unfolded protein. In addition y_f represents the y characteristic for the folded state and can be extrapolated from linear portions of the unfolding transition. The symbol y_u represents the y characteristic for the unfolded state and can be extrapolated from linear portions of the unfolding transition. Combining equations 5 and 6 the fraction of unfolded protein can be calculated from

$$f_U = (y - y_f)/(y_U - y_f) \quad (7)$$

Similarly the fraction of folded or native protein can be calculated from

$$f_f = (y_U - y)/ (y_U - y_f) \quad (8)$$

Then the equilibrium constant (K_{eq}) for the unfolding reaction (K_U) is

$$K_U = f_U/f_f \quad (9)$$

So therefore if equations 7 and 8 are substituted into 9

$$K_U = (y - y_f)/(y_U - y) \quad (10)$$

And

$$\Delta G = - RT \ln K_{eq} \quad (11)$$

Where ΔG_f is the free energy of unfolding, R is the Gas constant, T is temperature in Kelvin and K_{eq} is the equilibrium constant for a reaction.

In order to determine $\Delta G(H_20)$ it is assumed that ΔG has a linear dependence on denaturant concentration [D] for all urea concentrations (Pace, 1986). Therefore

$$\Delta G = \Delta G(H_2O) - m\,[D] \qquad (12)$$

$\Delta G(H_2O)$ represents the free energy required in the absence of denaturant to destabilize the protein. m is the m-value for unfolding and [D] is the denaturant concentration. Combining equations 10, 11 and 12 and rearranging them gives

$$y = [\, y_f + y_U * e^{-(\Delta G(H_20) - m[D])/RT} \,]/[1 + e^{-(\Delta G(H_20) - m[D])/RT}] \qquad (13)$$

Data obtained was fitted to equation 13 using SigmaPlot® v8.0 and the parameters $\Delta G(H_20)$ and m were obtained.

2.6.2.2. Three-state model ($N_2 \leftrightarrow 2I \leftrightarrow 2U$)

Previously, equilibrium unfolding studies revealed that wild-type rGST M1-1 unfolds via a three-state process involving a structured monomeric intermediate (Hornby *et al.*, 2000).

For the three state model, two equilibrium constants can be defined as:

$$K_1=[I]^2/[N_2] \text{ and } \quad K_2=[I]/[U] \qquad (14)$$

Then

$$[U] = 0.25K_1K_2[-(1+K_2) + \sqrt{(1+K_2)^2 + 8(P_t)/K_1}] \qquad (15)$$

The contribution of the three species to the probe signal for the observed property is:

$$Y_{obs} = Y_N + [(Z^*[I] + [U])/P_t\,](Y_U - Y_N) \text{ where } Z = (Y_I - Y_N)/(Y_U - Y_N) \quad (16)$$

From equations 14 and 15, K_1 and K_2 can be calculated, and using equation 12, m_1, m_2, $\Delta G(H_2O)_1$ and $\Delta G(H_2O)_2$ can be derived.

2.6.3. Data analysis and molecular graphics

Unless otherwise stated, all linear and non-linear least-squares fitting of data was performed using Sigma Plot v8.0 (Jandel Corporation). Images of the three-dimensional structures of GSTs reported in this dissertation were either generated using Molscript v2.0 (Kraulis, 1991) interfaced with Pymol (DeLano, 2002); http://www.pymol.org/) or Swiss-pdb Viewer v3.7 (Guex and Peitsch, 1997). Sequence alignments were performed using the COMPARER on-line server at http://www.cryst.bioc.cam.ac.uk/~robert/cpgs/COMPARER/comparer.html (Sali and Blundell, 1990), which uses the DiCE structural alignment program (unpublished).

CHAPTER 3

RESULTS AND DISCUSSION

3.1. Generation of wild-type rGST M1-1 and F56S/R81A rGST M1 expression plasmids

Oligonucleotide directed site-directed mutagenesis was performed as per the manufacturers instructions (Braman *et al.*, 1996) on template F56S rGST M1 plasmid DNA. The PCR reaction mixture was transformed into *Escherichia coli* XLI-Blue supercompetent cells and five subclones were selected for plasmid purification. The purified plasmids were analysed by restriction digest with the results from subclone one shown as an example in Figure 7.

The restriction map of wild-type rGST M1 plasmid is illustrated in Figure 6. Highlighted in red, is the position of engineered restriction site *blp1* which was silently incorporated into the F56S/R81A rGST M1 plasmid. This site was engineered to facilitate differentiation of the template F56S rGST M1 and F56S/R81A rGST M1 plasmids by restriction analyses, as the template F56S rGST M1 plasmid lacks the *blp1* site. The single restriction site for *Sac II*, was used as a selection marker site for the wild-type rGST M1 plasmid.

Restriction enzyme digestion fragments separated on 1 % agarose gel stained with ethidium bromide are shown in Figure 7. Lanes 2 and 6 contain uncut control template F56S rGST M1 plasmid DNA and template F56S rGST M1 plasmid incubated with *blp1*, respectively. The template F56S rGST M1 plasmid was not cut with *blp1* (lane 6). Lanes 3 and 4 contain subclone one plasmid DNA incubated with *blp1* and control subclone one plasmid DNA, respectively. Lane 5 contains wild-type rGST M1 plasmid DNA incubated with *Sac II.* Wild-type rGST M1 plasmid DNA

was linearised by *Sac II.* Uncut subclone one, wild-type rGST M1 and template F56S rGST M1 plasmid DNA share identical migration distances, confirming a similar plasmid DNA size. These results suggest that subclone one plasmid DNA contained the desired R81A mutation.

The plasmid DNA from the five subclones selected (only subclone one shown), were sequenced to confirm the presence of the engineered mutation, and to ensure no undesirable mutations had occurred as a result of the error prone *Taq* polymerase utilised for mutagenesis. A segment of the nucleotide sequence of subclone one containing the *blp1* diagnostic site, and the R81A mutated codon sites are shown in Figure 8. The rest of the sequence was found to be error free.

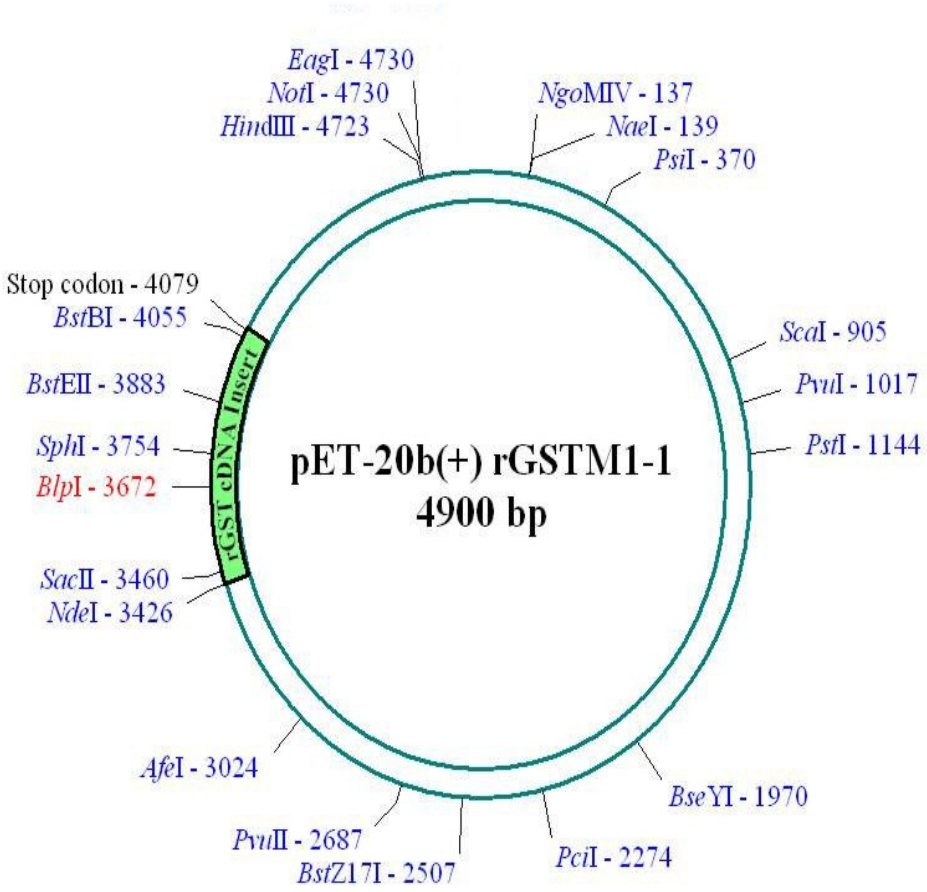


Figure 6. Restriction digest map of pET-20b(+) rGST M1 expression plasmid containing the wild-type rGST M1 encoding cDNA insert.

The diagram shows the unique restriction enzyme site *Sac II*. Marked in green is the position of the rGST M1 cDNA insert. Highlighted in red is the position of the engineered translationally silent *blp1* restriction site incorporated into the F56S/R81A rGST M1 plasmid, which was used to differentiate between parental F56S rGST M1 plasmid and F56S/R81A rGST M1 plasmid.

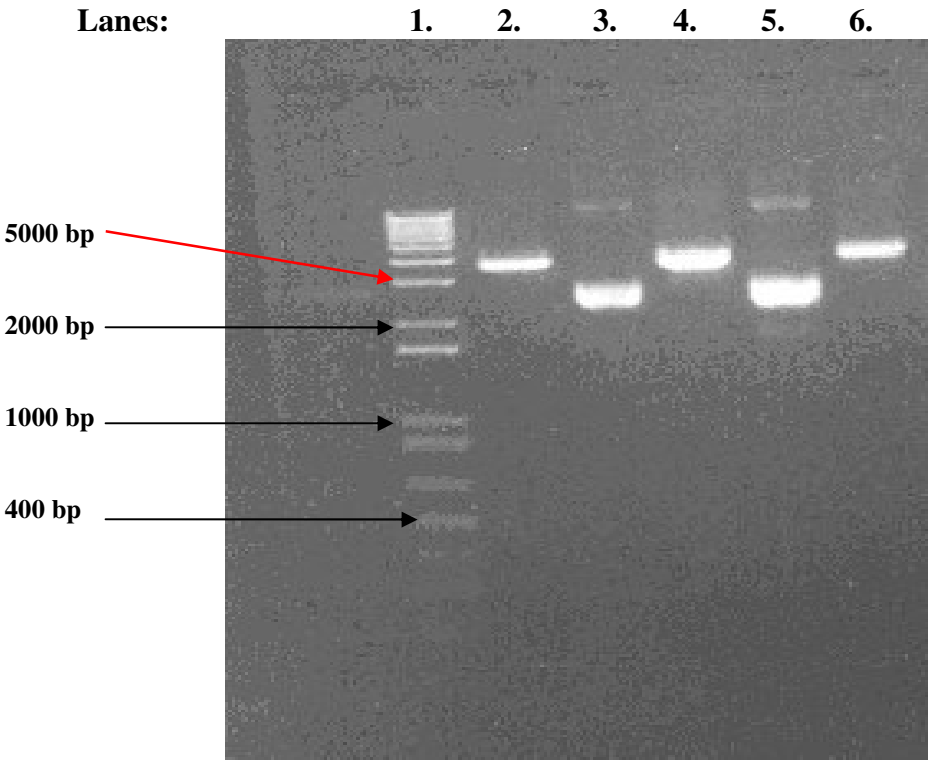


Figure 7. Separation of restriction enzyme digest fragments of the wild-type rGST M1-1 plasmid, F56S/R81A rGST M1 mutant plasmid and F56S rGST M1 plasmid on 1 % agarose gel.

Lane 1 contains a 1 Kb DNA ladder. Lanes 2 and 6 contain uncut control F56S rGST M1 plasmid DNA and template F56S rGST M1 plasmid incubated with *blp1*, respectively. The template F56S rGST M1 plasmid was not cut with *blp1* (lane 6). Lanes 3 and 4 contain subclone one plasmid DNA incubated with *blp1* and control subclone one plasmid DNA, respectively. Lane 5 contains wild-type rGST M1 plasmid DNA incubated with *Sac II.* Wild-type rGST M1 plasmid DNA was cut by *Sac II.*

A.

5' -TAC CTT GCC CGC AAG CAC CAC CTG TGT GGA- 3'

Tyr Leu Ala Arg Lys His His Leu Cys Gly

B.

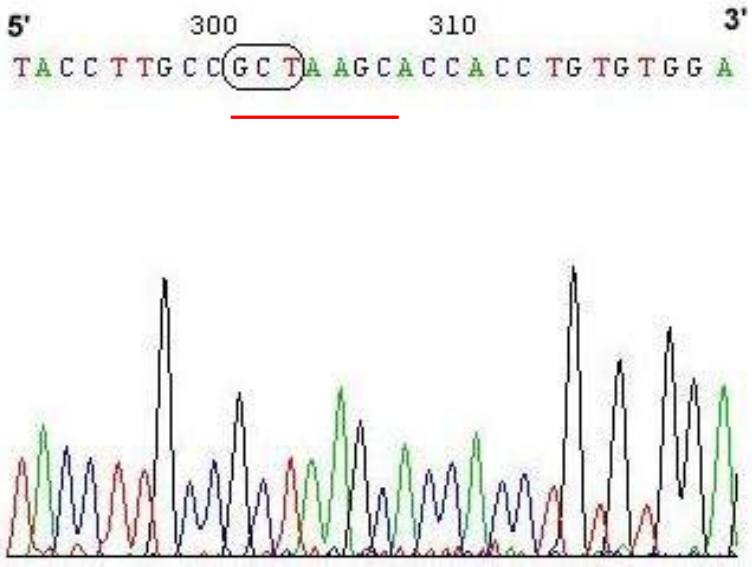


Figure 8. Section of the nucleotide sequence coding region of rGST M1 wild-type with corresponding amino acid sequence and the nucleotide sequence coding region of F56S/R81A rGST M1.

(A) Highlighted in red is the wild-type Arg-81 codon targeted for mutagenesis and corresponding arginine amino acid. **(B)** The nucleotide sequence of subclone one showing the presence of an alanine codon (GCT; encircled) at position 81 of rGST M1, confirming the success of the mutagenesis of F56S rGST M1 to F56S/R81A rGST M1. The engineered *blp1* restriction site is underlined in red.

3.2. Purification of wild-type rGST M1-1 and F56S/R81A rGST M1

The purification of F56S/R81A rGST M1 followed a basic protocol devised by Hornby *et al*. (2002). The F56S/R81A rGST M1 expression plasmid was transformed in to *Escherichia coli* BL-21 (DE3) cells for over-expression. A growth study was performed in order to identify mid-log phase (OD_{600} = 1). An induction study was performed at various times to optimise F56S/R81A rGST M1 expression. IPTG was used in accordance with plasmid manufacturer's recommendations at a concentration of 0.5 mM.

SDS-PAGE separation of the total bacterial protein complement from F56S/R81A rGST M1 over-expression is illustrated in Figure 9A. Figure 9A shows expression of a ~ 26 kDa protein (using 0.5 mM IPTG from 1-16 hours), which corresponds to that of monomeric wild-type rGST M1-1. The ~ 26 kDa protein band is visible throughout the different sample times in Figure 9A and is much darker compared to all the bacterial protein background bands. Although maximum induction was achieved after 3 hours, maximum protein concentration was seen at 16 hours. This was due to the increase in cell mass over time. Cell growth was restricted to 16 hours to avoid cell death typically associated with longer growth times. Thus, the optimal IPTG induction time of the F56S/R81A rGST M1 was 16 hours. Low molecular weight markers were used to determine the apparent molecular weight of F56S/R81A rGST M1. The calibration curves shown in Figures 9B, was constructed from the electrophoretic mobility of these markers.

To determine the solubility of F56S/R81A rGST M1, whole cell extract, lysed cell supernatant and cell pellet were analysed by SDS-PAGE.

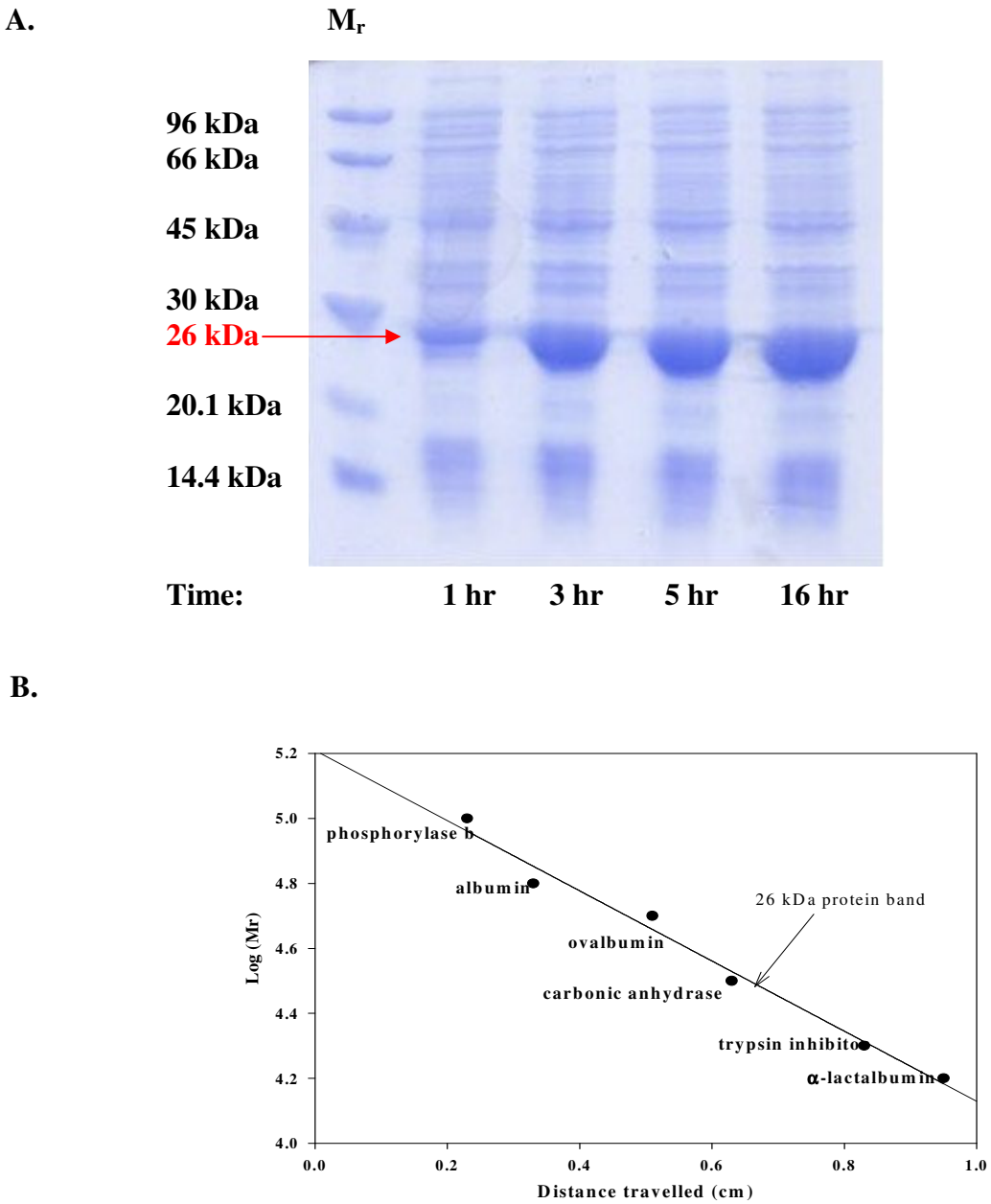


Figure 9. SDS-PAGE separation of the F56S/R81A rGST M1 over-expression induction study.

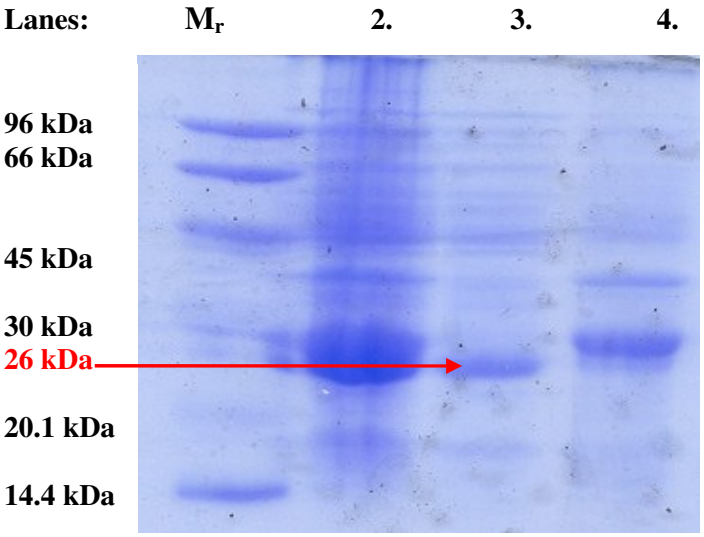
(A) IPTG (0.5 mM) induction study; sample times were between 1 and 16 hours. Lane M_r contains low molecular mass markers phosphorylase b (97 kDa), albumin (66 kDa), ovalbumin (45 kDa), carbonic anhydrase (30 kDa), trypsin inhibitor (20.1 kDa) and α-lactalbumin (14.4 kDa). **(B)** Calibration curve of molecular mass versus distance migrated for molecular weight marker proteins using the same conditions as for purified F56S/R81A rGST M1. Electrophoretic mobility of over-expressed protein corresponds to a size of ~ 26 kDa. The position of heterologously expressed protein and its apparent molecular mass is indicated by the black arrow.

Figure 10A shows the electrophoretogram containing the pellet (lane 4) and cytosolic extracts (lane 3) of F56S/R81A rGST M1. A band of ~ 26 kDa, which represents monomeric wild-type rGST M1-1, was present in the supernatant (Figure 10A, lane 3), confirming the ~ 26 kDa protein band remained in the soluble fraction.

SP-Sepharose cation exchange chromatography was employed to purify F56S/R81A rGST M1. The protein eluted as a single peak using a NaCl gradient as illustrated in Figure 11A. Peak fractions were pooled and concentrated. Peak fractions (Fractions 10-25) assessed using SDS-PAGE were judged to be electrophoretically pure (Figure 12A, lane 3). SEC-HPLC of the concentrated sample detected a singular contaminant with a chromatographic mobility corresponding to a ~ 60 kDa protein. To remove this contaminating protein, a G-75 Sephadex gel filtration chromatography step was added. Figure 11B shows the G-75 Sephadex gel filtration chromatography elution profile. The chromatogram has two overlapping peaks which shows partial separation between the proteins. Fractions completely separate from the overlapping contaminant peak were pooled and their purity evaluated by SDS-PAGE (Figure 12A). SDS-PAGE of pooled fractions of the first peak detected a singular contaminant with an electrophoretic mobility corresponding to a ~ 30 kDa protein (Figure 12A, lane 4). Lane 5 of Figure 12A shows a final purified single band of mass ~ 26 kDa corresponding to the expected size of the GST subunit.

Eight litres of LB cultures yielded ~ 50 mg and ~ 10 mg of wild-type and F56S/R81A rGST M1 proteins, respectively. A possible reason for the difference in protein yield could be attributed to possible lower stability of the F56S/R81A rGST M1. Mutant protein stability *in vitro* is generally correlated with their steady-state level of expression *in vivo* because unstable mutants are subject to degradation (Bowie and Sauer, 1989).

A



B

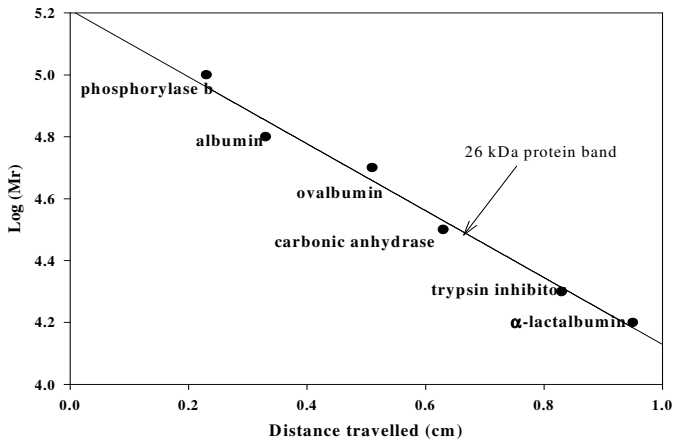
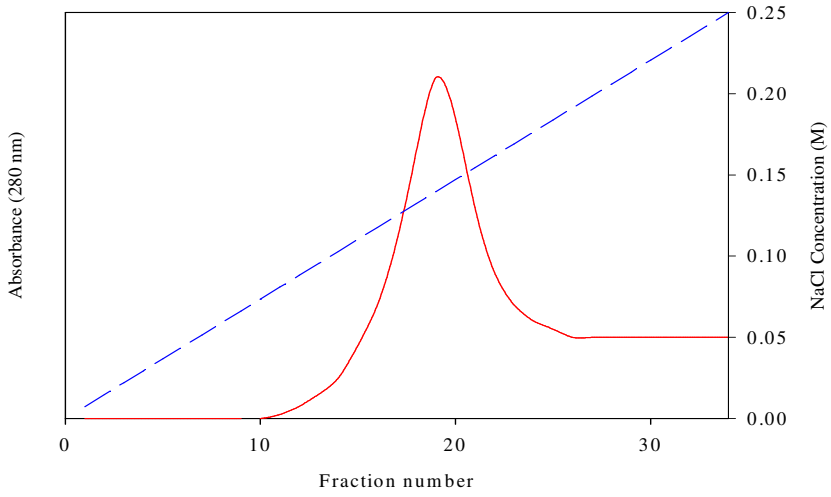


Figure 10. Solubility determination of the F56S/R81A rGST M1.
(A) Lane M_r contains low molecular mass markers. Lane 2: over-expressed total crude cell extract. Lane 3: soluble fraction of the crude cell extract. Lane 4: insoluble fraction of the crude cell extract. **(B)** Calibration curve of molecular mass versus distance migrated for molecular weight marker proteins using the same conditions as for purified F56S/R81A rGST M1. Electrophoretic mobility of over-expressed protein corresponds to a size of ~ 26 kDa. The position of heterologously expressed protein and its apparent molecular mass is indicated by the black arrow.

A.



B.

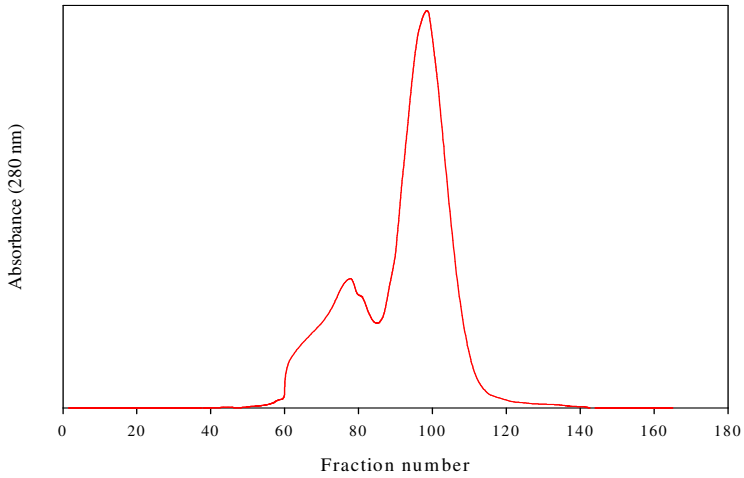
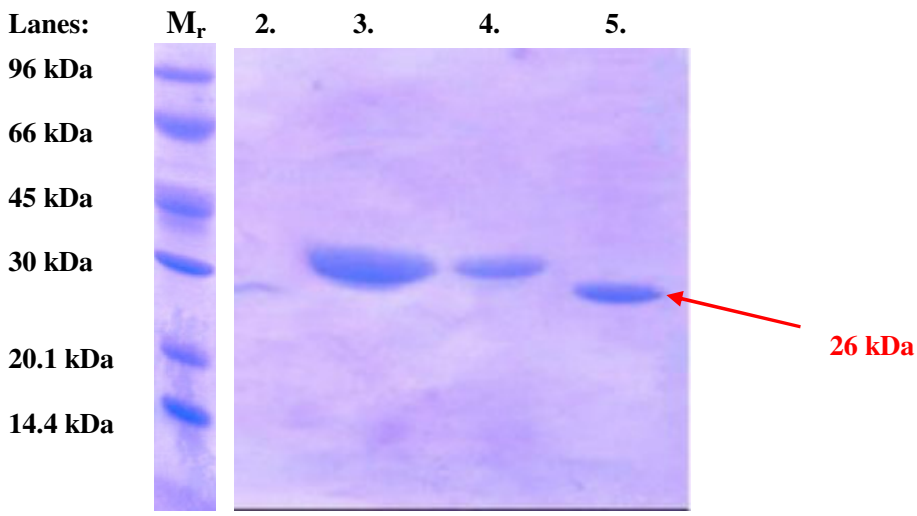


Figure 11. Elution profiles of the chromatography purification steps of F56S/R81A rGST M1.

(A) Elution chromatogram of F56S/R81A rGST M1 protein (—) purified by means of SP-Sepharose cation exchange chromatography. Fractions were eluted using 20 mM MOPS buffer, pH 7.0, containing 0.02 % NaN_3. The protein eluted as a single peak using a 0-0.5 M NaCl gradient (- - -). **(B)** Elution profile of the F56S/R81A rGST M1 protein (—) separated using G-75 Sephadex gel filtration chromatography. Fractions were eluted with a 20 mM sodium phosphate buffer, pH 6.5, containing 0.02 % NaN_3.

A.



B.

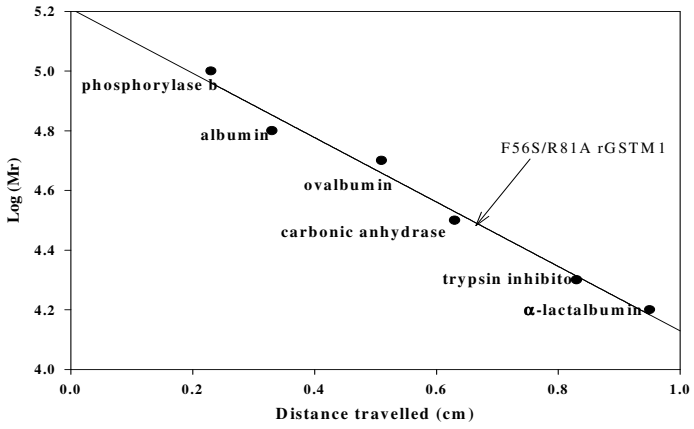


Figure 12. SDS-PAGE separation showing progressive F56S/R81A rGST M1 purification.

(A) Lane M_r: SDS molecular mass marker. Lane 2: hGST A1-1 molecular mass marker. Lane 3: Combined SP-Sepharose fractions, indicating a macromolecular contaminant together with F56S/R81A rGST M1. Lane 4: Combined fractions of the 1st peak obtained from Sephadex G-75. Lane 5: Combined fractions of the 2nd peak (purified F56S/R81A rGST M1) obtained from Sephadex G-75. **(B)** Calibration curve of molecular mass versus distance migrated for molecular weight marker proteins. Electrophoretic mobility of F56S/R81A rGST M1 corresponds to 26 kDa. The position of heterologously expressed protein and its apparent molecular mass is indicated by the black arrow.

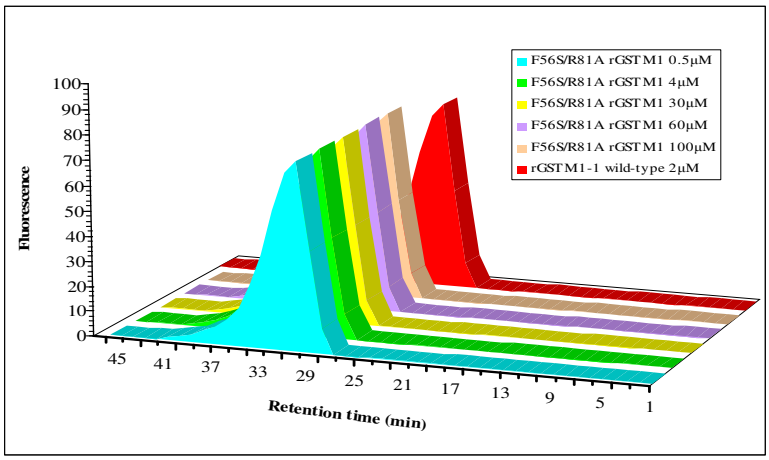
3.3. Physiochemical properties of wild-type rGST M1-1 and F56S/R81A rGST M1

3.3.1. Oligomeric Status

The hydrodynamic volumes and oligomeric state of native wild-type rGST M1-1, and native F56S/R81A rGST M1, were determined using SEC-HPLC as described in section 2.4.2.

The SEC-HPLC elution profiles for wild-type rGST M1-1 and F56S/R81A rGST M1 are shown in Figure 13A. All protein samples eluted as single peaks signifying homogeneous and contaminant-free samples. A calibration curve of log molecular mass vs. retention time in minutes was generated using a set of low molecular weight proteins and is shown in Figure 13B. The arrows in Figure 13B show the retention time for wild-type rGST M1-1 and the F56S/R81A rGST M1 samples. The apparent molecular mass for wild-type rGST M1-1 and F56S/R81A rGST M1 was obtained from the calibration curve using linear regression analysis. The same buffer conditions were used for both the low molecular mass markers and the various samples. The relative molecular mass of wild-type rGST M1-1 was estimated to be 52 kDa (Figure 13B) which compares well with previously published data (Ji *et al.*, 1992), and this confirmed the dimeric status of the wild-type enzyme. The relative molecular mass of F56S/R81A rGST M1 measured at 4 μM was estimated to be ~ 26 kDa, exactly half that of the wild-type. This result corresponds to a monomeric molecular mass of the protein suggesting the possibility of a stable monomeric species.

A.



B.

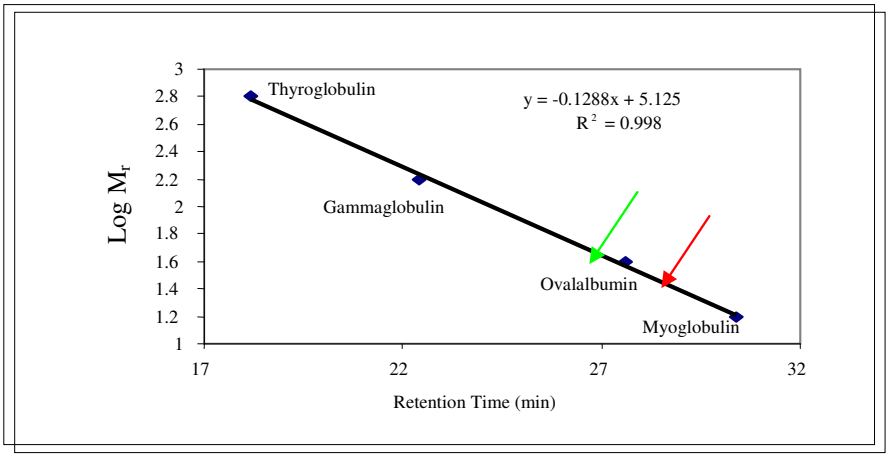
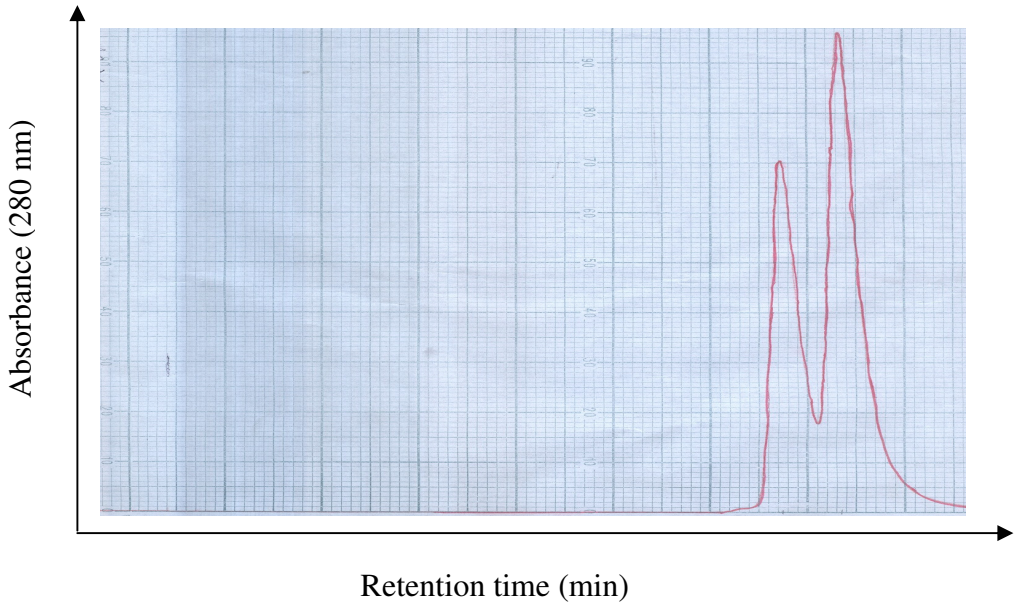


Figure 13. SEC-HPLC elution profiles of wild-type rGST M1-1 and F56S/R81A rGST M1.

(A) Wild-type rGST M1-1 and F56S/R81A rGST M1 in 0.1 M sodium phosphate; 0.1 M sodium sulphate; 0.05 % NaN_3 pH 6.7 buffer separated on SEC-HPLC column at a flow rate of 0.5 ml/min. Elution of the sample peaks was monitored continuously by fluorescence detection. The excitation wavelength was set at 295 nm and the emission at 340 nm and 345 nm for wild-type rGST M1-1 and F56S/R81A rGST M1, respectively. The retention time was recorded using a chart recorder at a chart speed of 30 cm/hr. **(B)** Calibration curve of protein standards: thyroglobulin (670 kDa); gammaglobulin (158 kDa); ovalbumin (44 kDa) and myoglobin (17 kDa). Elution in 0.1 M sodium phosphate; 0.1 M sodium sulphate; 0.05 % NaN_3 pH 6.7 buffer was measured using absorbance at 280 nm. Green and red arrows indicate the retention times of wild-type rGST M1-1 and F56S/R81A rGST M1, respectively.

A



B.

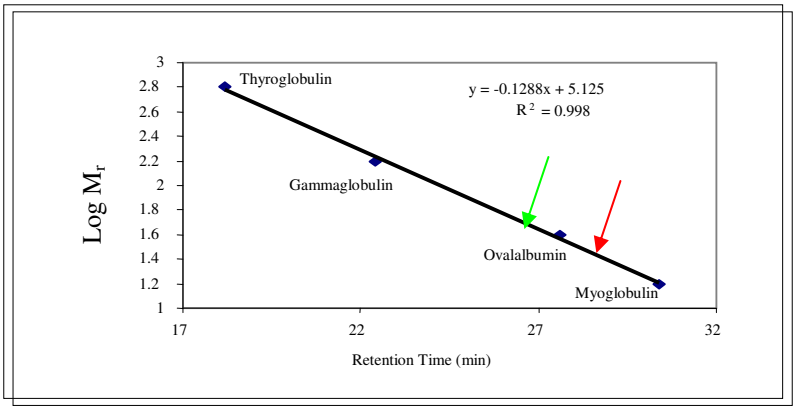


Figure 14. SEC-HPLC elution profiles of wild-type rGST M1-1 spiked with F56S/R81A rGST M1.

(A) SEC-HPLC elution of wild-type rGST M1-1 spiked with F56S/R81A rGST M1 in buffer of 0.1 M sodium phosphate; 0.1 M sodium sulphate; 0.05 % NaN_3 buffer pH 6.7. Elution of the sample peaks was monitored continuously using absorbance at 280 nm at a flow rate of 0.5 ml/min. The retention time was recorded using a chart recorder at a chart speed of 30 cm/hr. **(B)** Calibration curve of protein standards: thyroglobulin (670 kDa); gammaglobulin (158 kDa); ovalbumin (44 kDa) and myoglobin (17 kDa). Elution in 0.1 M sodium phosphate; 0.1 M sodium sulphate; 0.05 % NaN_3 buffer pH 6 was measured using absorbance at 280 nm. Green and red arrows indicate the retention times of wild-type rGST M1-1 and F56S/R81A rGST M1, respectively.

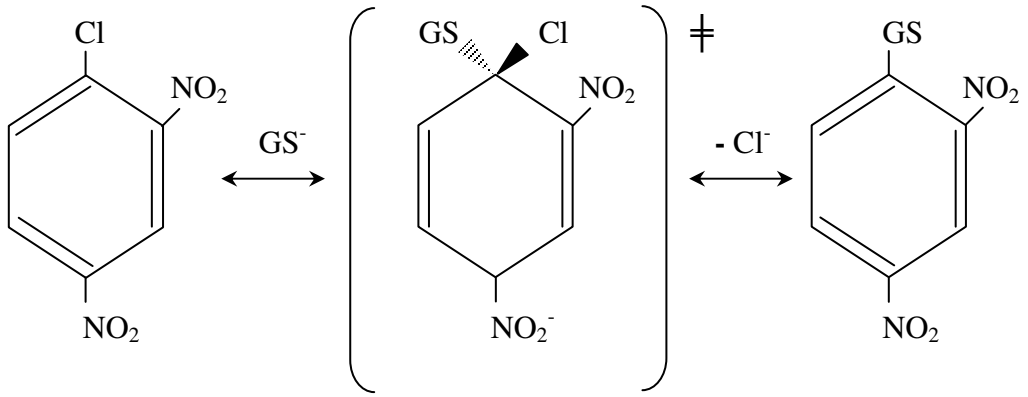
Further concentration dependent investigations were conducted (Figure 13A) up to 100 μM and confirmed the initial size measurement of F56S/R81A rGST M1. There was no peak widening for all concentrations of F56S/R81A rGST M1 indicated no dimer-monomer equilibrium existed. Thus, the F56S/R81A rGST M1 monomeric status was independent of concentration. This suggests the F56S/R81A rGST M1 exists as a stable monomer within the concentration range used. A sample comprising wild-type rGST M1-1 spiked with F56S/R81A rGSTM1 was analysed on the SEC-HPLC system (Figure 14A). The spiked sample consisted of two overlapping peaks corresponding to ~26kDa and ~52kDa, thereby indicating the relative molecular mass of F56S/R81A rGST M1 is not equal the that of wild-type rGST M1-1.

Equilibrium unfolding studies of wild-type rGST M1-1 detected a ~20kDa monomeric intermediate at low protein concentration in the unfolding transition (Hornby *et al.*, 2000). Further investigation in which the lock-and-key motif at the subunit interface of wild-type rGST M1-1 was disrupted confirmed the presence of a monomeric intermediate. In this study Phe-56, the key residue, was replaced with a serine, arginine or glutamic acid. At low protein concentrations substantial amounts of a ~30 kDa monomer were detected when Phe-56 was replaced with serine or arginine. The relative molecular masses of the monomeric intermediates determined in each study are similar to the ~26 kDa F56S/R81A rGST M1. However the F56S and F56R variant proteins which exist in concentration dependent dimer-monomer equilibrium, the F56S/R81A rGST M1 monomer stability is concentration independent. A monomeric form of GST P1-1 was created by introducing 10 site-specific mutations in the subunit interface (Abdalla *et al.*, 2002). The relative molecular mass of monomeric GST P1 was determined to be 26.9 kDa, which compares well to the relative molecular mass of F56S/R81A rGST M1. Native GST P1-1 monomeric subunits can not be isolated. However, the highly mutated GST P1 free monomer subunits do not associate with each other or with a native dimer subunit if the reduced state is maintained. Similarly, F56S/R81A rGST M1 monomers do not associate with each other or with native wild-type rGST M1-1 subunits. Recently, Vargo *et al*. (2004) investigated the importance of two main areas

of interaction at the inter-subunit interface of hGST A1-1: (i) the hydrophobic lock-and-key motif consisting of the Phe-52 key residue, and (ii) the Arg/Glu region consisting of Arg-69 and Glu-97 from both subunits. These areas of interaction are equivalent to the hydrophobic lock-and-key motif consisting of the Phe-56 key residue and the mixed-charge cluster at the dimer interface in rGST M1-1. The investigators found that a single mutation at either Phe-52 or Arg-69 in hGST A1-1 greatly shifted the dimer-monomer equilibrium towards monomer. It was concluded that Phe-52 and Arg-69 are major determinants of dimer formation and that a single mutation at either position substantially hinders dimerisation (Vargo *et al.*, 2004). However, disruption of the hydrophobic lock-and-key together with the mixed-charge cluster at the dimer interface in rGST M1-1 did not shift dimer-monomer equilibrium towards monomer, rather they prevented dimer formation. This proves the mixed-charge cluster at the dimer interface in rGST M1-1 is essential for dimerisation in rGST M1-1.

3.3.2. Specific activity

The GST-catalysed conjugation of GSH to the aryl chloride CDNB (Habig and Jakoby, 1981), is shown in Figure 15. The CDNB nucleophilic aromatic substitution reaction (S_NAR reaction) takes place via formation of 1-chloro-1-(S-glutathionyl)2,4-dinitrocyclohexadienate, a Meisenheimer complex transition state (Graminski *et al.*, 1989). The GST active site stabilises the formation of this dead-end Meisenheimer complex for efficient S_NAR reactions (Bico *et al.*, 1994; Graminski *et al.*, 1989).



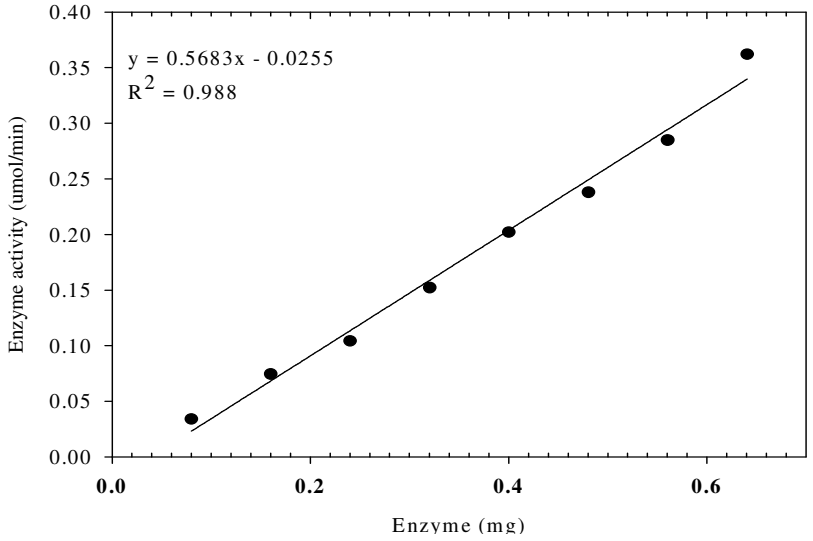
1-chloro-2,4-dinitrobenzene **1-chloro-1-(S-glutathionyl)-2,4-dinitrobenzene** **1-(S-glutathionyl)-2,4-dinitrobenzene**

Figure 15. A schematic representation of the GST catalysed conjugation of GSH to CDNB.

This S_NAR reaction involves the nucleophilic attack by the GSH thiolate anion (GS^-) on the electrophilic carbon on the phenyl ring, followed by expulsion of the chloride ion (Cl^-). The brackets indicate the proposed Meisenheimer (σ) complex in which the sulphur and the chloride atoms are linked to the same carbon atom. Adapted from (Graminski *et al.*, 1989).

GST/CDNB-conjugation assays (Habig and Jakoby, 1981), using 1 mM concentrations of GSH and CDNB and 0.025-0.5 μM concentrations of F56S/R81A rGST M1 indicated that F56S/R81A rGST M1 was catalytically inactive. However, further investigations utilising higher final protein concentrations ranging from 1-8 mM, as shown in Figure 16, resulted in a specific activity of 0.6 μmol/min/mg. The F56S/R81A rGST M1 protein displayed less than 1 % activity relative to the reported wild-type value of 50 μmol/min/mg (Hornby *et al.*, 2002).

A.



B.

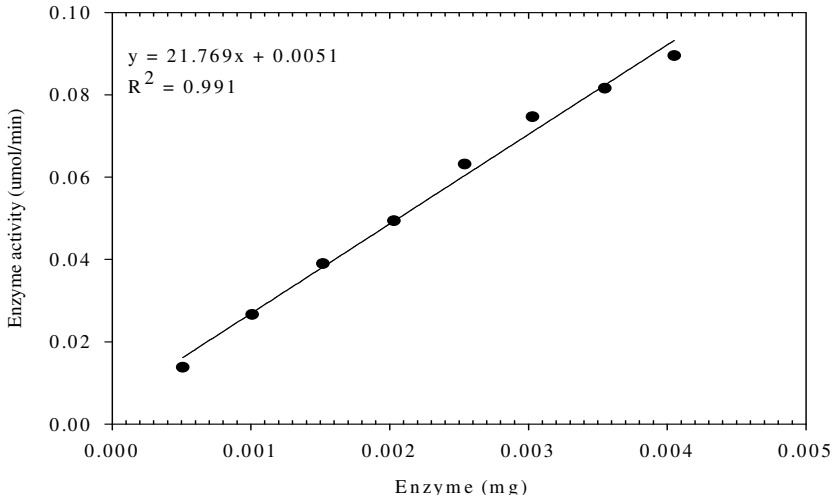


Figure 16. Specific activity of (A) F56S/R81A rGST M1 and (B) wild-type rGST M1-1 proteins.

The solid lines represent the linear fit to the experimental data (●). The specific activity of each protein was determined from the slope of the graph using Sigma Plot v8.0, with linear regression analysis (R^2 = 0.98 for F56S/R81A rGST M1; R^2 = 0.99 for wild-type rGST M1-1.

Other GST subunit interface variants, for example the F52S (Sayed *et al.*, 2000) and F52A (Vargo *et al.*, 2004) of hGST A1-1, and the F56R, F56S and F56E of rGST M1-1 (Hornby *et al.*, 2002) show similar results where dramatically decreased activity was observed for each protein relative to the wild-type protein. Previous functional studies of rGST M1-1 have suggested that there is a close link between the structure of the dimer interface and the GSH binding site (Hornby *et al.*, 2002). Flanking the lock-and-key motif in rGST M1-1 is the active site as well as α-helix 2 and β-strand 3. The Phe-56 is located on a connecting loop between α-helix 2 and β-strand 3. The base of the 56-loop is in van der Waals contact with Tyr-6, which directly interacts with the sulphur of GSH. The amino acid substitution within this loop possibly contributes to the decreased activity. Asn-58 is one of the major residues interacting with bound GSH in the complexed G-site (Ji *et al.*, 1992). Asn-58 shares a relatively close proximity with the Phe-56 residue at the rGST M1-1 subunit interface. Together, the substitution of the Phe-56 residue and the absence of this Asn-58/Phe-56 positioning perhaps play a role in the decreased activity of the F56S/R81A rGST M1 protein. Replacement of Phe-56 'key' residue at the subunit interface of rGST M1-1 (Hornby *et al.*, 2002), as well as that of the topologically equivalent hGST P1-1 (Tyr-50; Stenberg *et al.*, 2000), affects the catalytic activity of these enzymes. Small residue (Y50A in Pi), hydrophilic residue (F56R, F56S, F56E in Mu; Y50R and Y50S in Pi) or bulky hydrophobic residue (Y50L in Pi) substitutions diminish catalytic activity in the respective enzymes. As with the F56S/R81A rGST M1, this reduction in activity is partly a result of the altered conformation of the loop which forms part of the G-site of the protein. On the contrary, aromatic substitution favours enzyme activity (Y50F in Pi; F52Y in Alpha). Several GST isoenzymes have either tyrosine or phenylalanine at this position; therefore it would be unlikely for aromatic substitution to hinder catalytic functioning. Additional electrostatic interactions at the dimer interface are provided by the presence of tyrosine in the key position compared with the more commonly occurring phenylalanine residue. Thus, a less conformationally restricted enzyme is seen with the removal of the hydroxyl group of tyrosine, as in the Y50F hGST P1-1 variant protein, and consequently a higher catalytic activity is observed (Stenberg *et*

al., 2000). Although not directly a part of the G-site, the lock-and-key motif of the class Alpha, Mu and Pi GSTs impacts on the structure involved in the GSH binding region, and consequently affect enzyme activity.

The contribution of inter-subunit interactions to the GST dimeric structure, and consequently catalytic functionality, remains uncertain. The dimeric structure of GSTs is necessary for the construction of fully functional catalytic sites situated near the subunit interface in the Alpha, Mu, Pi, and Sigma class enzymes. The majority of interactions between GSTs and glutathione are contributed by domain 1. Domain 1 contains all of the G-site residues bar 1 aspartate from domain 2 of the other subunit (Mu, Asp-105; Pi, Asp-96; Alpha, Asp-101). This Asp residue interacts with the γ-glutamyl moiety of GSH. Only alpha class has an additional contribution from Arg-130 of the other subunit. Dissociation of the dimer has resulted in catalytic inactivation as the γ-glutamyl moiety of GSH seems to be the major G-site binding determinant (Adang *et al.*, 1989; Adang *et al.*, 1990). This can be expected since it is the interaction with the enzyme that is most similar amongst the different gene classes (Dirr *et al.*, 1994). Mutagenic studies have highlighted the integral role of this residue in homologous Pi class GST.

The involvement of inter-subunit interactions in protein functionality was investigated in the generation of a structurally stable but catalytically inactive monomeric species of hGST P1-1 (Abdalla *et al.*, 2002). By introducing 10 site-specific mutations in the dimer interface of hGST P1-1, a monomeric GST P1 variant protein was constructed. Three residues involved in the lock-and-key motif were mutated (Met-92 into Glu, Gly-96 into Gln and Tyr-50 into Glu). Structural changes in the G-site of the protein resulted in the inactive catalytic nature of the monomer. In addition, the monomer may affect the binding of GSH to the G-site as the stabilising contacts to GSH from the adjacent subunit are not present. Altogether, it is unlikely that a single point mutation alone was responsible for the loss in activity in GST P1. A heterodimeric protein composed of subunits from the fully functional wild-type hGST P1-1 and the nearly inactive Y50A hGST P1-1 was constructed to identify the

importance of Tyr-50 (Hegazy *et al.*, 2004; Stenberg *et al.*, 2000). Tyr-50 is located far from the catalytic site of hGST P1-1. The specific activity of Y50A hGST P1-1 decreased 25 000 fold in comparison with the wild-type enzyme. The heterodimer displayed a 100 fold decrease in activity relative to the wild-type protein (Hegazy *et al.*, 2004). For the heterodimers, the decrease in activity might be due to failure of dimerisation. The diminished activity of both Y50A hGST P1-1 and the heterodimers was due to a reduction in the conformational flexibility of α-helix 2, and consequently a change in the G-site geometry (Hegazy *et al.*, 2004).

3.4. Spectral properties of the wild-type rGST M1-1 and F56S/R81A rGST M1

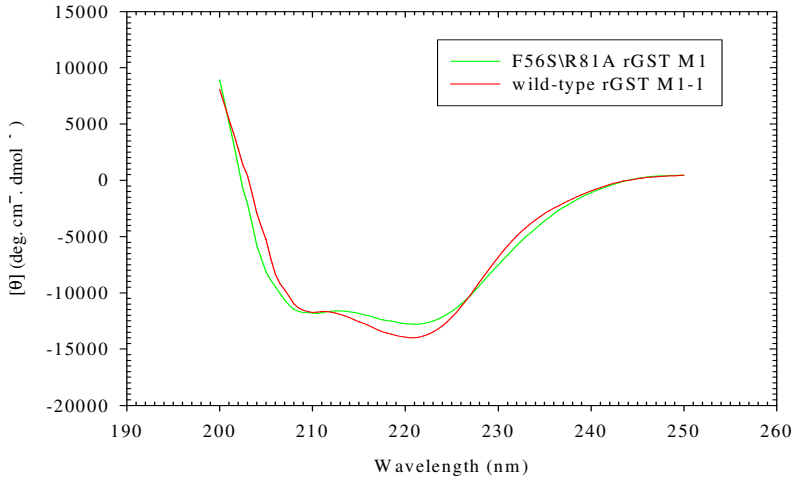
3.4.1. Far-UV circular dichroism

In proteins, far-UV CD spectra (210-250 nm) are primarily influenced by the amide interactions in the polypeptide backbone (Johnson, 1990). Accordingly, this makes it an excellent probe to measure the secondary structural content of proteins. Thus, far-UV CD was used to compare wild-type rGST M1-1 and F56S/R81A rGST M1.

Figure 17A shows far-UV CD spectra of wild-type rGST M1-1 and F56S/R81A rGST M1 proteins in their native conformational states. Wild-type rGST M1-1 and F56S/R81A rGST M1 CD spectra exhibit two minima occurring around 208 nm and 222 nm, characteristic of a predominantly α-helical protein. The native-like F56S/R81A rGST M1 spectrum differs slightly from the wild-type in the 228-238 nm range, and exhibits a reduced intensity minimum at 222 nm. After one week F56S/R81A rGST M1 lost approximately 30 % of its α- helical content, signifying a less stable conformation than the wild-type relative to time.

A monomeric form of GST P1-1 was created by introducing 10 site-specific mutations in the subunit interface (Abdalla *et al.*, 2002). Overall, secondary structure of the highly mutated engineered monomeric GST P1-1 monomer was similar to that of wild-type GST P1-1. In a similar study by Vargo *et al*. (2004), the hGST A1-1 monomer shared overall secondary structural similarity to that of wild-type hGST A1-1. Similarly, the global secondary structure of F56S/R81A rGST M1 was relatively unaltered in comparison to the wild-type, even though the introduced amino acid substitutions prohibited subunit association. Although the far-UV CD spectra of wild-type rGST M1-1 and F56S/R81A rGST M1 proteins are very similar and follow the same trend, they are not superimposable. Dissimilar environments surrounding the tryptophan residues could be responsible for the minor variation observed in the 228-238 nm range (Vuilleumier *et al.*, 1993). Secondary structural modification can occur due to the assembly or disassembly of a quaternary structure (Hennessey *et al.*, 1982). Consequently, preventing protein association in F56S/R81A rGST M1 may account for the loss of intensity at 222 nm, and is indicative of a diminished α-helical content.

A.



B.

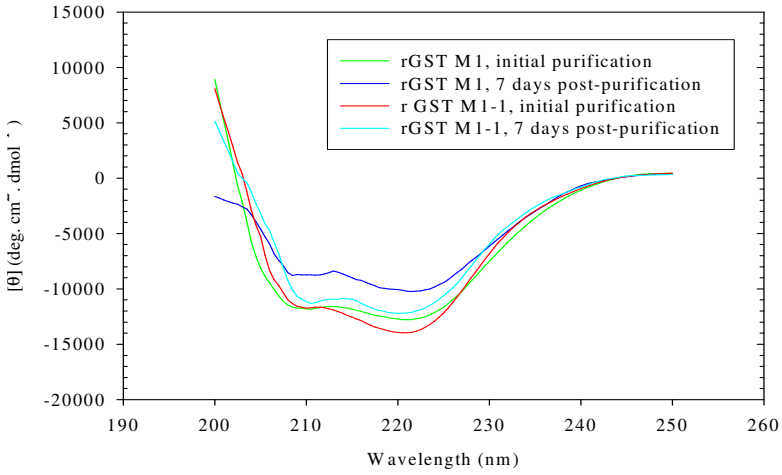


Figure 17. Far-UV circular dichroism spectra of wild-type rGST M1-1 and F56S/R81A rGST M1.

(A) Spectra for the 2 µM native folded wild-type and 4 µM F56S/R81A rGST M1. The spectral analyses were performed in 5 mM sodium phosphate buffer, pH 6.5, containing 0.02 % NaN_3. Each spectrum represents an average of 20 accumulations. **(B)** Spectra of wild-type and F56S/R81A rGST M1 at different stages after purification. Readings were performed in 5 mM sodium phosphate buffer, pH 6.5, containing 0.02 % NaN3, and were taken immediately after purification and seven days post-purification to monitor conformational changes with increasing time. Each spectrum represents an average of 20 accumulations.

3.4.2. Near-UV circular dichroism

A protein's ability to absorb radiation in the near-UV (255-350 nm) spectral region originates from the orientations and interactions of the aromatic amino acid side chains in relation to their surrounding environment. The conformational information generated reflects exclusively on the localized surroundings of the aromatic residues. However, the general cooperative nature of protein structure means that alterations in the aromatic near-UV CD spectra reveal more global tertiary changes in protein conformation. Accordingly, near-UV CD was used to compare the tertiary structures of wild-type rGST M1-1 and F56S/R81A rGST M1.

The near UV-CD spectra, obtained from wild-type rGST M1-1 and F56S/R81A rGST M1, are displayed in Figure 18. Both spectra show similar patterns and are dominated by four tryptophan residues, two of which, (Trp-7 and Trp-45), are located in the GSH binding domain near the active site, and act as hydrogen bond donors for enzyme-bound GSH. The wild-type rGST M1-1 spectrum illustrates a minor trough at ~ 295 nm attributed to tryptophan. Additionally, tryptophan and tyrosine residues predominantly create the two large positive peaks at ~ 280 and ~ 288 nm, which is preceded by two minor troughs at 262 nm and 268 nm resultant from phenylalanine. This spectral pattern is consistent with previously published spectra (Zhang and Armstrong, 1990). The F56S/R81A rGST M1 spectrum shows greatly reduced troughs in the phenylalanine-dependent range, indicating an altered local environment around the phenylalanines. This could be a result of the F56S mutation. F56S/R81A rGST M1 CD signal enhancement between the ~275 nm and ~288 nm region suggests the micro-environments around the tryptophan and tyrosine residues of F56S/R81A rGST M1 differs from those of the wild-type enzyme.

Replacement of Phe-56 at the subunit interface of rGST M1-1 revealed the presence of a monomeric intermediate in the unfolding pathway at low protein concentration (Hornby *et al.*, 2002).

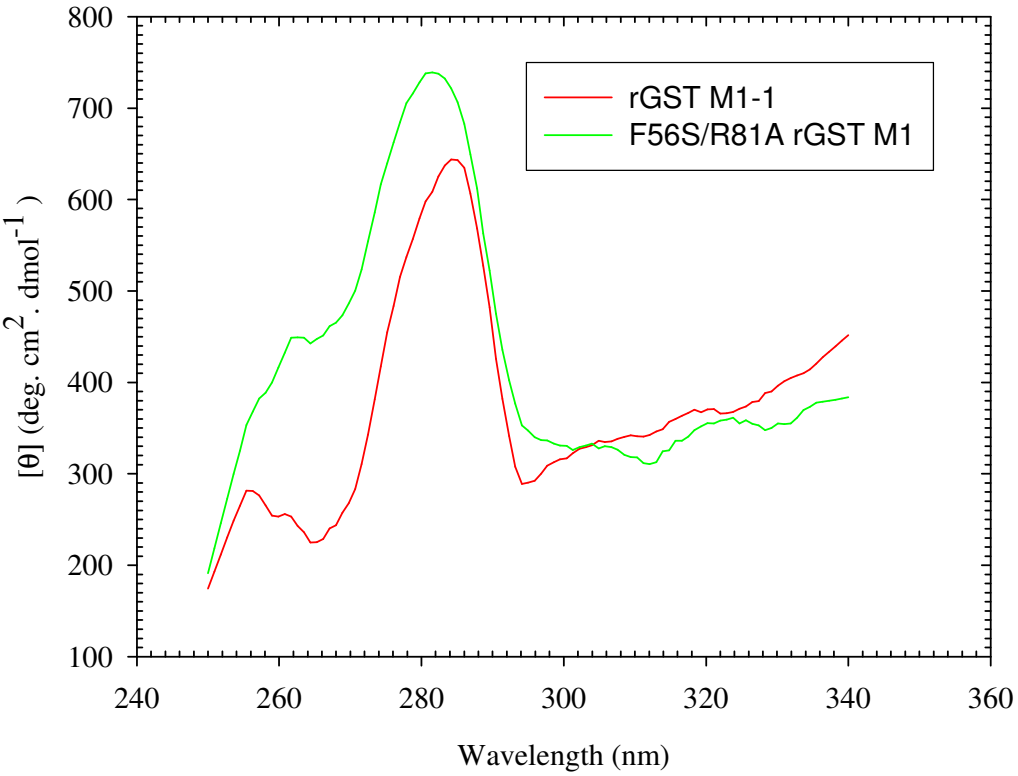


Figure 18. Near-UV circular dichroism spectra of wild-type rGST M1-1 and F56S/R81A rGST M1.

Spectral analyses were performed using 2 μM native folded wild-type and 4 μM F56S/R81A rGST M1 in 5 mM sodium phosphate buffer, pH 6.5, containing 0.02 % NaN_3. Each spectrum represents an average of 20 accumulations.

The tertiary structures of the various Phe-56 rGST M1-1 monomeric intermediates are not identical to the subunit in dimeric wild-type rGST M1-1 (Hornby *et al.*, 2002). These structural differences are a result of the loop region containing Phe-56 and the α-helix 4 and α-helix 5 assuming an alternate, more dynamic conformation. Dimerisation induces the loop to assume a native-like conformation, enabling the correct docking of the Phe-56 side chain into the neighbouring subunit (Codreanu *et al.*, 2005). Similarly, these results reflect on the overall packing of the F56S/R81A rGST M1, and suggest a more loosely packed conformation surrounding the loop region containing Phe-56 due to the absence of dimerisation.

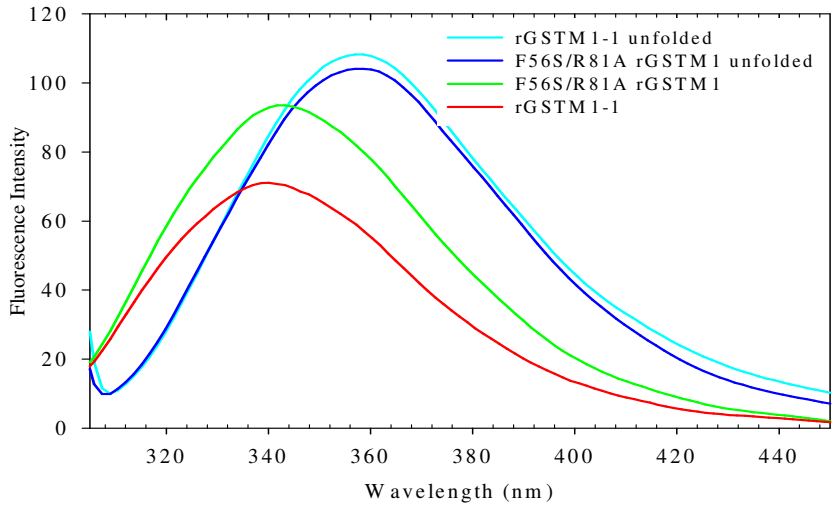
3.4.3. Fluorescence spectroscopy

Each subunit of wild-type rGST M1-1 and F56S/R81A rGST M1 contains four tryptophans located at positions 7, 45, 146 and 214. Crystallographic analysis of the rGST M1-1 structure suggests that the indole side chain of residues Trp-146 and Trp-214 are partially exposed to solvent. Two of these tryptophans are in domain 1 (positions 7 and 45) and the other two in domain 2 (positions 146 and 214), thus providing an ideal probe to monitor global structural changes.

Figure 19 shows emission spectra generated by selective tryptophan excitation at 295 nm for wild-type rGST M1-1 and F56S/R81A rGST M1 proteins in the folded and unfolded states. Fluorescence emission maxima values are dependant on the environment of the fluorophore (Lakowicz, 1999).The emission maxima describing the native conformations of the wild-type rGST M1-1 and F56S/R81A rGST M1 are 340 and 345 nm, respectively indicating the tryptophans are located in a hydrophobic environment. However, the F56S/R81A rGST M1 tryptophans seem less buried in their environment, due to the lack of association of the individual F56S/R81A rGST M1 subunits. The red-shifted F56S/R81A rGST M1 spectrum suggests that the mutations at the lock-and-key motif and the mixed-charge cluster modify the environment of the tryptophan residues in F56S/R81A rGST M1. The fluorescence

intensities of the folded F56S/R81A rGST M1 were approximately 20 % weaker than that observed in the wild-type. Both unfolded wild-type and unfolded F56S/R81A rGST M1 enzymes show red-shifted maximum emission intensities at a wavelength of 360 nm, indicating similar denatured states and increased if (not complete) tryptophan exposure to solvent. The red-shift in unfolding was accompanied by enhanced fluorescence intensity relative to that of the native proteins. This suggests that the tryptophan residues, in the folded state of each protein, are in a quenching environment. From the emission spectra, it is clear that the tryptophans in domain 2 of F56S/R81A rGST M1 have a less packed environment which is confirmed by the near-UV CD data.

A.



B.

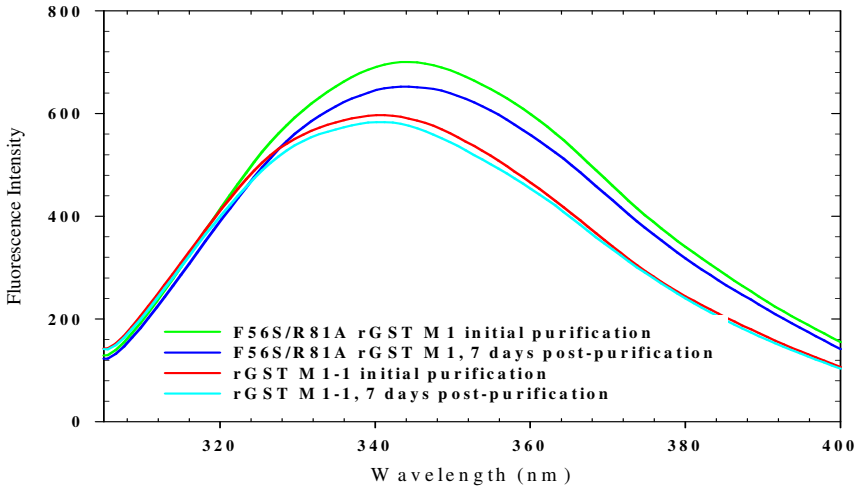


Figure 19. Fluorescence emission spectra of rGST M1-1 and F56S/R81A rGST M1.

(A) Fluorescence emission spectra of 2 μM rGST M1-1 and 4 μM F56S/R81A rGST M1 in 5 mM sodium phosphate buffer, pH 6.5, excited at 295 nm. The wavelengths of maximum emission intensity of native rGST M1-1 and native F56S/R81A rGST M1 are 340 and 345 nm, respectively. Unfolded rGST M1-1 and unfolded F56S/R81A rGST M1 enzyme both show maximum emission intensity at a wavelength of 360 nm. **(B)** Emission spectra of wild-type and F56S/R81A rGST M1 at different stages after purification. Readings were performed in 5 mM sodium phosphate buffer, pH 6.5, containing 0.02 % NaN_3, and were taken immediately after purification and seven days post-purification to monitor conformational changes

3.4.3.1. Non-substrate ligand binding properties

The anionic dye, ANS, has been used as a probe for the detection of non-polar surfaces on proteins (Abdalla *et al.*, 2002; Hornby *et al.*, 2000). When ANS binds a hydrophobic surface its fluorescence emission maximum shifts to a lower wavelength. This mechanism of ANS binding has been questioned. A study on ANS binding to BSA has revealed that there are ion pair interactions involved between the sulfonate moeity of ANS and cationic moeities in the protein (Matulis and Lovrien, 1998). Consequently, ANS binding would be dependent on pH and amino acid composition of the protein. It has been predicted by Sluis-Cremer *et al.* (1996) that ANS binds a single site (L site; ligand-binding site) in class Pi which is found at the dimer interface. Sayed *et al.* (2002) have shown using isothermal titration calorimetry which is a direct and sensitive method of measuring binding, that ANS binds two sites per dimer and that the sulfonate moeity does not contribute significantly to the binding. Ji *et al.* (1996) identified an L-site in GST S1-1, at the dimer interface, using *S*-(3-iodobenzyl) glutathione as a ligand. Nichole Kinsley (unpublished MSc results), has demonstrated that ANS binds at the H-site in class Mu. All these sites are characteristically hydrophobic and it has not been reported that ANS binds any GST via ion pair formation. However, ANS is an amphipathic molecule. Recently, the interactions between ANS and class Alpha have been shown not to be exclusively hydrophobic, as previously assumed. The hydrophobic anilino and naphthyl rings of ANS occupy the nonpolar H-site, whereas the negatively charged sulfonate group can occupy the interface between the G-site and H-site (Dirr *et al.*, 2005). Consequently, it was considered appropriate to use ANS as a probe in this study to measure overall surface hydrophobicity of F56S/R81A rGST M1, due to increased subunit interface exposure.

Fluorescence emission spectra of ANS bound to wild-type rGST M1-1 and F56S/R81A rGST M1 proteins are shown in Figure 20. In the absence of protein, ANS was observed to have an emission maximum at 530 nm. Upon binding of the

ligand to the wild-type and F56S/R81A rGST M1, a blue-shift in emission maxima was observed. The emission maxima for the wild-type and F56S/R81A rGST M1 were found to be 485 nm and 475 nm, respectively. When compared with the emission maximum of 545 nm for ANS in water and 454 nm for ANS bound to the highly hydrophobic site in apomyoglobin (Stryer, 1965), it is apparent that the ANS binding sites of the wild-type and F56S/R81A rGST M1 are not exclusively hydrophobic, which agrees with previous publications regarding the wild-type protein (Hornby *et al.*, 2002). In addition, the shift in emission maximum is indicative of differences between the overall surface hydrophobicity and to a certain extent the hydrophobicity of the H-sites of wild-type rGST M1-1 and F56S/R81A rGST M1. The extent of the blue-shift in emission maximum is determined by the polarity of the ANS binding site; the lower the polarity, the greater the blue-shift (Lakowicz, 1999; Sayed *et al.*, 2002). Therefore, the ANS binding site of F56S/R81A rGST M1 is less polar than that of the wild-type rGST M1-1, but both are more polar than class Alpha (Sayed *et al.*, 2002).

The spectral blue-shift of the rGST M1-1 and F56S/R81A rGST M1 was accompanied by a fluorescence intensity enhancement (see Figure 20). Compared to wild-type rGST M1-1 bound to ANS, a 2.5 fold enhanced fluorescence intensity is observed for the F56S/R81A rGST M1 when bound to ANS. ANS fluorescence is quenched by water. Consequently, the fluorescence intensity of protein-bound ANS is highly dependent on its accessibility to water (Kirk *et al.*, 1996). The fluorescence intensity enhancement of ANS bound to F56S/R81A rGST M1 (Figure 20) indicates that the dye is less exposed to solvent than when the dye is bound to wild-type rGST M1-1. This is possibly due to the decrease in the solvent exposure at the H-site, the region to which ANS has been shown to bind in class Alpha (Dirr *et al.*, 2005) and Mu (Nichole Kinsley unpublished MSc results) enzymes. The fluorescence enhancement of ANS bound to rGST M1-1 is low, indicating a greater exposure of ANS to solvent when bound to rGST M1-1. This result is similar in class Alpha (Sayed *et al.*, 2002). This is indicative of the decreased exposure of ANS to solvent in the respective proteins.

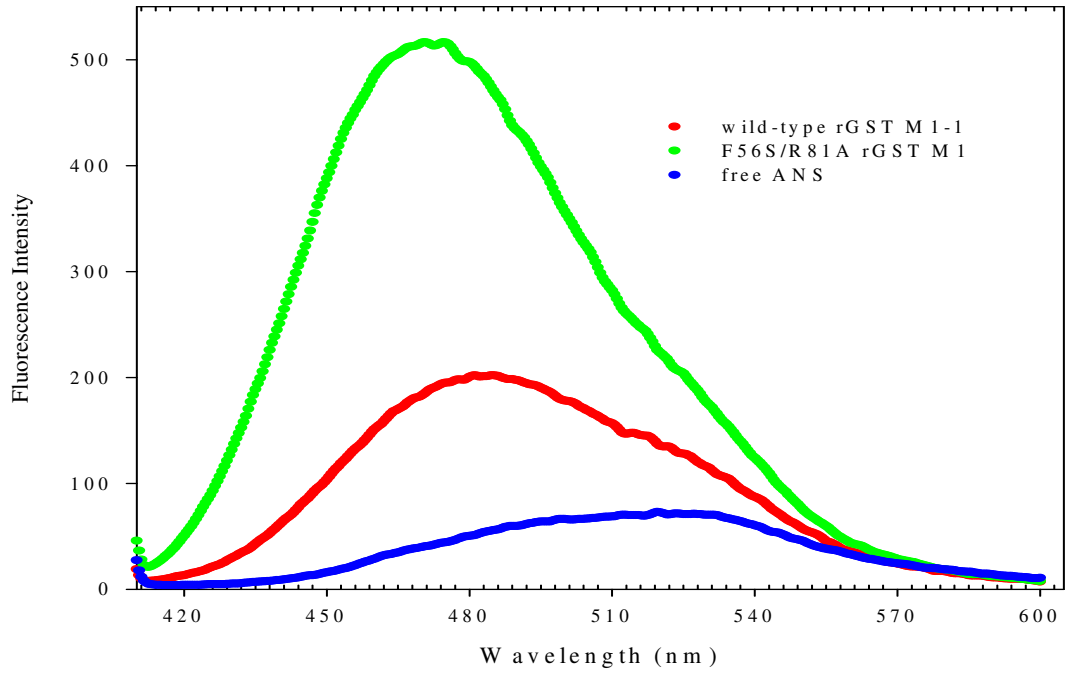


Figure 20. Fluorescence emission spectra of ANS bound to wild-type rGST M1-1 and the F56S/R81A rGST M1.

Experiments were performed by adding 200 μM ANS in 5 mM sodium phosphate buffer, pH 6.5, to 2 μM wild-type, and 4 μM F56S/R81A rGST M1 protein concentrations. ANS was selectively excited at 400 nm and the emission spectra were measured in the wavelength range 400-600 nm. The wavelengths of maximum emission intensity of the wild-type and F56S/R81A rGST M1 are 485 nm and 475 nm, respectively.

The lock-and-key inter-subunit motif in class Mu GST alone also impacts on the binding of ANS (Hornby *et al.*, 2002). Similar to the F56S/R81A rGST M1, the polarity and degree of solvent exposure of the ANS binding site in the F56S and F56R rGST M1-1 proteins was reduced relative to the wild-type protein (Hornby *et al.*, 2002). Disruption at the dimer interface of GSTs induces conformational changes at/near the H-site site. These results suggest that ANS non-specifically binds the hydrophobic regions exposed at the dimer interface of class Mu GSTs if subunit association is prevented. The dimer interface of class Mu GST is therefore likely to be substantially hydrophobic.

3.5. Thermal inactivation of the wild-type rGST M1-1 and F56S/R81A rGST M1-1

The thermal denaturation of wild-type rGST M1-1 and F56S/R81A rGST M1 was investigated between 20 °C - 70 °C and is shown in Figure 21. The wild-type is observed to have a loss of α-helical content with a midpoint (T_m) of 58°C compared to 52°C for F56S/R81A rGST M1. Wild-type rGST M1-1 has a higher thermal stability than F56S/R81A rGST M1. This suggests that the overall stability of F56S/R81A rGST M1 has been compromised. The lower T_m value seen for F56S/R81A rGST M1 coincides with lower $\Delta G(H_20)$ values (see section 3.6.3). Unlike solvent-induced denaturation, thermal denaturation is irreversible due to protein aggregation at high temperatures.

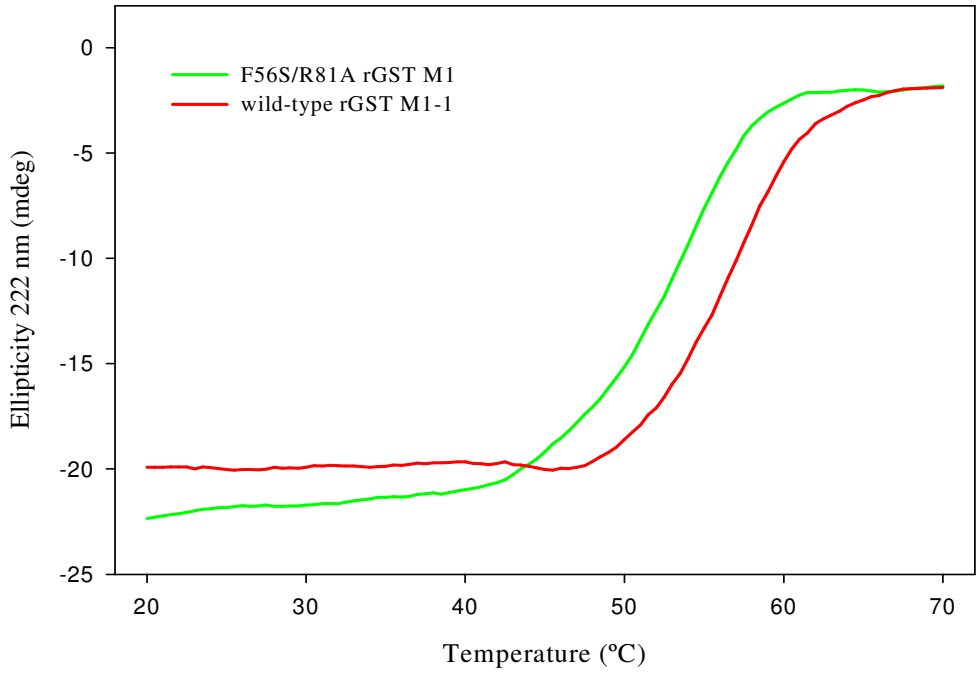


Figure 21. Thermal denaturation curves of F56S/R81A rGST M1 and wild-type rGST M1-1 from 20 - 70°C monitored by circular dichroism spectroscopy at 222 nm.

Spectral analyses were performed using 2 μM native folded wild-type and 4 μM F56S/R81A rGST M1 in 5 mM sodium phosphate buffer, pH 6.5, containing 0.02 % NaN_3. Each thermal denaturation profile represents an average of 20 accumulations.

3.6. Urea-induced equilibrium unfolding of the wild-type rGST M1-1 and F56S/R81A rGST M1

3.6.1. Techniques monitoring unfolding transitions

Ideally, multiple probes should be used to monitor the unfolding of proteins to enhance detection of any intermediates in the unfolding process. An accurate examination of the equilibrium unfolding process of the protein can thus be determined (Neet and Timm, 1994).

Far-UV CD, tryptophan fluorescence and ANS binding have been used as probes to monitor the unfolding process of GSTs (Erhardt and Dirr, 1995; Hornby *et al.*, 2000; Kaplan *et al.*, 1997; Stevens *et al.*, 1998; Wallace *et al.*, 1998b). Accordingly, these spectroscopic probes have been employed to assess the unfolding process and subsequently the conformational stability of the F56S/R81A rGST M1.

3.6.2. Reversibility of unfolding

In order to derive thermodynamic parameters for the protein unfolding processes, the reversibility of unfolding needs to be established (Pace, 1986). Tryptophan fluorescence was used to monitor the reversibility of the urea-induced equilibrium unfolding process for the wild-type rGST M1-1 and F56S/R81A rGST M1.

Illustrated in Figure 22 are the native and refolded fluorescence spectra for wild-type rGST M1-1 and F56S/R81A rGST M1. Following a ten-fold dilution of unfolded wild-type rGST M1-1 in 8 M urea to 0.8 M urea, wild-type rGST M1-1 was observed to regain around 98 % of its native structure (Figure 22).

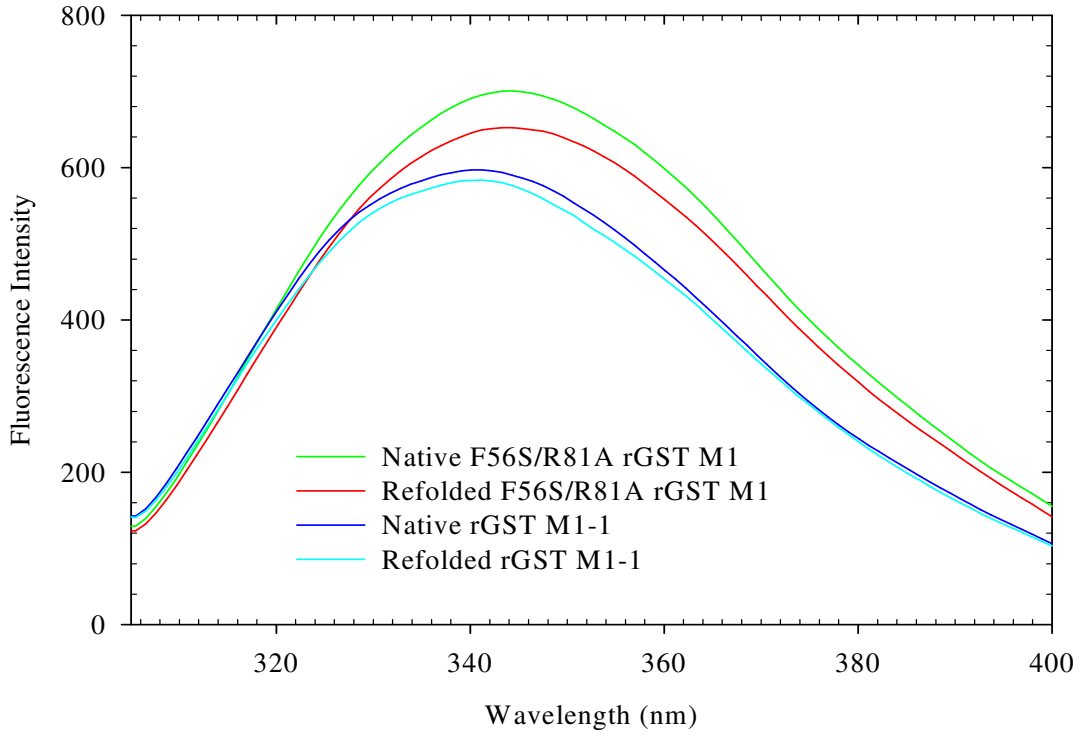


Figure 22. Fluorescence emission spectra of native and refolded wild-type rGST M1-1 and F56S/R81A rGST M1.

Fluorescence emission spectra of 2 μM wild-type rGST M1-1 and 4 μM F56S/R81A rGST M1 in 5mM sodium phosphate buffer, pH 6.5, excited at 295 nm. The wavelengths of maximum emission intensity of native and refolded rGST M1-1 and F56S/R81A rGST M1 are 340 and 345 nm, respectively.

This is indicative that the unfolded protein can refold and regain the same structure as that of the native protein in 0.8 M urea. In a similar experiment, the F56S/R81A rGST M1 regained approximately 90 % of its native structure (Figure 22). Aggregation was detected during the unfolding transition (see section 3.6.3). The diminished ability of the F56S/R81A rGST M1 to regain its native structure can possibly be due to aggregation of the protein in the unfolded state. Similar recovery data has been seen for class Alpha (Wallace *et al.*, 1998b) and class Pi (Erhardt and Dirr, 1995).

Aggregation is difficult to predict or prevent, and according to Booth *et al.* (1997), any protein can form aggregates under the appropriate conditions. The mechanism of protein aggregation is poorly understood. The most accepted explanation is that partially folded intermediates are formed when they have some hydrophobic surface area exposed (Fink, 1998; London *et al.*, 1974; Speed *et al.*, 1995). These exposed hydrophobic regions can then promote association and subsequent aggregation. These aggregates posses either an ordered morphology (amyloid fibrils) or are amorphous (inclusion bodies). The principle of this hypothesis is that there is an aggregation-prone intermediate formed, where this intermediate is the precursor for the association process. Considering this, it is possible that due to the prevention of dimerisation, conformational changes in the F56S/R81A rGSTM1 native state may have increased exposure of hydrophobic surface area, ultimately promoting non-specific association and subsequent aggregation. The ANS binding studies discussed in section 3.4.4 provide support for this.

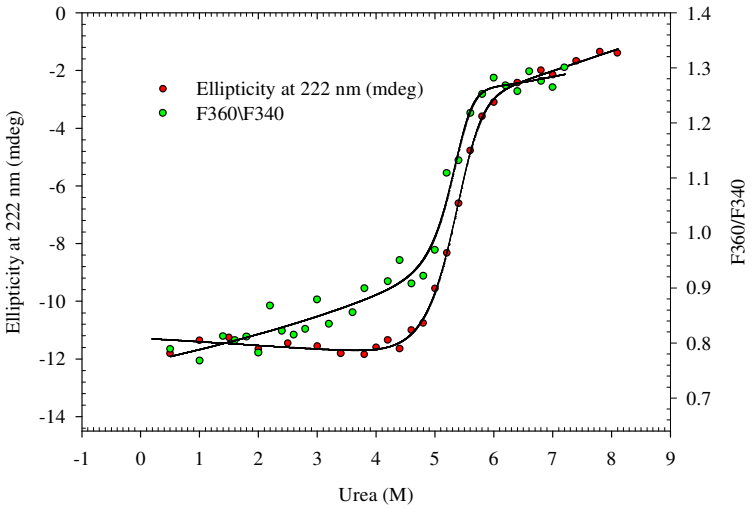
3.6.3. Urea induced equilibrium unfolding transitions

Illustrated in Figure 23 are the urea-induced equilibrium unfolding curves for (A) wild-type rGST M1-1 and (B) F56S/R81A rGST M1 monitored by tryptophan fluorescence and ellipticity. Illustrated in Figure 24 are the urea-induced equilibrium unfolding curves for wild-type rGST M1-1 and F56S/R81A rGST M1 monitored by (A) Tryptophan fluorescence (fluorescence emission at 340 nm for wild-type rGST

M1-1 and 345 nm for F56S/R81A rGST M1 which is characteristic of the folded state), (B) ellipticity, and (C) ANS binding.

The 222 nm CD unfolding transition curves in Figure 23 and 24B for F56S/R81A rGST M1 and wild-type rGST M1-1 are monophasic and sigmoidal, suggestive of a two-state transition. The CD transitions for the F56S/R81A rGST M1 and wild-type rGST M1-1 do not overlay. This non-coincidence implies dissimilar conformational stability between monomeric F56S/R81A rGST M1 and dimeric wild-type rGST M1-1. This sigmoidal transition corresponds to the second fluorescence transition characterising the unfolding of the monomeric intermediate of wild-type rGST M1-1, and the stable monomeric F56S/R81A rGST M1 (I↔U) in the F56S/R81A rGST M1. Comparison of C_m values revealed a shift towards lower urea concentrations of the unfolding transition mid-point for F56S/R81A rGST M1 relative to wild-type rGST M1-1. This data, represented in Table 2, indicates a diminished stability of F56S/R81A rGST M1 relative to wild-type rGST M1-1 during urea denaturation.

A.



B.

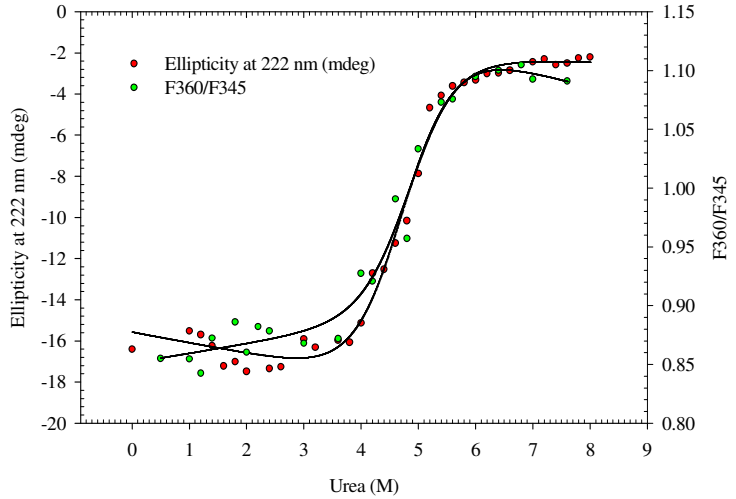


Figure 23. Urea-induced equilibrium unfolding of 4 μM F56S/R81A rGST M1 and 2 μM wild-type rGST M1-1.

Transitions were monitored by (●) tryptophan fluorescence and (●) far-UV circular dichroism. Experiments were performed in 20 mM sodium phosphate buffer, pH 6.5. Tryptophan residues were selectively excited at 295 nm. For each urea concentration the ratio of the emission intensity at 360 nm to 345 nm for the F56S/R81A rGST M1 and 340 nm for the wild-type rGST M1-1, is plotted. The unfolding curves for **(A)** wild-type rGST M1-1 were fitted according to a three-state model for dimeric proteins whereby the native dimer dissociates into a structured monomeric intermediate ($N_2 \leftrightarrow 2I \leftrightarrow 2U$). Unfolding curves for **(B)** F56S/R81A rGST M1 were fitted according to a two-state model for monomeric proteins ($I \leftrightarrow U$) using Sigma Plot v8.0, with a non-linear regression analysis ($R^2 = 0.99$ for CD data; $R^2 = 0.98$ for fluorescence data).

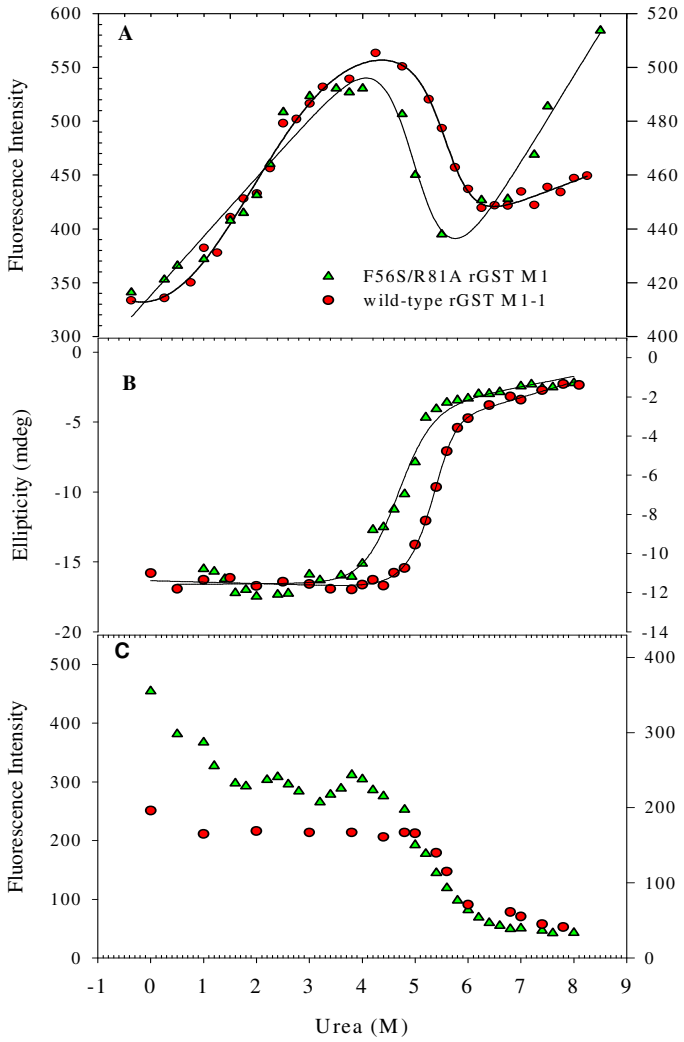


Figure 24. Urea-induced equilibrium unfolding of 4 μM F56S/R81A rGST M1 and 2 μM wild-type rGST M1-1.

Transitions were monitored by (**A**) tryptophan fluorescence (**B**) far-UV circular dichroism and (**C**) ANS Binding. Experiments were performed in 20 mM sodium phosphate buffer, pH 6.5. Tryptophan residues were selectively excited at 295 nm. The unfolding curves for wild-type rGST M1-1 were fitted according to a three-state model for dimeric proteins whereby the native dimer dissociates into a structured monomeric intermediate ($N_2 \leftrightarrow 2I \leftrightarrow 2U$). Unfolding curves for F56S/R81A rGST M1 were fitted according to a two-state model for monomeric proteins ($I \leftrightarrow U$). The data was fitted using Sigma Plot v8.0, with a non-linear regression analysis (R^2 = 0.99 for CD data; R^2 = 0.98 for fluorescence data).

Urea-induced equilibrium unfolding curves for wild-type rGST M1-1 and F56S/R81A rGST M1 monitored by tryptophan fluorescence are illustrated in Figure 25. For each urea concentration, the ratio of the emission intensity at 360 nm to 345 nm for the F56S/R81A rGST M1 and 360 nm to 340 nm for the wild-type rGST M1-1 is plotted. The far-UV CD and fluorescence unfolding events measured for wild-type rGST M1-1 and F56S/R81A rGST M1 do not overlay. Changes in the unfolding transition mid-points (C_m value), as well as differences in the slope of the transition regions, shape the non-superimposable unfolding curves of F56S/R81A rGST M1 and wild-type rGST M1-1.

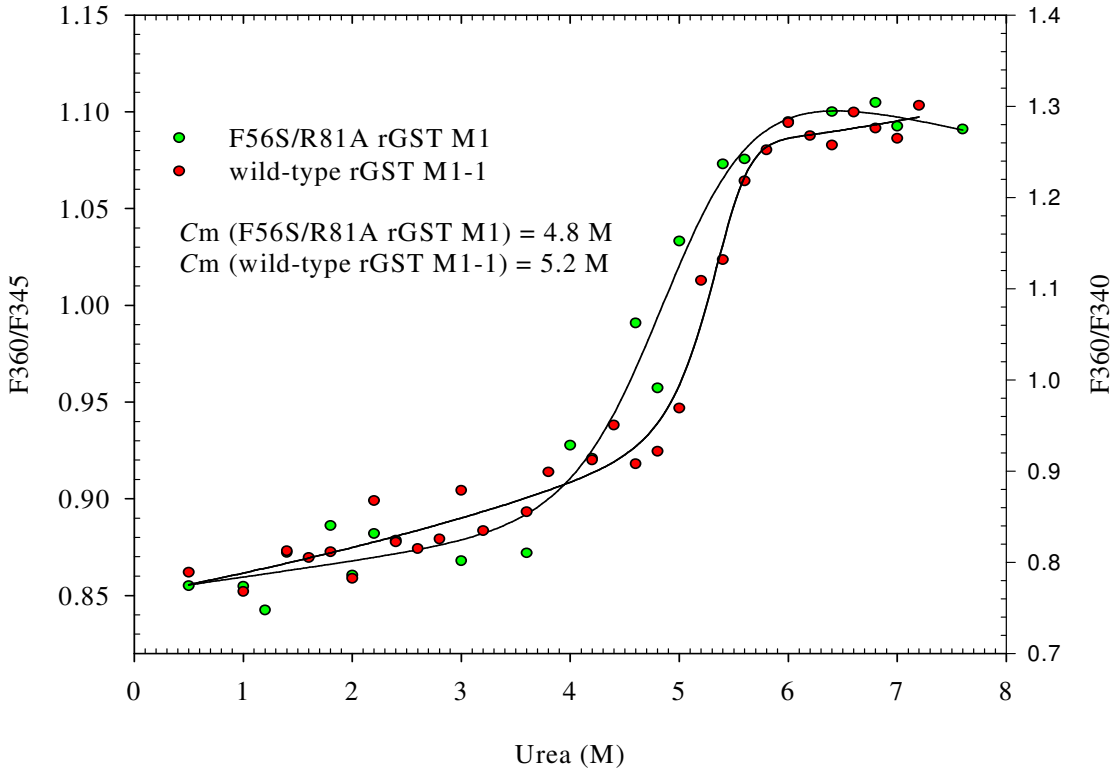


Figure 25. Urea-induced unfolding of 4 μM F56S/R81A rGST M1 and 2 μM wild-type rGST M1-1 monitored by tryptophan fluorescence.
Experiments were performed in 20 mM sodium phosphate buffer, pH 6.5. Tryptophan residues were selectively excited at 295 nm. For each urea concentration the ratio of the emission intensity at 360 nm to 345 nm for the F56S/R81A rGST M1 and 340 nm for the wild-type rGST M1-1 is plotted. Unfolding curves for wild-type rGST M1-1 were fitted according to a three-state model for dimeric proteins whereby the native dimer dissociates into a structured monomeric intermediate ($N_2 \leftrightarrow 2I \leftrightarrow 2U$). Unfolding curves for F56S/R81A rGST M1 were fitted according to a two-state model for monomeric proteins ($I \leftrightarrow U$). The data was fitted using Sigma Plot v8.0, with a non-linear regression analysis (R^2 = 0.99 for CD data; R^2 = 0.98 for fluorescence data).

Table 2: Thermodynamic parameters obtained from urea equilibrium denaturation of wild-type rGST M1-1 and F56S/R81A rGST M1 monitored by fluorescence and circular dichroism.

rGST M1-1	Model	$\Delta G(H_20)$ (kcal/mol)	*m*-value (kcal mol^{-1} M^{-1})	C_m (M)
wild-type rGST M1-1	$N_2 \leftrightarrow 2I$	8.21 ± 2.2	0.75 ± 0.72	2.4
	$I \leftrightarrow U$	13.59 ± 3.5	2.57 ± 0.64	5.2
		Total 21.8	Total 3.32	
F56S/R81A rGST M1	$I \leftrightarrow U$	6.7 ± 2.5	2.11 ± 0.5	4.8

The tryptophan fluorescence unfolding transition curves of wild-type rGST M1-1 are biphasic and suggestive of a three-state transition which corresponds to previous published data (Hornby *et al.*, 2002). This is similar to the reported mechanism for certain Pi class enzymes (Aceto *et al.*, 1992; Sacchetta *et al.*, 1999), but in contrast to the two-state transition for several other classes (Erhardt and Dirr, 1995; Kaplan *et al.*, 1997; Wallace *et al.*, 1998b) and the four-state mechanism of the Sigma class (Stevens *et al.*, 1998). The wild-type rGST M1-1 display relatively steep slopes in their unfolding transitions, while F56S/R81A rGST M1 displays an extremely steep unfolding post-transition. The pre-transition of F56S/R81A rGST M1 between 0-4 M urea shows tryptophan fluorescence emission is enhanced (hyperfluorescence). The presence of hyperfluorescence during the pre-transitional unfolding process has been explained by Ervin *et al.* (2002) as a pre-transition conformationally loosened state; however, the contribution of general solvent effects cannot be ruled out when examining the pre-transition of F56S/R81A rGST M1.

Changes in the unfolding transition mid-points (C_m value), as well as differences in the slopes of the pre and post transition regions, are observed for F56S/R81A rGST M1 relative to wild-type protein. The fluorescence transition of F56S/R81A rGST M1 corresponds to the dimer disassociation event seen at low urea concentration (C_m

2.4M urea, transition absent in F56S/R81A rGST M1). Both wild-type rGST M1-1 and F56S/R81A rGST M1 show similar unfolding trends but do not overlay each other (see Figure 24, 25). Non-simultaneous loss of secondary and tertiary structure, as well as loss of packing about the tryptophans, can be assumed in both wild-type rGST M1-1 and F56S/R81A rGST M1, as the far-UV CD and fluorescence unfolding events measured for each protein are non-coincident when unfolding in the presence of urea (see Figure 23).

Relative to wild-type protein, the unfolding transitions observed for F56S/R81A rGST M1 are slightly broader and indicate a less cooperative unfolding process. This data suggests a decrease in the dependence of free energy change of unfolding upon denaturant concentration (i.e. a decreased *m*-value; (Pace, 1990). For further discussion see section 3.6.2.

Figure 26 illustrates Rayleigh scattering due to aggregation of F56S/R81A rGST M1. This Rayleigh scattering is observed to increase with increasing protein concentration in the urea concentration range corresponding to that of the unfolding transition regions (~ 4.0-6.0 M urea). If aggregates are present in the unfolding process, the system is no longer considered to be in equilibrium (i.e., the process is no longer considered to be reversible). Thus, in order to determine the thermodynamic parameters of F56S/R81A rGST M1, the effects associated with aggregation need to be excluded. Urea-induced equilibrium unfolding experiments were performed using 4 μM protein concentrations, the lowest possible concentration for accurate comparison to wild-type rGST M1-1. Aggregate formation was minimised, but not abolished, for F56S/R81A rGST M1 at this protein concentration.

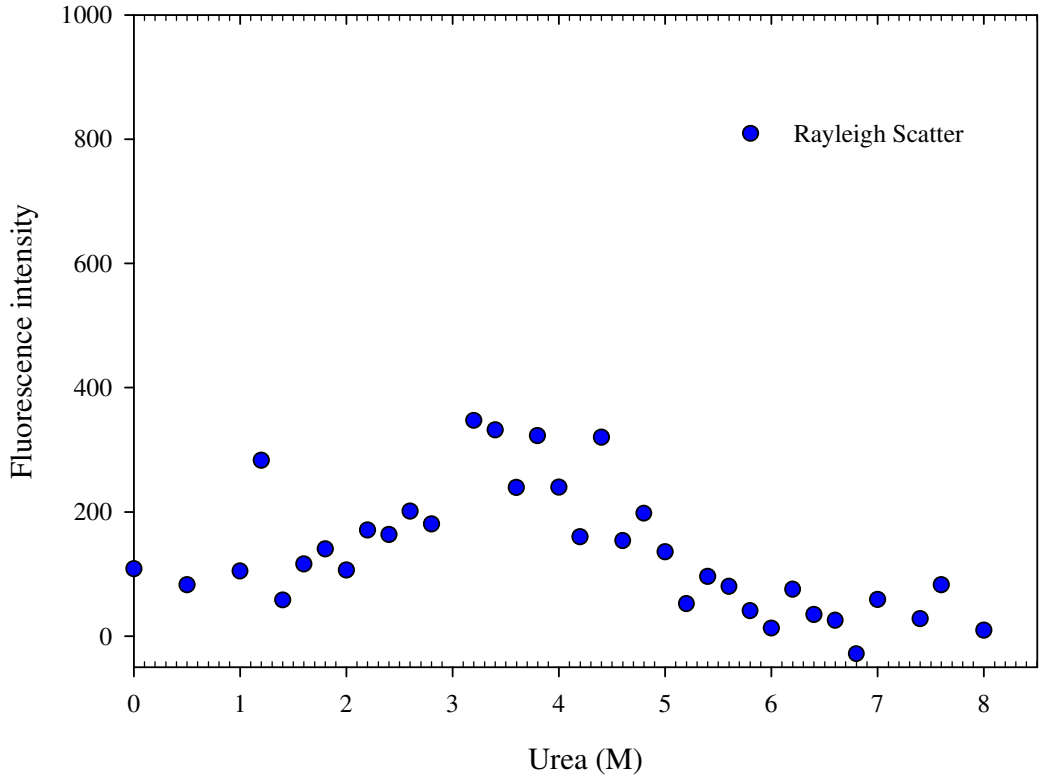


Figure 26. Rayleigh scattering of 4 μM F56S/R81A rGST M1, at different concentrations of urea.

Rayleigh scattering was measured by setting excitation and emission wavelengths to 295 nm. Experiments were performed in 20 mM sodium phosphate buffer, pH 6.5.

3.6.4. Analysis of the equilibrium unfolding transitions according to a two-state unfolding pathway (I ↔ U)

The differences observed in the urea-induced equilibrium unfolding process of wild-type rGST M1-1 and F56S/R81A rGST M1 suggest dissimilar thermodynamic stabilities of each protein. In order to determine the thermodynamic parameters of unfolding for the wild-type rGST M1-1 and F56S/R81A rGST M1, a satisfactory model describing the urea-induced equilibrium unfolding of these proteins has to be employed.

Urea-induced equilibrium unfolding curves of F56S/R81A rGST M1 monitored by 222 nm CD are characterised by single sigmoidal transitions, suggesting the absence of any thermodynamically stable intermediates. Accordingly, F56S/R81A rGST M1 unfolding curves were fitted to a two-state model. This model assumes the presence of only two species of protein, the native (I) and unfolded (U) forms, within the transition region (Pace, 1986). Urea-induced equilibrium unfolding curves for F56S/R81A rGST M1 were analysed according to the linear extrapolation method of Pace (1986). In order to obtain values for $\Delta G(H_2O)$ and the *m*-value, linear dependence of Gibbs free energy change of unfolding on denaturant concentration (Pace, 1986) was assumed.

The thermodynamic parameters defining the unfolding transitions of each variant protein are shown in Table 2. The R^2 values and dependency values determined for the fitted curves, together with the standard errors derived from the thermodynamic parameters, were satisfactory. Thus, the fit to a two-state model was acceptable.

Protein conformational stability can be measured as Gibbs free energy change in the absence of denaturant $\Delta G(H_2O)$. The $\Delta G(H_2O)$ values for F56S/R81A rGST M1 determined using fluorescence (6.7 ± 1.9 kcal/mol) and far-UV CD (6.31 ± 0.85 kcal/mol) are similar to each other, but differ significantly to the previously reported

wild-type rGST M1-1 value (19.60 ± 0.21 kcal/mol;(Hornby *et al.*, 2000) and to the experimentally determined value (25.19 ± 1.15 kcal/mol). The diminished Gibbs free energy of unfolding is indicative of F56S/R81A rGST M1 being less stable than wild-type rGST M1-1. The data in Table 2 suggests that disruption of the mixed-charge cluster, and the lock-and-key motif, is unfavourable to protein stability, and that dimerisation contributes to individual subunit as well as overall protein stability. Consequently, the prevention of dimerisation in rGST M1-1 removes other inter-subunit interactions which are possibly required for subunit and dimer stabilisation.

The *m*-value of a denaturation curve is related to the increase in solvent-accessible surface area due to unfolding (Myers *et al.*, 1995). The theoretical *m*-value of a protein can therefore be estimated by relating the amount of change in buried surface area to the number of amino acid residues within that protein (Myers *et al.*, 1995). For urea as denaturant:

$$m = 374 + 0.11\ (\Delta ASA) \qquad (16)$$

Using equation 16, the predicted *m*-values are 4.7 and 2.5 $kcal.mol^{-1}M^{-1}$ for the wild-type rGST M1-1 and F56S/R81A rGST M1, respectively. This theoretical value is greater than the previously published *m*-value of 3.3 $kcal.mol^{-1}M^{-1}$ (Hornby *et al.*, 2000) and experimentally determined *m*-value of 3.39 $kcal.mol^{-1}M^{-1}$ for wild-type rGST M1-1. The *m*-values obtained for F56S/R81A rGST M1 using fluorescence (2.11 ± 0.5) and far-UV CD (1.3 ± 0.2) data were similar to each other. Possible reasons for reduced *m*-values are either differences in the exposed surface area of the protein in the native and/or the unfolded state (Myers *et al.*, 1995; Shortle, 1995), or changes in the interactions of denaturant with the protein molecules in the denatured state (Arakawa and Timasheff, 1984). Urea is known to induce complete unfolding of GSTs (Wallace *et al.*, 1998b), however, which rules out the possibility that the F56S/R81A rGST M1 reduced *m*-value is a result of changes in the interactions of the denaturant with the protein molecules in the denatured state (Arakawa and Timasheff, 1984). The most likely explanation is that less surface area has been exposed upon

unfolding of F56S/R81A rGST M1, however an alternative reason could be if the protein was incorrectly folded initially.

Decreased *m*-values may result from differences in exposed surface area of the native and denatured state. A decrease in sensitivity of the equilibrium between folded dimer and unfolded monomer to denaturant would also be reflected by a decreased *m*-value (Shortle, 1995). In section 3.4.4, ANS binding studies indicated enhanced exposure of hydrophobic surface area in F56S/R81A rGST M1. However, enhancement is localised due to exposure of the dimer interface rather than a global effect. Decreased *m*-values may result from diminished cooperativity of unfolding transitions for F56S/R81A rGST M1. Nonetheless, the *m*-value obtained for F56S/R81A rGST M1 indicates a highly cooperative folding pathway. Cooperativity of protein folding arises when hydrophobic and hydrophilic copolymer sequences collapse to states that are compact and also have good hydrophobic cores (Dill *et al.*, 1995). The hydrophobic nature of the dimer interface of rGST M1-1 replicates a protein hydrophobic core. Disruption of the dimer interface residues and prevention of quaternary interaction implementation alter the nature of this hydrophobic core, hence lowering the cooperativity of the unfolding process of the F56S/R81A rGST M1 slightly. The folding of F56S/R81A rGST M1 is highly cooperative although slightly less than wild-type rGST M1-1. Decreased cooperativity could suggest the presence of additional state(s) of intermediate structure and stability at equilibrium (Wallace *et al.*, 2000). The binding of the amphipathic ligand, ANS, is frequently used to probe for the presence of intermediates during protein unfolding events. Figures 24C illustrates the effect of urea on the binding of ANS to wild-type rGST M1-1 and F56S/R81A rGST M1. The unfolding processes of wild-type rGST M1-1 and F56S/R81A rGST M1 in the presence of ANS are similar to each other, and follow a similar trend to that seen with tryptophan fluorescence. Thus, no intermediate in F56S/R81A rGST M1 folding pathway was detected from ANS unfolding studies. Consequently, the possibility of F56S/R81A rGST M1 following an alternate unfolding pathway is unlikely.

GSTs are obligate dimeric proteins and a major question is the reliance of stability on the quaternary structure. It is well known that interactions at the subunit interface play an important role in stabilising the subunit tertiary structure as GSTs are obligate functional dimers (for review see Dirr, 2001). The extent of subunit stabilization is class dependent. Classes Mu and Sigma form molten globule-like monomeric intermediates during unfolding whilst the classes Pi, Alpha and Sj26 GST unfold with a concerted dimer dissociation-unfolding process. The unfolding process of class Pi GST is unclear in terms of whether it unfolds via a two-state (Dirr and Reinemer, 1991; Erhardt and Dirr, 1995) or three-state (Aceto *et al.*, 1992) unfolding process. Wallace and Dirr (unpublished) have addressed this issue and found that class Pi, in fact, does unfold via a three-state unfolding process where the intermediate is dimeric and that intermediate formation is due to the highly flexible/mobile α-helix 2. It is clear that the reliance on quaternary structure for subunit stabilization is different for the different classes of GST. There is no clear trend regarding the hydropathy of the dimer interface, as GSTs with hydrophobic (e.g. GST M1-1) and hydrophilic dimer interfaces (e.g. GST S1-1) form monomeric intermediates during the unfolding process. These results reflect the diverse range of stabilities of these enzymes, for both monomer and dimer – a diversity resulting, in part, from the differences in inter-subunit interactions.

Conclusion

The aim of this investigation was to determine the role of the mixed-charge cluster at the dimer interface of rGST M1-1 in dimerisation. From the results discussed, it is clear that the mixed-charge cluster at the dimer interface of rGST M1-1 is essential for subunit association and contributes significantly to protein stability and catalytic function. The monomeric F56S/R81A rGST M1 seemed structurally similar to the wild-type however domain 2 seems to be more loosely packed. ANS binding studies indicated a large increase in the accessible hydrophobic surface area of the monomer. The urea-induced conformational stability studies indicated destabilisation of F56S/R81A rGST M1 as a result of the obliteration of any quaternary interaction, resulting in an increase in aggregate structures. Subunit association is suggested to not be a highly cooperative reaction suggested by the creation of stable monomers in F56S/R81A rGST M1. Dimerisation of rGST M1-1 is critical for the stabilisation and functionality of the enzyme.

References

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

Aceto, A., Sacchetta, P., Bucciarelli, T., Dragani, B., Angelucci, S., Radatti, G.L., and Di Ilio, C. (1995) Structural and functional properties of the 34-kDa fragment produced by the N-terminal chymotryptic cleavage of glutathione transferase P1-1. *Arch. Biochem. Biophys.* **316**, 873-878.

Aceto, A., Caccuri, A.M., Sacchetta, P., Bucciarelli, T., Dragani, B., Rosato, N., Federici, G., and Di Ilio, C. (1992) Dissociation and unfolding of Pi-class glutathione transferase. Evidence for a monomeric inactive intermediate. *Biochem. J.* **285**, 241-245.

Adang, A.E., Brussee, J., Van der, G.A., and Mulder, G.J. (1990) The glutathione-binding site in glutathione S-transferases. Investigation of the cysteinyl, glycyl and gamma-glutamyl domains. *Biochem. J.* **269**, 47-54.

Adang, A.E., Meyer, D.J., Brussee, J., Van der, G.A., Ketterer, B., and Mulder, G.J. (1989) Interaction of rat glutathione S-transferases 7-7 and 8-8 with gamma-glutamyl- or glycyl-modified glutathione analogues. *Biochem. J.* **264**, 759-764.

Agianian, B., Tucker, P.A., Schouten, A., Leonard, K., Bullard, B., and Gros, P. (2003) Structure of a Drosophila sigma class glutathione S-transferase reveals a novel active site topography suited for lipid peroxidation products. *J. Mol. Biol.* **326**, 151-165.

Arakawa, T. and Timasheff, S.N. (1984) Protein stabilization and destabilization by guanidinium salts. *Biochemistry* **23**, 5924-5929.

Argos, P. (1988) An investigation of protein subunit and domain interfaces. *Protein Eng* **2**, 101-113.

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

Bahadur, R.P., Chakrabarti, P., Rodier, F., and Janin, J. (2003) Dissecting subunit interfaces in homodimeric proteins. *Proteins* **53**, 708-719.

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

Bico, P., Chen, C.Y., Jones, M., Erhardt, J., and Dirr, H. (1994) Class pi glutathione S-transferase: Meisenheimer complex formation. *Biochem. Mol. Biol. Int.* **33**, 887-892.

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

Board, P.G., Coggan, M., Chelvanayagam, G., Easteal, S., Jermiin, L.S., Schulte, G.K., Danely, 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 transferase. *J. Biol. Chem*, 24798-24806.

Bogan, A.A. and Thorn, K.S. (1998) Anatomy of hot spots in protein interfaces. *J. Mol. Biol.* **280**, 1-9.

Booth, D.R., Sunde, M., Bellotti, V., Robinson, C.V., Hutchinson, W.L., Fraser, P.E., Hawkins, P.N., Dobson, C.M., Radford, S.E., Blake, C.C., and Pepys, M.B. (1997) Instability, unfolding and aggregation of human lysozyme variants underlying amyloid fibrillogenesis. *Nature* **385**, 787-793.

Bowie, J.U. and Sauer, R.T. (1989) Equilibrium dissociation and unfolding of the Arc repressor dimer. *Biochemistry* **28**, 7139-7143.

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

Brinda, K.V., Kannan, N., and Vishveshwara, S. (2002) Analysis of homodimeric protein interfaces by graph-spectral methods. *Protein Eng* **15**, 265-277.

Chothia, C. and Janin, J. (1975) Principles of protein-protein recognition. *Nature* **256**, 705-708.

Codreanu, S.G., Thompson, L.C., Hachey, D.L., Dirr, H.W., and Armstrong, R.N. (2005) Influence of the dimer interface on glutathione transferase structure and dynamics revealed by amide H/D exchange mass spectrometry. *Biochemistry* **44**, 10605-10612.

DeLano, W.L. (2002) The PyMOL Molecular Graphics System, DeLano Scientific, San Carlos, USA.

Dill, K.A. (1990) Dominant forces in protein folding. *Biochemistry* **29**, 7133-7155.

Dill, K.A., Bromberg, S., Yue, K., Fiebig, K.M., Yee, D.P., Thomas, P.D., and Chan, H.S. (1995) Principles of protein folding--a perspective from simple exact models. *Protein Sci.* **4**, 561-602.

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

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

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.

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

Duquerroy, S., Cherfils, J., and Janin, J. (1991) Protein-protein interaction: an analysis by computer simulation. *Ciba Found. Symp.* **161**, 237-249.

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.

Ervin, J., Larios, E., Osvath, S., Schulten, K. and Gruebele, M. (2002) What causes hyperfluorescence: folding intermediates or conformationally flexible native states? *Biophys J*. **83**, 473-83

Fersht, A.R. (1984) Basis of biological specificity. *Trends Biochem. Sci.* **9**, 145-147.

Fink, A.L. (1998) Protein aggregation: folding aggregates, inclusion bodies and amyloid. *Fold. Des* **3**, R9-23.

Goodsell, D.S. and Olson, A.J. (1993) Soluble proteins: size, shape and function. *Trends Biochem. Sci.* **18**, 65-68.

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.

Gulick, A.M., Fahl, W.E., and Goihl, A.L. (1992) Structural studies on human glutathione S-transferase pi. Family of native-specific monoclonal antibodies used to block catalysis. *J. Biol. Chem* **267**, 18946-18952.

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

Hayes, J.D. and McLellan, L.I. (1999) Glutathione and glutathione-dependent enzymes represent a co-ordinately regulated defence against oxidative stress. *Free Radic. Res* **31**, 273-300.

Hayes, J.D. and Strange, R.C. (2000) Glutathione S-transferase polymorphisms and their biological consequences. *Pharmacology* **61**, 154-166.

Hegazy, U.M., Mannervik, B., and Stenberg, G. (2004) Functional role of the lock-and-key motif at the subunit interface of glutathione transferase p1-1. *J. Biol. Chem.* **279**, 9586-9596.

Hennessey, J.P., Johnson, W.C.J., Bahler, C., and Wood, H.G. (1982) Subunit interactions of transcarboxylase as studied by circular dichroism. *Biochemistry* **21**, 642-646.

Heringa, J. and Argos, P. (1991) Side-chain clusters in protein structures and their role in protein folding. *J. Mol. Biol.* **220**, 151-171.

Hornby, J.A., Codreanu, S.G., Armstrong, R.N., and Dirr, H.W. (2002) Molecular recognition at the dimer interface of a class mu glutathione transferase: role of a hydrophobic interaction motif in dimer stability and protein function. *Biochemistry* **41**, 14238-14247.

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.

Hubbard, S.J. and Argos, P. (1994) Cavities and packing at protein interfaces. *Protein Sci.* **3**, 2194-2206.

Jakobsson, P.-J., Morgenstern, R., Mancini, J., Ford-Hutchinson, A., and Persson, B. (1994) Common structural features of MAPEG- a widespread superfamily of membrane associated proteins with highly divergent functions in eicosanoid and glutathione metabolism. *Protein Sci.* **8**, 689-692.

Janin, J. and Chothia, C. (1990) The structure of protein-protein recognition sites. *J. Biol. Chem.* **265**, 16027-16030.

Janin, J., Miller, S., and Chothia, C. (1988) Surface, subunit interfaces and interior of oligomeric proteins. *J. Mol. Biol.* **204**, 155-164.

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

Johnson, W.C.J. (1990) Protein secondary structure and circular dichroism: a practical guide. *Proteins* **7**, 205-214.

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

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

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

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.

Kirk, W.R., Kurian, E., and Prendergast, F.G. (1996) Characterization of the sources of protein-ligand affinity: 1-sulfonato-8-(1')anilinonaphthalene binding to intestinal fatty acid binding protein. *Biophys. J.* **70**, 69-83.

Kraulis, P.J. (1991) MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures. *J. App. Crystallogr.* **24**, 946-950.

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

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

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

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

Lawrence, M.C. and Colman, P.M. (1993) Shape complementarity at protein/protein interfaces. *J. Mol. Biol.* **234**, 946-950.

Lo, C.L., Chothia, C., and Janin, J. (1999) The atomic structure of protein-protein recognition sites. *J. Mol. Biol.* **285**, 2177-2198.

London, J., Skrzynia, C., and Goldberg, M.E. (1974) Renaturation of Escherichia coli tryptophanase after exposure to 8 M urea. Evidence for the existence of nucleation centers. *Eur. J. Biochem.* **47**, 409-415.

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

Mannervik, B. (2003) The isoenzymes of glutathione transferase. *Adv. Enzymol. Relat. Areas Mol. Biol* **57**, 357-417.

Martini, F., Aceto, A., Sacchetta, P., Bucciarelli, T., Dragani, B., and Di Ilio, C. (1993) Investigation of intra-domain and inter-domain interactions of glutathione transferase P1-1 by limited chymotryptic cleavage. *Eur. J. Biochem.* **218**, 845-851.

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.

McCammon, J.A. (1998) Theory of biomolecular recognition. *Curr. Opin. Struct. Biol.* **8**, 245-249.

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.

Meyer, D.J. and Thomas, M. (1995) Characterization of rat spleen prostaglandin-H D-isomerase as a sigma class GSH transferase. *Biochem J* **311**, 739-742.

Meyer, M., Wilson, P., and Schomburg, D. (1996) Hydrogen bonding and molecular surface shape complementarity as a basis for protein docking. *J. Mol. Biol.* **264**, 199-210.

Monod, J., Wyman, J., and Changeux, J.-P. (1965) On the nature of allosteric interactions: a plausible model. *J. Mol. Biol.* **12**, 88-118.

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.

Nayal, M., Hitz, B.C., and Honig, B. (1999) GRASS: a server for the graphical representation and analysis of structures. *Protein Sci.* **8**, 676-679.

Neet, K.E. and Timm, D.E. (1994) Conformational stability of dimeric proteins: quantitative studies by equilibrium denaturation. *Protein Sci.* **3**, 2167-2174.

Nooren, I.M. and Thornton, J.M. (2003) Diversity of protein-protein interactions. *EMBO J.* **22**, 3486-3492.

Ottosen, M.B., Bjornberg, O., Norager, S., Larsen, S., Palfey, B.A., and Jensen, K.F. (2002) The dimeric dihydroorotate dehydrogenase A from Lactococcus lactis dissociates reversibly into inactive monomers. *Protein Sci.* **11**, 2575-2583.

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

Pace, C.N. (1990) Conformational stability of globular proteins. *Trends Biochem. Sci.* **15**, 14-17.

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.

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

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

Perkins, S.J. (1986) Protein volumes and hydration effects. *Eur. J. Biochem.* **319**, 749-754.

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

Rossjohn, J., Polekhina, G., Feil, S.C., Allocati, N., Masulli, M., De Illio, C., and Parker, M.W. (1998) A mixed disulfide bond in bacterial glutathione transferase: functional and evolutionary implications. *Structure.* **6**, 721-734.

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

Sacchetta, P., Pennelli, A., Bucciarelli, T., Cornelio, L., Amicarelli, F., Miranda, M., and Di Ilio, C. (1999) Multiple unfolded states of glutathione transferase bbGSTP1-1 by guanidinium chloride. *Arch. Biochem. Biophys.* **369**, 100-106.

Sali, A. and Blundell, T.L. (1990) Definition of general topological equivalence in protein structures. A procedure involving comparison of properties and relationships through simulated annealing and dynamic programming. *J. Mol. Biol.* **212**, 403-428.

Sayed, Y., Hornby, J.A., Lopez, M., and Dirr, H. (2002) Thermodynamics of the ligandin function of human class Alpha glutathione transferase A1-1: energetics of organic anion ligand binding. *Biochem. J.* **363**, 341-346.

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.

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

Sheinerman, F.B., Norel, R., and Honig, B. (2000) Electrostatic aspects of protein-protein interactions. *Curr. Opin. Struct. Biol.* **10**, 153-159.

Shortle, D. (1995) Staphylococcal nuclease: a showcase of m-value effects. *Adv. Protein Chem.* **46**, 217-247.

Sluis-Cremer, N., Naidoo, N., Kaplan, W.H., Manoharan, T.H., Fahl, W.E. and Dirr, H.W. (1996) Determination of a binding site for non-substrate ligands in mammalian cytosolic glutathione S-transferases by means of fluorescence resonance energy transfer. *Eur. J. Biochem.* **241,** 484-488.

Speed, M.A., Wang, D.I., and King, J. (1995) Multimeric intermediates in the pathway to the aggregated inclusion body state for P22 tailspike polypeptide chains. *Protein Sci.* **4**, 900-908.

Stenberg, G., Abdalla, A.M., and Mannervik, B. (2000) Tyrosine 50 at the subunit interface of dimeric human glutathione transferase P1-1 is a structural key residue for modulating protein stability and catalytic function. *Biochem. Biophys. Res. Commun.* **271**, 59-63.

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.

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.

Tsai, C.J., Lin, S.L., Wolfson, H.J., and Nussinov, R. (1997) Studies of protein-protein interfaces: a statistical analysis of the hydrophobic effect. *Protein Sci.* **6**, 53-64.

Tsai, C.J. and Nussinov, R. (1997) Hydrophobic folding units at protein-protein interfaces: implications to protein folding and to protein-protein association. *Protein Sci.* **6**, 1426-1437.

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

Vuilleumier, S., Sancho, J., Loewenthal, R., and Fersht, A.R. (1993) Circular dichroism studies of barnase and its mutants: characterization of the contribution of aromatic side chains. *Biochemistry* **32**, 10303-10313.

Wallace, L.A., Blatch, G.L., and Dirr, H.W. (1998a) A topologically conserved aliphatic residue in alpha-helix 6 stabilizes 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., Sluis-Cremer, N., and Dirr, H.W. (1998b) Equilibrium and kinetic unfolding properties of dimeric human glutathione transferase A1-1. *Biochemistry* **37**, 5320-5328.

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

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

Xu, D., Tsai, C.J., and Nussinov, R. (1998) Mechanism and evolution of protein dimerization. *Protein Sci.* **7**, 533-544.

Zhang, P.H. and Armstrong, R.N. (1990) Construction, expression, and preliminary characterization of chimeric class mu glutathione S-transferases with altered catalytic properties. *Biopolymers* **29**, 159-169.

Zhu, Z.Y. and Karlin, S. (1996) Clusters of charged residues in protein three-dimensional structures. *Proc. Natl. Acad. Sci. U. S. A* **93**, 8350-8355.