

https://dspace.org/

The stability of monomeric class Mu GST M1

Mashapa, Tshepo

2013-02-07
http://hdl.handle.net/10539/12407

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

THE STABILITY OF MONOMERIC CLASS Mu GST M1

TSHEPO MASHAPA

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, October 2011

DECLARATION

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



_____

Tshepo Mashapa

_____ day of_____, 2011

ABSTRACT

Cytosolic glutathione transferases (GSTs) are obligate homo- or heterodimers with a complex canonical structure whose assembly involves interactions at the domain interface and at the subunit interface. Each subunit is composed of an N-terminal GSH binding domain (domain 1) and a C-terminal all-α-helical domain (domain 2). Currently, little is known about the stability, folding and assembly of GST subunits in the absence of quaternary interactions. The stabilities of several GST dimers, including class Mu GST, are dependent upon two groups of interactions between domains 1 and 2 of opposing subunits, namely, a hydrophobic lock-and-key motif and a buried charge cluster motif. In class Mu GST M1-1 from rat, the mutation of the two key residues , Phe-56 and Arg-81, involved in subunit-subunit interaction motifs, into Ser and Ala, respectively, has been reported to result in a stable monomer (Thompson *et al.* (2006) *Biochemistry* 45, 2267-2273). In this study, using gel filtration chromatography, it was confirmed that simultaneous disruption of both motifs in rGST M1-1 to create the F56S/R81A mutant results in complete dissociation of the dimer to a monomeric protein. Comparative studies reveal the monomeric mutant to be catalytically inactive due to the loss of the subunit-subunit contacts, displaying approximately 2.7 % activity relative to the wild type enzyme. Moreover, the stability and folding of a monomeric form of rGST M1-1 were investigated under equilibrium conditions by circular dichroism, intrinsic tryptophan fluorescence, and ANS binding. The results show that the monomer is less stable, with a $\Delta G(H_2O)$ value of 13.3 kcal/mol relative to the reported $\Delta G(H_2O)$ value of 16.5 kcal/mol (Hornby *et al.* (2000) *Biochemistry* 39, 12336-12344) for the second transition (i.e. I⇔U) of the dimeric wild type enzyme.

This work is dedicated to my late grandmother Kehumile "Zepps" Kebawetse, who taught me that there is nothing that is beyond my reach, so long as I want it bad enough.

"At every single moment of one's life one is what one is going to be and no less than what one has been."

Oscar Wilde

ACKNOWLEDGMENTS

I wish to extend my gratitude and thanks to the following people:

My supervisor, Prof. Heini Dirr, for his patience, guidance and support throughout the course of my studies. Thank you for believing in me and giving me the opportunity of working in your laboratory.

My co-supervisor, Dr. Samantha Gildenhuys, for her continuous guidance and helpful discussions over the years. Thanks for extending your valuable knowledge and understanding on to me.

All current and past members of the Protein-Structure Function Research Unit, with special thanks to Nishal Parbhoo, thanks for passing on some of your skills.

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

TABLE OF CONTENTS

DECLARATION ii

ABSTRACT iii

ACKNOWLEDGEMENTS v

TABLE OF CONTENTS vi

LIST OF FIGURES viii

LIST OF TABLES ix

ABBREVIATIONS x

CHAPTER 1: INTRODUCTION 1

1.1. Protein Structural Hierarchy 1

1.2. The Protein Folding Problem 1

1.3. Protein Conformational Stability 2

1.4. Protein Folding Models 3

1.4.1. Two-state folding mechanism 4

1.4.2. Three-state folding mechanism 6

1.5. Protein-Protein Recognition Sites 6

1.5.1. Molecular complementarity 7

1.5.2. Interface physical and chemical properties 8

1.5.2.1. Amino acid composition 9

1.5.2.2. Electrostatic complementarity 9

1.5.2.3. Hydrophobic effect 10

1.6. Glutathione Transferases 11

1.7. Three Dimension Structure of Class Mu GST M1-1 12

1.8. Intersubunit Interactions in GSTs 12

1.9. Stability and Folding of GSTs 15

1.10. Objective 18

CHAPTER 2: EXPERIMENTAL PROCEDURES 19

2.1. Materials 19

2.2. Sequencing of the Mutant and Wild-type Plasmids 19

2.3. Transformation, Overexpression and Purification of rGST M1-1 and F56S/R81A rGST M1 20

2.4. Protein Concentration Determination 22
2.5. Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis 23
2.6. Size Exclusion High-Performance Liquid Chromatography 24
2.7. Enzyme Specific Activity 24
2.8. Spectroscopic Studies 25
2.8.1. Far-UV circular dichroism (CD) 25
2.8.2. Tryptophan fluorescence 26
2.8.3. Extrinsic fluorescence spectroscopy – ANS binding 27
2.9. Conformational Stability 27
2.9.1. Recovery of native F56S/R81A 27
2.9.2. Urea-induced equilibrium unfolding 28
2.10. Data Fitting 29
CHAPTER 3: RESULTS 31
3.1. cDNA Sequence Verification 31
3.2. Purification of Wild-type rGST M1-1 and F56S/R81A rGST M1 31
3.3. Protein Size Determination 34
3.4. Functional and Structural Characterisation 38
3.4.1. Enzymatic activity 38
3.4.2. Spectroscopic properties of rGST M1-1 and rGST M1 38
3.4.2.1. Far-UV circular dichroism 38
3.4.2.2. Intrinsic tryptophan fluorescence 40
3.4.2.3. ANS binding studies 43
3.5. Equilibrium Unfolding of the F56S/R81A rGST M1 Monomeric Mutant 45
3.5.1. Recovery of native structure 45
3.5.2. Urea-induced unfolding of the F56S/R81A mutant protein 45
3.5.3. Urea-induced unfolding in the presence of ANS 48
CHAPTER 4: DISCUSSION 51
4.1. Quaternary Interactions are Important for the Catalytic Function of rGST M1-1 51
4.2. Loss of Quaternary Contacts Results in Enhanced Exposure of Hydrophobic Surface Area 54
4.3. Quaternary Interactions Increase the Stability of the Individual Subunits 55
CHAPTER 5: REFERENCES 58

LIST OF FIGURES

Figure 1-1: Commonly observed folding pathways for dimeric proteins. 5

Figure 1-2: Three dimensional structure ribbon representation of homodimeric rGST M1-1 viewed down its two fold axis. 13

Figure 3-1: A portion of the nucleotide sequencing chromatograms for the rGST M1-1 cDNA ORF indicating the codon mutations responsible for the corresponding F56S/R81A rGST M1-1 double mutant protein sequence. 32

Figure 3-2: Analysis of the F56S/R81A mutant protein overexpression 33

Figure 3-3: Elution profiles of wild-type rGST M1-1 and F56S/R81A rGST M1. 35

Figure 3-4: SDS-PAGE analysis of purified proteins. 36

Figure 3-5: SE-HPLC elution profile of rGST M1-1 and F56S/R81A monomeric mutant. 37

Figure 3-6: Specific activity of wild-type rGST M1-1 and F56S/R81A mutant. 39

Figure 3-7: Far-UV circular dichroism spectra of native wild-type rGST M1-1 and F56S/R81A rGST. 41

Figure 3-8: Fluorescence emission spectra of wild-type GST M1-1 and F56S/R81A monomeric mutant. 42

Figure 3-9: Fluorescence emission spectra of ANS bound to wild-type rGST M1-1 and the F56S/R81A rGST M1 mutant. 44

Figure 3-10: Recovery of the F56S/R81A rGST M1. 46

Figure 3-11: Urea-induced equilibrium unfolding of F56S/R81A. 47

Figure 3-12: GST M1 urea induced unfolding monitored by following ANS fluorescence. 50

Figure 4-1: Important residues at the G-site of rGST M1-1. 52

LIST OF TABLES

Table 3-1: Equilibrium unfolding thermodynamic parameters. 49

ABBREVIATIONS

A	Absorbance
A_{280}	Absorbance at 280 nm
Å	Ångström
A1-1	Alpha class glutathione *S*-transferase with two type 1 subunits
ANS	8-Anilino-1-naphthalenesulphonate
CD	Circular dichroism
cDNA	Complementary DNA
CDNB	1-Chloro-2,4-dinitrobenzene
C_m	Chemically-induced midpoint of unfolding
Da	Dalton
ΔASA	Change in solvent accessible surface area
Δ*G*	Change in Gibbs free energy
ΔG_f	Change in Gibbs free energy of unfolding
Δ*G*(H2O)	Change in Gibbs free energy in the absence of denaturant
Δ*H*	Change in enthalpy
Δ*S*	Change in entropy
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
DTT	Dithiothreitol
E. coli	*Escherichia coli*
EDTA	Ethylene diamine tetraacetic acid
ε	Molecular extinction coefficient
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 state
I_2	Dimeric intermediate state
IPTG	Isopropyl-β-D-thiogalactoside
ITC	Isothermal titration calorimetry

K_d	Dissociation constant
K_{eq}	Equilibrium constant
λ_{em}	Emission wavelength
λ_{ex}	Excitation wavelength
λ_{max}	Maximum emission wavelength
L-site	Non-substrate ligand binding site
M1-1	Mu class glutathione *S*-transferase with two type 1 subunits
[θ]	Mean residue ellipticity
MOPS	4-Morpholinepropanesulfonic acid
Mr	Relative Molecular Mass
m-value	Susceptibility of protein to denaturant, i.e., cooperativity
N	Native state
N_2	Native dimeric state
OD	Optical density
P1-1	Pi class glutahione *S*-transferase with two type 1 subunits
PDB	Protein Data Bank
R	Universal gas constant
rGST	Glutathione *S*-transferase from rat
rGST M1	Mu class rat glutathione *S*-transferase monomer
rpm	Revolutions per minute
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
SE-HPLC	Size-exclusion high performance liquid chromatography
Sj26GST	26 kDa GST from *Schistosoma japonicum*
T	Temperature in Kelvin
TEMED	N,N,N'N'-tetramethylethylenediamine
Tris	Tris(hydroxymethyl)-aminomethane
U	Unfolded state
UV	Ultra-violet
YT	Yeast tryptone

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

CHAPTER 1
INTRODUCTION

1.1. Protein Structural Hierarchy

Protein structural levels of organisation were first described by Linderstrom-Lang and Bernal (Jaenicke and Rudolph, 1986). They referred to the primary structural level as the specific amino acid sequence along the covalent polypeptide chain. This sequence is governed by the genetic code, which specifies which nucleic acid codon corresponds to which amino acid. The regular arrangement of this linear amino acid sequence into structural motifs such as α-helices and β-sheets is defined as the proteins secondary structural level, and further arrangement of these motifs gives rise to what are referred to as supersecondary structures. The tertiary structural level is achieved by the folding of the secondary structural elements into one or more functional domains. In some cases, proteins are made up of only one polypeptide chain and hence have the tertiary structural level as their highest level of organisation, and these are known as monomeric proteins. Having said that, it should however be noted that most proteins are composed of two or more polypeptide chain, and such proteins are referred to as oligomeric proteins (dimer – 2 chains; trimer – 3 chains; tetramer – 4 chains). Oligomeric proteins have the quaternary structure as their highest level of organisation. The quaternary structural level refers to the arrangement and interactions of the individual protein subunits (folded chains) with each other, and this assembly is mediated by protein-protein interactions.

1.2. The Protein Folding Problem

Inside the cell, proteins are synthesised, according to what has come to be known as the Central Dogma of molecular biology, that is, DNA → RNA → protein, as linear amino acid chains, which subsequently fold into unique 3-D structures known as the native states. Since a wide range of biochemical functions are regulated by the 3-D structure of proteins, understanding the mechanism by which the specific amino acid sequence of a protein directs its efficient folding to its functional native state represents one of the greatest unsolved problems in molecular biology and this has become known as the ‘protein folding problem’ (for a review see Dill *et al.*, 2008). The prediction of a proteins 3-D structure from its linear amino acid sequence has been an interesting task in molecular biology since the pioneering

work on protein folding, using bovine pancreatic ribonuclease, performed by Anfinsen. He proposed that a protein's amino acid sequence contains all the information required to direct its final 3-D structure and in turn its biological function (Anfinsen, 1973). As one might imagine, it would take an extremely long period of time for even the smallest protein to fold to its native conformation by randomly exploring all its possible conformations. Yet, despite the large degrees of freedom, many proteins have been found to fold into their native states in less than a few seconds, which suggests that they fold by some sort of pathway to reach their native state. This inference has become known as the Levinthal paradox (Levinthal, 1968; for a review see Dill and Chan ,1997). We now know that both structural changes and chemical interactions occur throughout the entire folding process, and to bring about the amino acid sequence to its native conformation within a very short time period all the mechanisms involved in the folding process need to be strongly cooperative (Notting, 1995).

It has been suggested that the folding problem can be broken down into three related questions (Creighton, 1990). Firstly, what the kinetic mechanism by which the protein adopts its native biologically active conformation is (i.e. the time limit within which the polypeptide chain must obtain its native conformation)? What is the physical basis for the native states stability? Finally, we need to understand why does a given amino acid sequence determine one particular folding process and resultant 3-D structure, as opposed to some other structure? Obtaining answers to these questions would make the prediction of the proteins 3-D structure based on the amino acid sequence of the polypeptide chain possible, however, as it has been found, this is not an easy task.

1.3. Protein Conformational Stability

The understanding of how proteins fold into stable biologically active 3-D structures is of great importance in most fields of molecular biology. There has been a lot of progress over the past couple of decades regarding our attempts of extending our knowledge on this subject, however, our understanding is still quite far from full. We now know that proteins fold spontaneously into their native state, and are guided by a decrease in free energy of the system. The stability of a given protein is defined by the difference between the interactions that favour the native (folded) state and those that favour the denatured (unfolded) state and this stabilisation may involve all levels of the protein structural hierarchy, local polypeptide chain packing, secondary and supersecondary structural elements, domains and subunits

(Jaenicke, 1987; Ragone and Colonna, 1995; Ragone, 1997; Facchiano *et al*., 1998; Jaenicke, 1998). This difference is best represented by the change in Gibbs free energy (ΔG) of the native and the denatured state upon complete unfolding of the native folded protein. Under isobaric conditions, ΔG is contributed to by two other thermodynamic parameters and can be described as follows:

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

where T is the temperature, ΔH and ΔS are the enthalpic and entropic contributions to the ΔG value. The enthalpic term contribution to ΔG arises predominantly from the loss of intermolecular and protein-solvent hydrogen bonds and van der Waals contacts. The formation of above mentioned interactions, as well as the formation of salt-bridges and solvent-solvent interactions at the surface of the protein molecule contributes to the binding enthalpy of the reaction (Dill, 1990). On the other hand, contributions to the ΔG by the entropic term arise from a combination of factors including the entropy contribution due to changes in conformational degrees of freedom of the polypeptide backbone as well as the amino acid side-chain (ΔS_{conf}) and the reorganisation of solvent molecules due to the burial of hydrophobic groups (ΔS_{solv}) (Murphy, 1999). Conformational entropy (ΔS_{conf}) is lost upon the folding of the linear polypeptide chain to a more compact 3-D structure, resulting in an unfavourable contribution to the folding process. However, this is compensated for by the formation of intramolecular and intermolecular interactions as the native state is acquired, resulting in the stabilisation of the native state (Dill, 1990).

The forces responsible for protein stability are the same as those responsible for protein folding since along the folding pathway and association, the polypeptide chain gains stability (Jaenicke, 1999). Globular proteins have been reported to exhibit marginal stabilities equivalent to a few intermolecular interactions (Dill, 1990; Jaenicke, 1991). On average, the values for the Gibbs free energy of stabilisation of small proteins are reported to lie anywhere between 5 and 15 kcal.mol^{-1} (Pace, 1990), and the strength of a single hydrogen bond is reported to range from 2 to 14 kcal.mol^{-1}, depending on the geometry of the interaction (Hagler *et al*., 1979; Dauber and Hagler, 1980; Privalov and Makhatadze, 1993).

1.4. Protein Folding Models

Dimeric proteins have been shown to be significantly more stable than their monomeric counterparts with stabilities ranging from 10-27 kcal/mol as compared to 5-15 kcal/mol (Neet and Timm, 1994; Rumfeldt *et al*., 2008). The additional stability is suggested to arise from

the subunit-subunit interactions contributing an excess of 65% of the dimeric protein stability (Neet and Timm, 1994). Dimeric proteins can fold via many different mechanisms ranging from the simplest mechanism of a two-state transition, involving only the native dimer (N_2) and unfolded monomers (U), to also forming a number of monomeric (M) or even dimeric intermediates (I_2) (see Figure 1-1, scheme 1-3). However, the most commonly observed mechanisms are the two-state and the three-state with monomeric and dimeric intermediates. Equilibrium chemical (e.g. urea or guanidinium chloride) denaturation and renaturation curves combined with measurements of the unfolding and refolding kinetics are commonly used in analysing the stability and folding mechanisms of proteins (Fersht, 1999). In addition to the Gibbs free energy (ΔG) for the structural transitions (i.e. folded, intermediate, and unfolded), the measured experimental quantities also include the denaturant dependence of ΔG (*m*-value),a value which is proportional to the change in solvent accessible surface area for the (ΔASA) for the structural transition (Myers *et al*.,1995). For the experimental results to be valid, it is important that protein folding be reversible, that equilibrium is reached before denaturation and renaturation curve measurements are taken, and that these curves are analysed using the appropriate model.

1.4.1. Two-state folding mechanism.

The two-state transition is the simplest for a dimeric protein (see Figure 1-1, scheme 1). Correspondence between the ΔG and the *m*-value obtained from the kinetic and equilibrium data is a key test for two-state behaviour. Dimeric proteins that exhibit two-state behaviour often fold with three or more state kinetics, whereas those that exhibit three-state equilibria have all been found to fold with at least three-state kinetics. The equilibrium unfolding curves of all dimeric proteins must be dependent on protein concentration as long as there is a change in molecularity upon unfolding of the native state (Ragone, 2000). Monomeric proteins that unfold via a two-state mechanism exhibit symmetrical denaturisation curves that are independent of protein concentration, in contrast, for dimeric proteins the transition is slightly asymmetric, being skewed toward the lower denaturant concentration. For dimeric proteins, the denaturation curve shape remains unchanged with increasing protein concentration, however, as a result of mass action, the mid-point shifts to higher denaturant concentration. This occurs because at a given denaturant concentration, as protein concentration increases, the ratio of dimer to unfolded monomer will increase so as to satisfy the equilibrium constant. In contrast, for monomeric proteins, the ratio of folded to unfolded

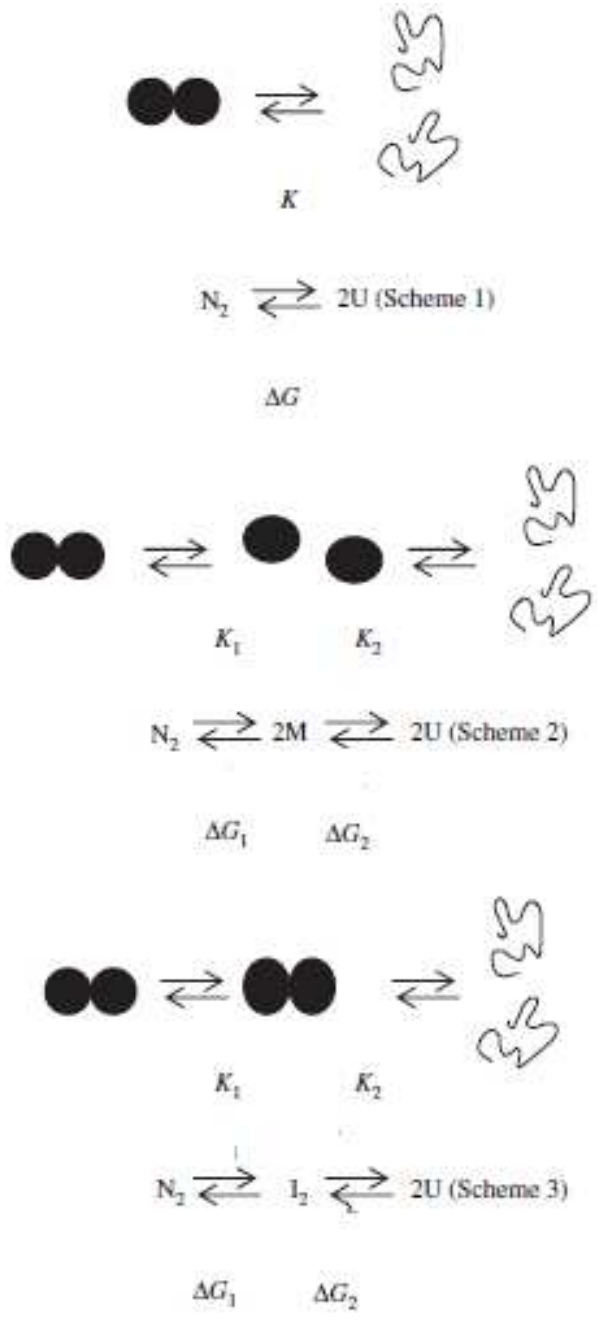


Figure 1-1: Commonly observed folding pathways for dimeric proteins.

Scheme 1 represents a simple two-state transition for a dimeric protein. Schemes 2 and 3 represent a three-state transition with the formation of monomeric intermediates and formation of a dimeric intermediate, respectively. N_2 is the native dimer; M is the monomeric intermediate; I_2 is the dimeric intermediate and U is the unfolded monomer. Adapted from Rumfeldt *et al.* (2008).

protein is independent on protein concentration since the transition does not involve a change in molecularity. Because there is no simple linear or proportional correspondence between

the ΔG values of dimers with those of monomers, one has to take caution when comparing ΔG values between dimeric and monomeric proteins (Park and Marqusee, 2004).

1.4.2. Three-state folding mechanism

The three-state folding mechanism of dimeric proteins involves the formation of either a monomeric or dimeric intermediates (Figure 1-1, scheme 2 and 3). Generally, the observation of a bimodal equilibrium curve is an indication for the population of an intermediate for both monomeric and dimeric proteins. For dimeric proteins the general trend is that a monomeric intermediate becomes more populated as protein concentration is decreased whereas when protein concentration is increased, dimeric intermediates become more populated (Hobart *et al.*, 2002; Mallam and Jackson, 2005; Park and Bedouelle, 1998). It should however be noted that, depending on the nature of the two transitions, for some protein concentrations only the native dimer and unfolded monomer will be significantly populated and as a result equilibrium transitions will appear to be two-state. Thus it is crucial to acquire equilibrium curves for a range of protein concentrations in the concentration range of the K_d in order to better define the mechanism, that is, to maximize the observed effect.

1.5. Protein-Protein Recognition Sites

The association of protein subunits plays an important role in stabilising the structure of oligomeric proteins and generating new functions, binding of ligands or regulation. Subunit-subunit associations that occur as a result of the formation of the protein quaternary structure involve specific complementary recognition between the individual macromolecules. The native quaternary structure of an oligomeric protein is directed by both the spatial arrangement of the individual subunits as well as the interactions between them (Anfinsen and Scheraga, 1975; Jaenicke, 1984). These associations are thought to be highly specific, and this specificity is achieved through complementary surface shape, hydrophobic patches, hydrogen bonding and ion pairing. Previous studies have suggested that, following folding of the linear amino acid sequence, there must be a unique pathway of association for oligomeric proteins (Jaenicke and Rudolph, 1986). This idea stems from the existence of well-populated folded protein intermediates capable of ‘subunit recognition’ within the mechanism of quaternary structure formation (Jaenicke and Rudolph, 1986; Jaenicke, 1987).

Protein-protein interactions determine a wide range of protein structures and functions, and thus understanding these interactions is of fundamental importance to molecular biology. In nature, many proteins exist as oligomers due to the association of protein subunits. The existence of oligomeric proteins is beneficiary since it introduces reduced solvent exposed surface areas, increased stability of individual subunit's tertiary structure, and novel function formation at the interface (Larsen *et al.*, 1998). Interactions governing these subunit-subunit associations are highly specific and display good geometric and physicochemical complementarity at the interacting surfaces. Geometric complementarity refers to the geometric interlocking, charge reciprocation, and amino acid side chain packing at the interface (Jones and Thornton, 1996), whereas physicochemical complementarity is concerned with the amino acid composition of the interface (Janin *et al.*, 1988), electrostatics (Sheinerman *et al.*, 2000) and hydrophobicity (Tsai *et al.*, 1997; Tsai and Nussinov, 1997). Having said this, it should be clear that in order for us to fully comprehend molecular complementarity, the physicochemical properties of the protein interfaces must be understood in great detail.

1.5.1. Molecular Complementarity

Proteins interact through their surfaces, thus it comes as no surprise that the analysis of protein-protein interactions usually focuses on protein surfaces. Molecular complementarity refers to the geometric interlocking, charge reciprocation, and amino acid side chain packing at the subunit-subunit interface (Jones and Thornton, 1996). Subunit-subunit interfaces vary in topology, ranging from those which are relatively flat to those with a range of convoluted protein interface surfaces. Furthermore, highly complex subunit-subunit interfaces exist which may contain extended arms, deep cavities or, in some cases, loops which completely surround the neighbouring subunits (Chothia and Janin, 1975; Duquerroy *et al.*, 1991). The subunit-subunit interface surface is significantly different from both the solvent accessible surface and protein core in that it is more hydrophobic compared to the protein surface and less hydrophobic than the core (Jones and Thornton, 1996). Although exceptions exist that demonstrate no hydrophobic interactions occur at the interface (Tsai *et al.*, 1997), the hydrophobic effect is considered to be the primary driving force in the stabilization of protein association. In many oligomeric proteins, the interfaces have been found to be composed of inter-dispersed hydrophobic patches with intersubunit hydrogen bonds and water molecules along the entire interface (Larsen *et al.*, 1998). Furthermore, the stabilisation of the interface

itself is brought about by the interactions within these hydrophobic patches (hydrophobic and van der Waals) (Chothia and Janin, 1975; Argos, 1988; Janin *et al.*, 1988; Janin and Chothia, 1990; Jones and Thornton, 1995; Jones and Thornton, 1996).

During the association event of two or more protein subunits, complementary molecular surfaces significantly reduce their solvent exposed surface area, and by so doing, gain stability for complex formation from intramolecular hydrophobic effects (Meyer *et al*., 1996). This determines the appropriate positioning and orientation of the individual subunits for the formation of complementary hydrogen bonds and other electrostatic and van der Waals interactions, with minimum steric hindrance (Meyer *et al.*, 1996). Although the hydrophobic effect is considered to be the primary driving force for the association of individual subunits, the specificity of the association is in fact brought about by the hydrogen bonds and the electrostatic interactions (Fersht, 1984). The folding of the polypeptide chain itself tends to reduce the surface of the polypeptide chain in contact with the solvent, thus replacing solvent-protein interactions with solvent-solvent interactions that are thermodynamically favourable, and this forms the structural basis of the hydrophobic effect (Janin e*t al.*, 1988).

1.5.2. Interface physical and chemical properties

The ability to predict the preferred manner by which protein subunits associate with one another is of great interest in molecular biology (Ehrlich *et al.*, 1998; Gallet *et al*., 2000; Fariselli *et al.*, 2002; Lichtarge and Sowa; 2002; Janin *et al.*, 2003; Ofran and Rost, 2003;) since many biological processes are regulated by the formation of protein-protein complexes (Klenthous *et al*., 2000). Quite a significant number of protein structure studies performed in the past have addressed protein-protein association, however, the principles governing these interactions are yet to be fully understood. For instance, although we now recognise protein-protein binding sites to be mainly hydrophobic, planar, globular, protruding (Chothia and Janin, 1975; Argos, 1988; Jones and Thornton, 1997) and composed of relatively large surfaces with good geometric and electrostatic complementarity (Jones and Thornton, 1996; Janin, 1995; Janin, 1997), no general patterns are observed amongst different protein-protein complexes. Thus, a lot of consideration needs to be taken on the physical and chemical properties of protein interfaces, including the amino acid composition of the interface (Janin *et al.*, 1988), electrostatic interactions at the interface (Sheinerman *et al.*, 2000), and its hydrophobicity (Tsai *et al.*, 1997; Tsai and Nussinov, 1997) in order to fully understand

molecular complementarity, and thus enhance our understanding of the protein's biological function.

1.5.2.1. Amino acid composition

Protein-protein interfaces display distinctive amino acid composition compared to the protein surface and core. The energetic contributions of individual amino acid side chains on protein-protein association, has been explored through the use of alanine scanning mutagenesis. Using this technique, previous studies have shown that single residues can contribute a large fraction of the binding free energy. Previous studies have indicated that the binding free energies are not distributed uniformly across the entire interface; rather, there are certain critical residues, referred to as "hot spots", which contribute the most to the free energy of binding (Cunningham and Wells, 1991; Clackson and Wells, 1995; Bogan and Thorn, 1998). Just over a decade ago, a dataset a of protein interface alanine mutants was compiled by Bogan and Thorn (1998). They then referred to those residues for which the change in binding free energy upon mutation to an alanine gave rise to a distinct increase in the absolute binding free energy of more than 2 kcal/mol as "hot spots". Hot spots are not homogeneously distributed in the interface, rather, they form one or more densely packed clusters of networked hot spots (Keskin *et al.*, 2005). Although protein interfaces are predominantly hydrophobic, specific conserved polar residues at specific locations serve as energy hot spots, furthermore, Keskin *et al.* (2005) have shown good correlation to exist between "hot spots" and structurally conserved residues. It should however be noted that, the proportion of polar and non-polar residues at these interfaces varies considerably (Janin *et al*., 1988). It has been found 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.

1.5.2.2. Electrostatic complementarity

Together with geometric complementarity, electrostatic interactions (more specifically, salt bridges as well as hydrogen bonds) at protein interfaces play a crucial role in generating unique structures and conferring conformational specificity (Fersht, 1984). The contribution of electrostatic interactions to protein structure may be attractive or repulsive, thus they can either stabilise or destabilise protein-protein interactions. The stabilizing effects arise due to the formation of ion pairs (or salt-bridges) between two oppositely charged groups within 4 Å

of each other (Barlow and Thornton, 1983), whilst the destabilising effects arise from the repulsion between similar charged groups. In proteins, electrostatic interactions are characterised by interaction between ionisable amino acid residues such as histidine, arginine, lysine aspartic and glutamic acid, as well as the N-terminal amide and C-terminal carboxyl group. Electrostatic interactions may occur between residues that are far from one another along the polypeptide chain, there by contributing to the tertiary conformation. In addition, these interactions may also occur between residues of two polypeptide chains (i.e. subunits), usually through charge clusters, thus contributing to the protein's quaternary conformation. A charge cluster refers to the distribution of charged amino acid residues within a protein with high charge concentration (or density) relative to the rest of the protein's charge distribution (Zhu and Karlin, 1996). Graph-spectral analysis has allowed for the identification of important charge cluster residues in protein dimerisation (Brinda *et al.*, 2002). Such clusters are generally found near the centre of the protein interface, serving to stabilise protein conformation and playing an important role in processes such as protein transport, localisation and regulation.

1.5.2.3. Hydrophobic effect

The hydrophobic effect is the primary driving force in the stabilisation of protein-protein associations (as well as protein folding), and arises from the absence of hydrogen bonding between water and non-polar groups rather than the presence of favourable interactions between the non-polar groups themselves (Dill, 1990). The importance of the hydrophobic effect on protein structure was first identified in 1959 by Kauzmann, who proposed that in aqueous environment, proteins tend to bury the non-polar amino acid residues, while orienting those with polar or charged side chains to interact with the solvent molecules (Kauzmann, 1959). Since the transfer of non-polar molecules to an aqueous environment disrupts the hydrogen bonding pattern of water, the transfer of non-polar amino acid residues from a non-polar environment into a polar environment is energetically costly, thus the burial of these residues via the association of protein subunits (or in the folding reaction) is energetically favourable (Dill, 1990). Although the contribution of the hydrophobic effect has been found not to be strong for protein associations as it is for protein folding process (Dill, 1990; Tsai and Nussiv, 1997), previous studies have identified the subunit interfaces of homodimers to be hydrophobic and to contain twice the buried surface area as compared to other protein complexes (Bahadur *et al.*, 2003)

1.6. Glutathione Transferases

The detoxification enzyme system consists of three major groups of enzymes which are responsible for the elimination of foreign, cytotoxic and genotoxic compounds (Jacoby and Ziegler, 1990). Enzymes that fall under the first group are known as Phase I enzymes, and these enzymes form reactive functional groups within the foreign compound. In some cases the addition of a functional group to the foreign compound results in a compound that is more reactive than the parent compound, and thus more toxic. Phase II enzymes (those belonging to the second group) are able to deactivate these reactive species by adding hydrophilic moieties to them, resulting into a more hydrophilic compound which can thereafter be eliminated via phase III pathways (Dirr *et al.*, 1994).

Glutathione transferases (GSTs, EC 2.5.1.18) are members of a super-gene family of multifunctional proteins that are grouped into various species-independent gene classes (Dirr *et al.*, 1994; Armstrong, 1997) and form a key component of the Phase II detoxification enzyme system (for a review see Hayes *et al.*, 2005). GSTs specifically function to protect cellular macromolecules, such as proteins and nucleic acids (RNA and DNA), from being attacked by a wide variety of reactive electrophilic endogenous and exogenous compounds by catalyzing the nucleophilic addition of the thiol group of reduced glutathione (GSH, γ-glutamyl-cysteinyl-glycine) to their electrophilic centers (Armstrong, 1991; Rushmore and Pickett, 1993). Furthermore, in addition to this catalytic function, GSTs have also been reported to play a role in the regulation of signalling pathways via protein-protein interactions (for a review see Sheehan *et al.*, 2001). These enzymes are divided into 2 distinct super-family members, namely the membrane bound microsomal and cytosolic family members (Hayes and Strange, 2000; Jakobsson *et al.*, 1994). The cytosolic GSTs exist as stable homo- or heterodiners with a conserved archetypical fold (Dirr *et al.*, 1994; Wilce and Parker, 1994; Armstrong, 1997), and dimerisation is reported to be highly specific occurring only between subunits within the same gene class. Cytosolic GSTs are classified into at least sixteen gene classes with subunit lengths ranging between 199 and 244 amino acids (Sheehan *et al.*, 2001; Hayes *et al.*, 2005). Eight of these sixteen Cytosolic GST classes are mammalian, including classes Alpha (A), Mu (M), Pi (P), Sigma (S), Theta (T), Omega (O), Kappa (K) and Zeta (Z) (Mannervik, 1985; Buetler and Eaton, 1992; Board *et al.*, 1997; Board *et al.*, 2000; Pemble *et al.*, 1996). Human and rodent cytosolic GSTs are reported to share more than 40% identity

for isoenzymes within the same gene class, and less than 25% identity for isoenzymes between classes (Sheehan *et al.*, 2001; Hayes *et al*., 2005)

1.7. Three Dimension Structure of Class Mu GST M1-1

Cytosolic class Mu GST M1-1 has been crystallized as a homodimer with a molecular mass of 51 500 Da (Ji *et al.*, 1992). Each subunit consists of 217 amino acids that fold into two structural domains with no disulfide bonds (Ji *et al.*, 1992) (Figure 1-2). The N-terminal domain (also referred to as domain 1) is made up of residue 1-81 in class Mu GST M1-1 and adopts a thioredoxin fold (Dirr *et al*., 1994; Wilce and Parker, 1994, Martin, 1995), consisting of four β-sheets with three flanking α-helices (Figure 1-2). The fold consists of an N-terminal βαβ and a C-terminal ββα arrangement linked by an α-helix (α-2) resulting in an overall βαβαββα motif (Dirr *et al*., 1994; Wilce and Parker, 1994). The four β-sheets are reported to be in the same plane, with α-helix 2 positioned above this plane facing the solvent and the other two helices (α-1 and α-3) being below the β-sheet plane (Sheehan *et al*., 2001). The N-terminal domain has been found to provide most of the GSH binding site and to be highly conserved through GST classes. The larger, all α-helical, C-terminal domain (domain 2) runs from residue 90-217 and is connected to the N-terminal domain by a short linker sequence. In the case of Mu class, the C-terminal domain consists of 5 α-helices (Ji *et al*, 1992) (Figure 1-2).

1.8. Intersubunit Interactions in GSTs

In GSTs, dimerisation is required for the formation of a fully functional active site, part of which is located at the dimer interface (Sayed *et al*., 2000; Stenberg *et al*., 2000; Dirr, 2001). In addition to maintaining the functional conformation of the active site, interactions at the GST dimer interface have also been reported to contribute substantially to the conformational stability (Aceto *et al*., 1992; Erhardt and Dirr, 1995; Kaplan *et al*., 1997; Stevens *et al*., 1998; Wallace *et al*., 1998; Hornby *et al*., 2000) and the dynamics (Codreanu *et al*., 2005; Thompson *et al*., 2006) of the individual subunits, as well as the cooperative behaviour between the two subunits (Caccurri *et al*., 1999; Wongsantichon and Ketterman, 2006; Fabrini *et al*., 2009).

Dimeric protein often tend to use large hydrophobic amino acid residues such as Phe, Trp, and Tyr, which protrude from one subunit and anchor into the other subunit to physically join

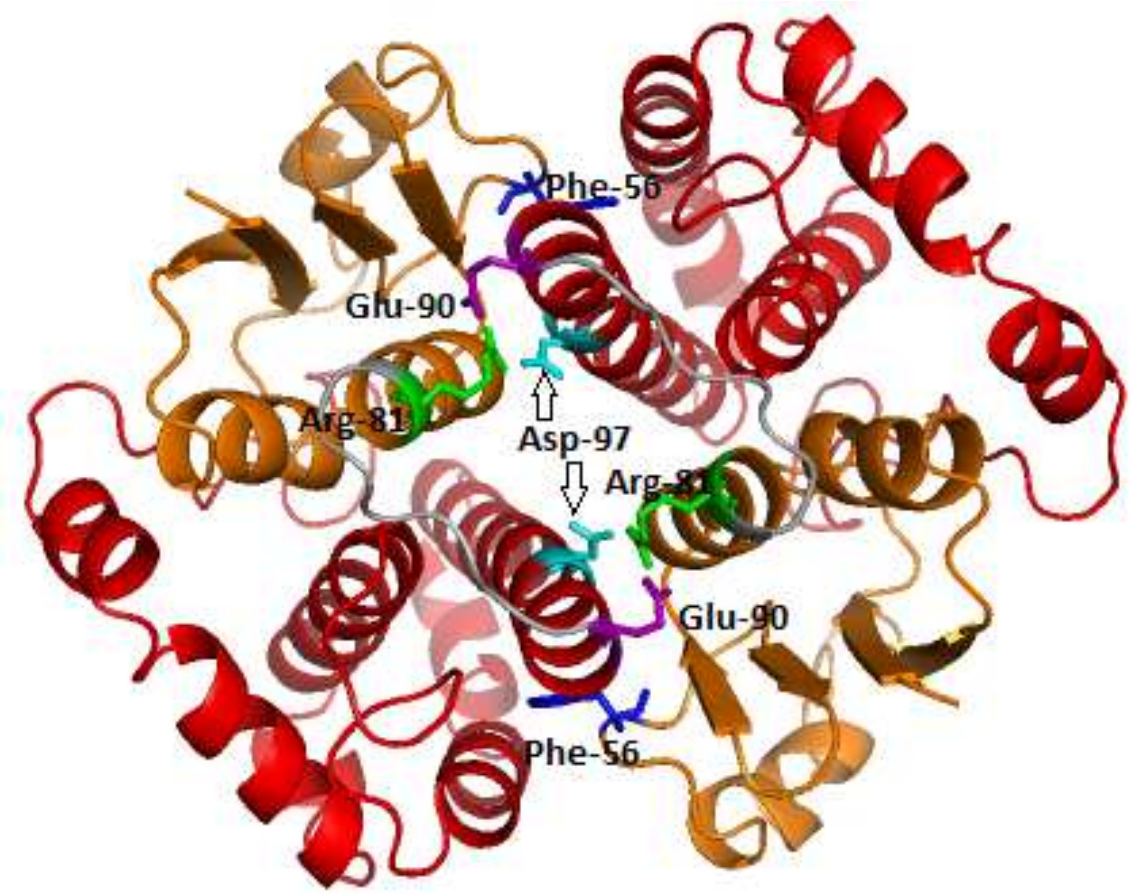


Figure 1-2: Three dimensional structure ribbon representation of homodimeric rGST M1-1 viewed down its two fold axis.

The N-terminal domain is represented in orange, while the C-terminal domain is in red. Phe-56 (blue) is located on the N-terminal domain and acts as the key of the hydrophobic lock-and-key motif, whereas Arg-81 (green) is involved in the mixed charge cluster, where it forms salt bridges with Glu-90 (purple) and Asp-97 (turquoise) of the adjacent subunit. This figure was generated using PyMOL™ v0.99 (DeLano Scientific, 2006) (PDB code 6gst).

the two subunits together (Jones and Thornton, 1995). The investigation of the structural features at the subunit interface of various GST gene classes has revealed the existence of two major types of interfaces; namely, the relatively flat and hydrophilic Sigma/Theta type with an extensive network of polar interactions and the Alpha/Mu/Pi/Sj26GST type which is characterised by a curved topography with a hydrophobic lock and key motif and a few polar interactions (Armstrong, 1997). Examination of rGST M1-1 as a representative for the curved interface group has revealed the existence of a hydrophobic lock-and key interaction motif at either end of the subunit interface while more polar contacts are located at the dimer's two-fold axis (Ji *et al*., 1992).

Two groups of interactions form the core of the interface contacts in class Mu GST M1-1 (Thompson *et al.*, 2006). The first is the hydrophobic contact between the subunits formed by the insertion of the phenyl ring of Phe-56 protruding from the N-terminal domain of one subunit into a hydrophobic pocket created by α-helix 4 and α-helix 5 within the C-terminal domain of the other subunit (Hornby *et al.*, 2002). The second is a mixed-charge electrostatic cluster between β-4/α-3 of the N-terminal domain and α-4 of the C-terminal domain. In this electrostatic interaction cluster, Arg-81, found within the N-terminal domain of one subunit, forms salt bridges with Glu-90 and Asp-97 in the C-terminal domain of the other subunit (Figure 1-2) (Zhu and Karlin, 1996; Luo *et al.*, 2002)

In an earlier study, the construction of a stable monomeric form of class Pi GST was archived by introducing a total of ten specific mutations in the hydrophobic region of the subunit interface (Abdalla *et al.*, 2002), however, since drastic changes had been made to the native proteins amino acid sequence, this monomer had lost its catalytic activity. Individual disruption of both of the above mentioned dimer interface interaction motifs (i.e. the hydrophobic lock-and-key motif and the mixed charge cluster), as well as combined mutations, has been reported to result in a shift in the equilibrium from dimer towards monomer (Vargo *et al.*, 2004). In order to gain insight on the nature and mechanics of subunit recognition in GST isoenzymes, Thompson *et al.* (2006) disrupted both subunit interface motifs in rGST M1-1 by using PCR based site-specific mutagenesis to mutate Phe-56 to Ser and Arg-81 to Ala which resulted in a stable GST M1 monomer. During the same period that this study was conducted, Hearne and Colman (2006) showed that a monomeric form of rGST M1-1 can also be constructed by disrupting only the interactions in the electrostatic

region of the dimer interface. In addition, a monomeric form of class Pi GST has also been constructed in a similar fashion (Huang *et al.*, 2008).

1.9. Stability and Folding of GSTs

Dimer interfaces can have a profound effect on the stability as well as the unfolding mechanisim of a protein (Rumfedlt *et al.*, 2008; Han *et al.*, 2008). GSTs are obligate dimeric proteins, with a complex canonical structure whose assembly involves interactions at the domain interface and at the subunit interface. Thus, it comes as no surprise that the dependence of the subunit stability on the quaternary interactions is of interest to molecular biologists. Previous studies have suggested that interdomain interactions make a significant contribution to the stability of the individual subunits that make up the GST dimer (Luo *et al*., 2002). Furthermore, unfolding studies on dimeric GSTs and monomeric GST-like proteins indicate that GST domains do not unfold independently of one another, rather, they do so as a single coopetative unit (Erhardt and Dirr, 1995; Kaplan et al., 1997; Stevens *et al.*, 1998; Wallace et al., 1998; Hornby *et al.*, 2000; Fanucchi *et al.*, 2008; Gildenhuys *et al.*, 2008; Fabrini *et al*., 2009).

Equilibrium unfolding studies of GST A1-1 (Wallace et al., 1998), P1-1 (Dirr and Reinemer, 1991; Erhardt and Dirr, 1995) and Sj26GST (Kaplan et al., 1997) indicate a tight coupling between association/dissociation and unfolding/refolding, following a cooperative two-state unfolding pathway. On the other hand, classes Mu (Hornby *et al.*, 2000) and Sigma (Stevens *et al.*, 1998) have been shown to display a multistate pathway, with stable monomeric or even dimeric folding intermediates.

For only a few oligomeric (Milla and Sauer, 1994; Gittelman and Matthews, 1990; Liang and Terwilliger, 1991) and many monomeric proteins (Kim and Baldwin, 1990), unfolding kinetics have been shown to be monophasic, agreeing with the "all-or-none" unfolding mechanism for a homogeneous population of protein. However, the unfolding kinetics of hGST A1-1 reveal that this enzyme unfolds via a biphasic mechanism, with a fast and a slower unfolding process (Wallace *et al.*, 1998). In this study, hGST A1-1 was proposed to unfold via the following model:

$$N_2 \rightarrow N_2^* \rightarrow 2U_A \qquad (2)$$

where N_2 represents the folded dimer, N_2^* represents the native-like dimeric intermediate, and U_A is the unfolded state of hGST A1-1. It has been reported that for hGST A1-1, no change is observed in the rate limiting step, and therefore the fast unfolding process represents a partial dissociation of the two domains and not the formation of the structured intermediate (Wallace *et al.*, 1998). Unlike GST M1-1, disruption of the lock-and-key motif in class Alpha GST A1-1 does not result in a stable monomer.

In class Mu GST M1-1, Tyr-6 is one of the amino acid residues found at the enzyme active site and is crucial for efficient catalysis. Residues that make up the 56-loop (the loop where Phe-56 is located) are in contact with this Tyr residue (Ji *et al.*, 1992). Backbone amide hydrogen-deuterium (H/D) exchange experiments have shown the 56-loop of the F56S/R81A mutant to display similar H/D exchange behaviour as the dimeric wild-type enzyme, however, native-like activity is not archived in the monomeric mutant (Thompson *et al,*. 2006). These findings suggest that the dimer interface interactions help in enforcing proper conformation of the 56-loop that is vital for GSH binding and efficient catalysis, and thus the disruption of the 56-loop is most probably the main reason for the low catalytic activity observed for the monomeric form of GST M1-1 (Thompson *et al.*, 2006)

In another study by Hearne and Colman (2006), it was found that the monomeric form of rGST M1-1 can also be attained by a single mutation in the electrostatic region of the subunit interface. In solution, GST M1-1 is at equilibrium between the monomeric and dimeric species, and the extent of either one of these species is dependent on a number of factors including enzyme concentration, conditions of the enzyme solution as well as the introduction of a mutation at the subunit interface. The mutation of Asp-97, which forms a salt bridge with Arg-81 of the adjacent subunit (Zhu and Karlin, 1996), to Lys results in a monomer which, like the F56S/R81A double mutant, exhibits decreased activity (Hearne and Colman, 2006). However, under the conditions stated in their study, Hearne and Colman (2006) found that the D97K mutant is almost as active as the native GST M1-1, which suggests that the decreased activity of the monomeric form of class Mu GST is attributed to intrasubunit changes rather than the generation of the monomeric species via the loss of intersubunit interactions (Hearne and Colman, 2006).

More recently, a similar study to the one described above was performed on human class Pi GST (Huang *et al.*, 2008). In this study, elimination of the charges of Arg-70, Arg-74, Asp-

90, and Asp-94 resulted in a decrease in enzyme activity and shifts the equilibrium toward the monomer. Since class Pi GST has been shown to retain substantial catalytic activity as a monomer (Ralat *et al.*, 2006; Ralat *et al.*, 2008), the observed loss in activity cannot simply be due to the dissociation of the dimer (Huang *et al.*, 2008). In contrast to the above mentioned mutations, mutating Thr-67 results in an enzyme similar to the wild-type, which exists as a monomer-dimer mixture but exhibits a higher enzymatic activity than the wild-type. This increased activity is most probably due to the proximity of Thr-67 of one subunit to Asp-94 of the other subunit, which can form hydrogen bonds. The replacement of Thr-67 with alanine removes the hydrogen bond and in turn increases the flexibility of the intersubunit cleft, and thus possibly making interactions within the protein that are more favourable for catalysis.

Previous studies have suggested that interactions at the subunit interface play an important role in stabilising the subunit tertiary structure as GSTs are obligate dimeric proteins (Dirr, 2001), however, the degree of subunit stabilization seems to be class dependent. There have been contradicting reports concerning the unfolding process of class Pi GST, making it unclear as to whether this enzyme follows a two-state (Dirr and Reinemer, 1991; Erhardt and Dirr, 1996) or three-state unfolding process (Aceto *et al.*, 1992). Recently this issue has been addressed and it has been suggested that class Pi GST P1-1 does, in fact, unfold via a three-state unfolding process where the intermediate state is dimeric and is formed due to a highly mobile α-helix 2 (Gildenhys *et al.*, 2010).

$$N_2(\alpha 2)_{f,m} \Leftrightarrow N_2(\alpha 2)_u \Leftrightarrow 2U \quad (3)$$

where N_2 is the native dimer with helix 2 (α2) in a folded and mobile conformation (subscripts f and m, respectively) or unfolded (subscript u) and U is the unfolded monomer.

1.10. Objectives

A number of previous studies have attempted to find experimental conditions that promote the dissociation of GSTs into single subunits and to study the catalytic and structural properties of the isolated monomers. Analysis of the structural data, as well as experimental data, reveals the major interactions across the dimer interface of rGST M1-1 to be the hydrophobic lock-and-key motif and the mixed charge cluster, and previous studies have shown that the disruption of these motifs results in a stable monomeric species (Thompson *et al.*, 2006). Thus, herein rGST M1-1 is used as a model to investigate the role played by subunit-subunit interactions in the stabilisation of the individual subunits in dimeric GST proteins. Furthermore since dimerisation is required for the formation of a fully functional active site, part of which is located at the dimer interface (Sayed *et al*., 2000; Stenberg *et al*., 2000; Dirr, 2001), the enzymatic activity of the variant protein is also compared to that of the wild-type enzyme to assess the catalytic functionality of the subunit in the absence of quaternary interactions.

CHAPTER 2

EXPERIMENTAL PROCEDURES

2.1. Materials

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/R81A rGST M1 expression plasmid was designed by switching the CGC (R) codon to GCC (A) in the F56S rGST M1 expression plasmid (kindly provided by Dr. J. A. T. Hornby , University of the Witwatersrand, Johannesburg) by John Walters (Protein Structure-Function Research Unit, University of the Witwatersrand, Johannesburg) (Thompson *et al.*, 2006). Inqaba Biotech (Pretoria, South Africa) conducted all sequencing to confirm the identity of the plasmids. DTT (dithiothreitol), IPTG (isopropyl-β-D-thiogalactopyranoside) and SDS-PAGE (sodium dodecyl sulphate polyacrylamide gel electrophoresis) molecular weight marker (SM0431) were purchased from Fermentas Life Sciences (St.Leon-Rot, Germany). ANS (8-Anilino-1-naphthalenesulfonic acid) and GSH (reduced L-glutathione) was purchased from Sigma-Aldrich (St. Louis, MO, USA). Ultrapure (99.5%) urea was purchased from Merck chemicals (Darmstadt, Germany). All other reagents used were of analytical grade. All solutions were prepared using distilled water and filtered through a 0.45 μm pore size filter.

2.2. Sequencing of the Mutant and Wild-type Plasmids

Samples of each of the purified wild-type rGST M1-1 and the F56S/R81A double mutant pET-20b(+) plasmids ware submitted to Inqaba Biotec (Pretoria, South Africa) for sequencing. The mutant pET-20b(+) cDNA open reading frame (ORF) which encodes the F56S/R81A rGST M1 protein, was sequenced to confirm the presence of the desired TTC→ GTC and CGC→GCC mutations and to confirm the absence of any other random mutations. In a similar manner, the wild-type pET-20b(+) cDNA open reading frame, encoding the wild-type rGST M1-1 protein, was sequenced to confirm the sequence of the wild-type enzyme. The nucleotide sequences were sequenced using modern techniques in accordance with the DNA sequencing technique suggested by Sanger *et al.* (1977). The primers previously designed to create the F56S/R81A rGST M1 expression plasmid by John Walters (Protein Structure-Function Research Unit, University of the Witwatersrand) had the following sequences:

Forward primer:

5' GC TAC CTT GCC ***GCT*** AAG CAC CAC CTG TGT GG **3'**

Reverse primer:

5' CC ACA CAG GTG GTG CTT ***AGC*** GGC AAG GTA GC **3'**

The bold codons represent the replacement of Arg-81 with Ala in the F56S/R81A rGST M1 primers. The determined nucleotide sequences were aligned with the known mRNA nucleotide sequence encoding glutathione *S*-transferase M1 using the alignment tool, BLASTN, in the NCBI BLAST suite of tools (Altschul *et al.*,1990).

2.3. Transformation, Overexpression and Purification of rGST M1-1 and F56S/R81A rGST M1.

The plasmids containing the cDNA insert that codes for the wild-type rGST M1-1 and the F56S/R81A rGST M1 mutant were used to transform *E. coli* BL21(DE3) pLysS cells (Lucigen, Middleton, WI, USA) and *E. coli* T7 Express *Iq* competent cells (New England Bio-labs Inc., Ipswich, MA, USA), respectively. The competent cells were transformed using a one-step method as described by Chung and colleagues (1989). Competent cells were thawed on ice for 15 minutes before adding 1 µL of 90 ng/µL plasmid DNA. Thereafter, the reaction mixture was incubated on ice for 30 minutes, heat-shocked at 42°C for 45 seconds (*E. coli* BL21(DE3) pLysS) and for 10 seconds (*E. coli* T7 Express *Iq*) on a heating block, and immediately transferred to an ice bath for 2 minutes. Cells were then grown by adding 900 µL of SOC medium (2% (w/v) tryptone, 0.5% (w/v) yeast extract, 250 mM KCl, 1 M glucose, 2 M $MgCl_2$) to the reaction mixture followed by incubation for 1 hour at 37°C with 250 rpm agitation. The cells were then plated on Luria Bertani-agar (LB-agar) plates (1% (w/v) tryptone, 0.5% (w/v) yeast extract, 1.0% (w/v) NaCl, 1.5% (w/v) agar) containing the antibiotics ampicillin and chloramphenicol to a final concentration of 100 $\mu g.mL^{-1}$ and 30 $\mu g.mL^{-1}$, respectively, and incubated overnight (~16 hours) at 37°C.

The transformed expression cells were added to 2xYT medium (1.6% (w/v) tryptone, 1.0% (w/v) yeast extract, 0.5% (w/v) NaCl) supplemented with the antibiotics at the same final

concentration as before and grown overnight at 37°C with 250 rpm agitation. A 100-fold dilution was then used to inoculate fresh, sterile 2xYT medium supplemented with ampicillin (100 μg.mL^{-1}) and chloramphenicol (30 μg.mL^{-1}). The cells were again then grown at 37°C with shaking at 250 rpm till an OD600 of ~1.0 was reached before adding IPTG to a final concentration of 0.7 mM (BL21(DE3) pLysS cells) and 0.1 mM (T7 Express *Iq* cells) in order to induce the over-expression of the class mu rGST proteins. The *E. coli* BL21(DE3) pLysS cells and *E. coli* T7 Express *Iq* cells were grown for a further 16 hours at 37°C and 20°C, respectively, with shaking at 250 rpm to achieve optimum protein expression.

Cells were harvested by centrifugation at 6500 × *g* for 20 minutes. *E. coli* BL21(DE3) pLysS cells (bearing rGST M1-1) and *E. coli* T7 Express *Iq* cells (bearing F56S/R81A mutant) re-suspended in 10 mM Tris-HCl, 1 mM EDTA, 200 mM NaCl, 0.02 % NaN_3, pH 7.8 and 20 mM MOPS, 1 mM EDTA, 1 mM DTT, 0.02 % NaN_3,pH 7.5, respectively. DNase I and lysozyme (to a final concentration of 10 mg.mL^{-1}) were added to the cell suspension and incubated at 4°C for 20 minutes before the cells were sonicated on ice.

Ampicillin resistance is conferred by the *Bla* gene of the pET-20(+) parental expression vector and chloramphenicol resistance stems from the pLysS plasmid which encodes the production an acetyl transferase and lysozyme. Lysozyme inhibits T7 RNA polymerase which keeps the target gene expression suppressed prior to IPTG induction.

The lysed cell suspensions were then centrifuged at 16 000 × *g* for 20 minutes at 4°C and samples of whole cell extract, soluble and insoluble fractions were electrophoresed on 12% acrylamide SDS-PAGE gels (section 2.5). Purification of the wild-type rGST M1-1 was archived by affinity chromatography as previously described (Hornby *et al.* 2000). The soluble fractions after centrifugation were loaded onto an *S*-hexylglutathione affinity column pre-equilibrated with 10 mM Tris-HCl, 1 mM EDTA, 200 mM NaCl, 0.02 % NaN_3, pH 7.8. After loading the column, 10 column volumes of equilibration buffer were passed through the column to remove any non-specifically bound protein. Bound wild-type rGST M1-1 was eluted off the column with 50 mM glycine-NaOH buffer, pH 10 (Cameron *et al.*, 1995). The eluent was collected in 4 mL fractions and the pH was neutralised by adding an equal volume of Mes-NaOH buffer, pH6.5 to all protein containing fractions.

The mutant protein was purified by cation exchange chromatography (Thompson *et al.* 2006). Protein containing fractions were pooled together and dialysed against 20 mM MOPS, 1mM EDTA, 1 mM DTT, 0.02 % NaN_3, pH 6.8 before being loaded onto an SP-Sepharose cation exchange column pre-equilibrated with the same buffer. The column was washed with equilibration buffer until the absorbance of column eluent at 280 nm was zero to ensure that all weakly interacting, unbound protein had been eluted. Bound F56S/R81A rGST M1 was eluted off the column with a linear salt gradient (0-300 mM NaCl) created by mixing the equilibration buffer with the elution buffer (20 mM MOPS buffer pH 6.8, containing 300 mM NaCl).

In all purification experiments, the presence of protein was detected by measuring the absorbance of the eluent, at 280 nm (A_{280}) in real-time, using the onboard spectrophotometer on the ÄKTAprime automated chromatographic system coupled to a computer with PrimeView 1.0 software (GE Healthcare Life Sciences, Uppsala, Sweden).

Purified wild-type rGST M1-1 and F56S/R81A mutant were dialysed against three (2 L) changes of protein storage buffer (20 mM sodium phosphate, 100 mM NaCl, 0.02% NaN_3, pH 6.5 for the wild-type protein; and 20 mM sodium phosphate, 1 mM DTT, 100 mM NaCl, 0.02% NaN_3, pH 6.5 for the mutant), lasting 3 hours each, after which it was snap frozen using liquid nitrogen and stored at -80ºC. Unless otherwise stated, all experiments for wild-type and mutant class Mu GST proteins were conducted in the corresponding storage buffer.

2.4. Protein Concentration Determination

The protein concentration for both protein forms (i.e. wild-type rGST M1-1 and the F56S/R81A variant) was calculated spectrophotometrically using the Beer-Lambert law:

$$A = \varepsilon_\lambda C l \qquad (4)$$

where A is the absorbance, ε_λ is the molar extinction coefficient ($M^{-1}cm^{-1}$) at a given wavelength (λ, which in the case of protein detection is at 280 nm), C is the molar concentration and *l* is the path length (cm).

A molar extinction coefficient of 40 060 $M^{-1}.cm^{-1}$ (Thompson *et al.*, 2006) was used to calculate the concentration of the F56S/R81A mutant, and for the wild-type protein, the

molar extinction coefficient at 280 nm was estimated using the method described by Perkins (1986):

$$\varepsilon(M^{-1}cm^{-1}) = 5\,500(\Sigma\text{Trp residues}) + 1\,340(\Sigma\text{Tyr residues}) + 150(\Sigma\text{Cys residues}) \quad (5)$$

where the values 5 500, 1 340 and 150 are the molecular extinction coefficients at 280 nm ($M^{-1}cm^{-1}$) of individual tryptophan, tyrosine and cysteine, respectively. Using this method, the extinction coefficient of the wild-type rGST M1-1 is calculated to be 81 480 $M^{-1}.cm^{-1}$ (Hornby *et al.*, 2000)

The absorbance at 280 nm was measured for a serially-diluted solution of each class mu rGST form and the concentrations were determined by fitting a linear regression to 10 points from a serial dilution. All absorbance readings were buffer corrected for by subtracting appropriate buffer reading used for the concentration determination. Substitution of the corrected absorbance, the extinction coefficient, and the pathlength (1 cm) into the Beer-Lambert law (see above) yields the molar concentration (M) of soluble protein.

2.5. Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis

Protein size, solubility, and purity was assessed using discontinuous SDS-PAGE according to the method described by Laemmli (1970). Protein samples were prepared at a 1:1 dilution with loading buffer (0.5 mM Tris-HCl, 20% (v/v) glycerol, 10% (w/v) SDS, 100 mM β-mercaptoethanol, 0.05 % bromophenol blue, pH 6.8) and incubated at 95 °C for 5 minutes before loading onto gel to ensure complete denaturation. The separating gel consisted of 15% (w/v) acrylamide: 0.4% bisacrylamide and the stacking gel consisted 5% acrylamide: 0.1% bisacrylamide. The protein samples (15 µL) were loaded onto the SDS-PAGE wells and electrophoresed at 140 V on a Hoefer MiniVE electrophoresis system (Holliston, MA, USA). The electrophoresis buffer used contained 25 mM Tris, 192 mM glycine and 1% (w/v) SDS, pH 8.5. The molecular weight marker used composed a mixture of seven proteins, namely, β-galactosidase (116.0 kDa), bovine serum albumin (66.0 kDa), ovalbumin (45.0 kDa), lactate dehydrogenase (35 kDa), restriction endonuclease Bsp98I (25.0 kDa), β-lactoglobulin (18.8 kDa) and lysozyme (14.4 kDa). The gels were stained for 2 hours in 2% (w/v)

Coomassie Blue R250 staining solution containing 13.5% (v/v) glacial acetic acid and 18.75% (v/v) ethanol and destained with 40% (v/v) ethanol and 10% (v/v) glacial acetic acid until the background was clear (overnight).

2.6. Size Exclusion High-Performance Liquid Chromatography

The hydrodynamic volume and oligomeric state of native wild-type rGST M1-1, and native F56S/R81A rGST M1, were determined using size exclusion high-performance liquid chromatography (SE-HPLC) in order to check whether the mutation resulted in a protein of lower molecular mass. Protein forms were loaded to the column at a final protein concentration of 20 µM F56S/R81A mutant and 10 µM rGST M1-1 (to achieve equal monomer concentrations). Protein samples were excited at a wavelength of 295 nm and elution from the column was monitored by following the fluorescence emission at 340 nm. All samples loaded had a final volume of 20 μL. A standard curve was constructed using the proteins thyroglobulin (from bovine; 670 kDa), γ-globulin (from bovine; 158 kDa), ovalbumin (from chicken; 44 kDa) and myoglobin (from horse; 17 kDa) for column calibration.

2.7. Enzyme Specific Activity

GSTs can catalyse the nucleophilic aromatic substitution reaction of CDNB and GSH, resulting in the formation of 1-(*S*-glutathionyl)-2,4-dintrobenzene, a compound which absorbs light maximally at 340 nm (Habig *et al.*, 1974; Habig and Jakoby, 1981). Thus, this reaction was used to determine the specific activity of the two class Mu GST forms.

Triplicate sample solutions of 1-7 nM wild-type rGST M1-1 protein and 25-800 nM F56S/R81A mutant were prepared in 0.1 M sodium phosphate, 1 mM EDTA, pH 6.5 (Habig and Jakoby, 1981). Enzyme activity was measured spectrophotometrically at 20 °C using a Jasco V–630 UV/Vis spectrophotometer by monitoring the formation of *S*-2,4-dinitrophenyl glutathione at 340 nm (Habig and Jakoby, 1981). GST enzyme was added to a 3 ml final volume solution with a final concentration of 1 mM GSH and 1mM CDNB (dissolved in ethanol). The ethanol concentration was held constant at 3% (v/v) in all reactions. The change in concentration of the 1-(*S*-glutathionyl) protein concentrations in the assay mix were chosen to ensure linear progress curves and the reaction was followed for 60 seconds. All reactions

were corrected for the non-enzymatic controls. The specific activity of the enzyme was then calculated using an extinction coefficient of 9 600 $M^{-1}.cm^{-1}$ for *S*-2,4-dinitrophenyl glutathione (Habig *et al.*, 1974). The specific activity (μmol. $min^{-1}.mg^{-1}$) was determined by linear regression as the slope of a plot between the initial velocity of complex formation (μmol.min^{-1}) versus protein amount (mg).

2.8. Spectroscopic Studies

2.8.1. Far-UV circular dichroism (CD)

Circular dichroism (CD) is a spectroscopic technique based on the principle that, optically active molecules interact differently with right- and left-handed polarised light. Like a vast majority of biological molecules (including proteins), circularly polarised light is chiral, that is, it exists in two non-superimposable forms which are mirror images of one another (Woody, 1995). Proteins have a number of optically active groups including the aromatic side chains, the polypeptide backbone and disulphide groups which have characteristic absorption bands at specific areas in the ultraviolet and visible region of the electromagnetic spectrum (Woody, 1995). Disulphide groups and aromatic amino acids have characteristic absorption bands in the near-UV range (250 - 300 nm). On the other hand, far-UV (190-250 nm) CD spectra of proteins are primarily influenced by the amide interactions in the polypeptide backbone (Johnson, 1990), with different secondary structural elements generating distinct CD spectra. Far–UV CD was thus used in this study to compare the secondary structures of wild-type rGST M1-1 and the F56S/R81A rGST M1 proteins, thus giving an indication as to whether the mutation has any impact on the global structural integrity of the subunits.

Far-UV CD spectra (180 - 250 nm) were recorded for wild-type rGST M1-1 and the F56S/R81A mutant protein prepared in 5 µM sodium phosphate buffer, pH 6.5. All far-UV CD spectra were recorded on a Jasco J-810 spectropolarimeter with Spectra Manager software v1.5.00 (Jasco Inc., Tokyo, Japan) at 20°C and each spectrum represent an average of 10 accumulations, at a scan speed of 200 $nm.min^{-1}$. The bandwidth was set to 1 nm and the data pitch 0.2 nm using a quartz cuvette with a pathlength of 2 mm. All spectra were buffer corrected. The spectra were normalised by calculating the mean residue ellipticity [θ] $deg.cm^2.dmol^{-1}.residue^{-1}$ using the following equation (Woody, 1995):

$$[\theta] = 100(S)/C.n.l \quad (6)$$

where S denotes the CD signal averaged over 10 accumulations in millidegrees after subtraction of the solvent baseline, C is the protein concentration in mM, n is the number of residues, and *l* is the path length in cm.

2.8.2. Tryptophan fluorescence

Fluorescence refers to the emission of radiation energy from a molecule which results from the return of an unpaired electron from the excited to the ground state (Lakowicz, 1999). There are three naturally occurring fluorophores in protein molecules, namely, tryptophan, tyrosine and phenylalanine residues. The fluorescence emission of phenylalanine residues in proteins is rarely observed as a result of this residue having a low quantum yield (Lakowicz, 1999). The quantum yield of tryptophan is reported to be more than twice that of tyrosine (Lakowicz, 1999). This is because, in the native state of most proteins, tyrosine emission is quenched as a result of fluorescence resonance energy transfer (FRET) to tryptophan nearby residues, resulting in tryptophan emission being the predominant contributor of protein fluorescence.

Since the fluorescence of the indole ring side chain of tryptophan residues is sensitive to their microenvironment, intrinsic tryptophan fluorescence is a useful technique of measuring the local environment of the tryptophan residues in the protein (Lakowicz, 1999). The location of tryptophan residues with in a protein molecule provides a useful means of probing for local or global conformational changes of a protein. In this study selective tryptophan fluorescence was used as a probe to compare global structural differences between wild-type rGST M1-1 and the F56S/R81A rGST M1 mutant.

Protein sample solutions of native and unfolded (in 8.0 M urea) 2 μM wild-type rGST M1-1 and 4 μM F56S/R81A rGST M1 protein were prepared in 20 mM sodium phosphate buffer, pH 6.5, containing 100 mM NaCl and 0.02% NaN_3 (and 1 mM DTT in the case of the mutant protein). Different concentrations of the two protein forms were used to account for the number of tryptophan residues in the monomeric mutant to those of the dimeric wild-type protein. Experiments were conducted at 20 °C, with samples being irradiated with planar polarised light at the wavelength 295 nm to selectively excite the tryptophan residues. Both excitation and emission bandwidth filters were set to 5 nm and a quartz cuvette with a

pathlength of 1 cm was used in all experiments. The intrinsic fluorescence emission was monitored between 290 and 500 nm using a scan speed of a 100 $nm.min^{-1}$. The obtained spectrum for each sample represents an average of three scan accumulates per sample. All contributions to the fluorescence spectra from the buffer solution were accounted for by subtracting the spectrum collected for the appropriate blank buffer solution.

2.8.3. Extrinsic fluorescence spectroscopy – ANS binding

ANS is an anionic dye which can be used as a probe for the detection of non-polar surfaces on proteins, since it binds to hydrophobic patches in proteins (Semisotnov *et al.*, 1991; Engelhard and Evans, 1995). In an aqueous environment ANS fluorescence is quenched by the solvent molecules, however, when bound to a hydrophobic surface or exposed to a hydrophobic environment its fluorescence emission maximum shifts to a lower wavelength (blue-shift) and there is an increase in its quantum yield (Engelhard and Evans, 1995).

A 20 mM ANS stock solution was prepared in F56S/R81A storage buffer. The concentration of ANS was confirmed using the Beer-Lambert law by recording the absorbance at 350 nm and using an extinction coefficient of 5000 $M^{-1}.cm^{-1}$ (Weber and Young, 1964). The ANS fluorescence spectra, in the absence and presence of 2 µM rGST M1-1 and 4 µM F56S/R81A proteins, were recorded at 20°C, buffer corrected, and are an average of three accumulations at a scan speed of 100 $nm.min^{-1}$ .For the urea induced unfolding studies, 4 µM of the F56S/R81 mutant was incubated for at least 2 hours, at different concentrations of urea (0 - 8 M) to archive equilibrium before the addition of ANS (to the protein-urea mixture) to a final concentration of 200 µM. The solution was incubated for at least an hour to achieve equilibrium. The samples were excited at 390 nm and emission spectra were recorded from 390 to 600 nm. Spectra were produced from an average of three accumulations at 300 $nm.min^{-1}$ scan speed. The excitation and emission slit widths were at 5 nm.

2.9. Conformational Stability

2.9.1. Recovery of native F56S/R81A

To derive thermodynamic parameters for the protein unfolding process, the reversibility of unfolding needs to be established (Pace, 1986). Establishing the extent by which the

unfolding process is reversible gives an indication as to whether equilibrium between the native and the denatured state exists or not. In this study, this was achieved by measuring the recovery of folded F56S/R81A rGST M1 using intrinsic tryptophan fluorescence as a tertiary structural probe. F56S/R81A rGST M1 (4 µM) was unfolded in the presence of 8 M urea and allowed to refold by diluting this protein-urea mixture 10-folds with protein storage buffer. The fluorescence spectrum of the resultant solution and of a 0.4 µM F56S/R81A in the presence of 0.8 M urea was then recorded (Section 2.9.2)

Urea used in all experiments was prepared according to a method described by Pace (1986) using F56S/R81A storage buffer as the solvent. Following preparation, the pH of the urea stock solution was adjusted to pH 6.5 before filtering the solution and having its concentration (10 M) confirmed through reading of its refractive index using an Atago R5000 refractometer (Tokyo, Japan) and using the equation formulated by Pace (1986):

$$[\text{Urea}] = 117.66(\Delta N) + 29.753(\Delta N)^2 + 18.56(\Delta N)^3 \quad (7)$$

where [Urea] is the molar concentration of urea and ΔN is the difference in refractive index between the buffered urea solution and the buffer solvent.

2.9.2. Urea-induced equilibrium unfolding.

Chemical-induced protein denaturation curves serve as a convenient method for estimating the conformational stability of a protein. In the present study, the effects of the engineered mutations on the stability of wild-type rGST M1-1 (i.e. the F56S/R81A mutant) were investigated. Using the chaotropic agent, urea, as a chemical denaturant, the conformational stability of the F56S/R81A mutant was determined by setting up a series of equilibrium unfolding reactions (0 - 8 M urea). In order to enhance detection of any intermediates in the unfolding process, it is ideal to use multiple probes to monitor the entire unfolding pathway of a protein (Neet and Timm, 1994). Far-UV CD (Section 2.8.1), intrinsic tryptophan fluorescence (Section 2.8.2) and ANS binding (Section 2.8.3) were used in this study to assess the unfolding process and subsequently the conformational stability of the F56S/R81A rGST M1.

2.10. Data Fitting

When a protein's unfolding process is reversible and occurs via a 2-state mechanism, a state of equilibrium is reached between the folded native (N) and the unfolded (U) species. Thus, only the native and the unfolded states are present at significant concentration levels (Pace, 1986):

$$N \Leftrightarrow U \tag{8}$$

Therefore, it follows that for a 2-state unfolding mechanism:

$$f_N + f_U = 1 \tag{9}$$

where f_N and f_U represent the fraction of the folded native protein species and that of the unfolded species, respectively. The obtained signal along any point of the unfolding process is contributed toward by the signal from the concentration of both species:

$$y = y_N f_N + y_U f_U \tag{10}$$

where y is the signal obtained from a specific spectroscopic probe. y_N represents the y value for the native state and can be extrapolated from linear pre-transition region of the unfolding data, whereas y_U represents the y value for the unfolded state and can be extrapolated from linear post-transition region of the unfolding data. Combining equation (3) and (4) yields:

$$f_U = (y_N - y)/(y_N - y_U) \tag{11}$$

The fraction of folded protein species can also be obtained in a similar manner:

$$f_N = (y - y_U)/(y_N - y_U) \tag{12}$$

the equilibrium constant (K_{eq}) for the unfolding process is expressed as:

$$K_{eq} = f_U/f_N \tag{13}$$

Thus, if equation (5) and (6) are substituted into equation (7), we get:

$$K_{eq} = (y_N - y)/(y - y_U) \tag{14}$$

and;

$$\Delta G = -RT \ln K_{eq} \tag{15}$$

where ΔG is the free energy of unfolding, R is the universal gas constant, and T is temperature in Kelvin. If it is assumed that ΔG has a linear dependence on the concentration of the denaturant [D] throughout the entire denaturant concentration range, the free energy difference between the folded and unfolded states, in the absence of denaturant ($\Delta G(H_2O)$), can be determined (Tanford, 1968). Thus;

$$\Delta G = \Delta G(H_2O) - m[D] \tag{16}$$

where m (m-value) represents the dependence of free energy on denaturant concentration which is also an indicator of co-operativity and can be related to the change in solvent accessible surface area (ΔASA).

The equilibrium unfolding data obtained in this study for the monomeric F56S/R81A mutant protein were fitted to equation (9) using SigmaPlot version 11.0. The thermodynamic parameters, $\Delta G(H_2O)$ and the *m*-value, were subsequently determined.

CHAPTER 3

RESULTS

3.1. cDNA Sequence Verification.

DNA sequencing results, from Inqaba Biotech (Pretoria, South Africa), using the T7 primers verified both the wild-type rGST M1-1 and the F56S/R81A mutant cDNA sequence inserts within the pET-20b(+) plasmid (Figure 3-1). The resultant rGST M1-1sequence showed 100% identity to wild-type rGST M1-1 sequence. Comparison of the mutant sequence to the wild-type sequence confirmed that the TTC codon, which codes for a phenylalanine residue at position 56 of the amino acid sequence of the wild-type enzyme, to be mutated to TCC which codes for a serine residue in the mutant sequence. Similarly, the CGC codon, which codes for an arginine residue at position 81 of the wild-type enzyme's amino acid sequence, is found to be mutated to GCC, which codes for alanine, in the mutant sequence (Figure 3-1).

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

The recombinant expression system for wild-type rGST M1-1 and the F56S/R81A mutant comprised *E. coli* BL21 (DE3) pLysS and *E. coli* T7 Express *Iq* Competent cells, respectively. The expression system for the F56S/R81A mutant protein initially contained the *E. coli* BL21 (DE3) pLysS system, however, the mutant protein was found not to be soluble at growth temperatures of 37°C and 20°C, in addition the mutant protein was found to be expressed at significantly higher levels in the T7 system (Figure 3-2), thus, *E. coli* T7 Express *Iq* Competent cells were used instead.

E. coli T7 Express *Iq* Competent cells are a derivative of the BL21 (DE3) pLysS cells. The difference between *E. coli* BL21 (DE3) pLysS and T7 Express *Iq* cell strains stems from the location of the T7 RNA polymerase gene. While the BL21 (DE3) pLysS cells carry the T7 RNA polymerase gene on a lysogenic prophage, the T7 Express *I*q cells have the gene

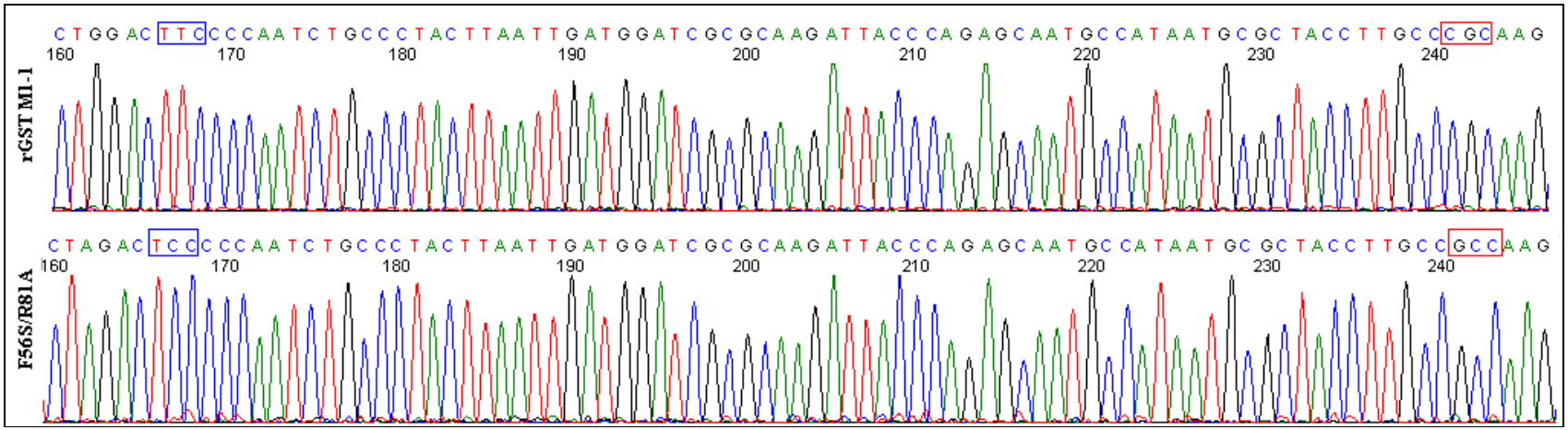


Figure 3-1: A portion of the nucleotide sequencing chromatograms for the rGST M1-1 cDNA ORF indicating the codon mutations responsible for the corresponding F56S/R81A rGST M1-1 double mutant protein sequence.

Shown here are the nucleic acid sequence of wild-type rGST M1-1 and the F56S/R81A sequence. The codon, TTC (blue box on wild-type sequence), which codes for Phe at position 56 of the amino acid sequence was mutated to TCC (blue box on mutant sequence), which codes for Ser. Similarly, CGC (red box on wild-type sequence) was mutated to GCC (red box on mutant sequence), which changes Arg to Ala at position 81of the amino acid sequence. The sequencing results were viewed using the program Finch TV version 1.4.0 (http://www.geospiza.com/FinchTV: Geospiza Inc.).

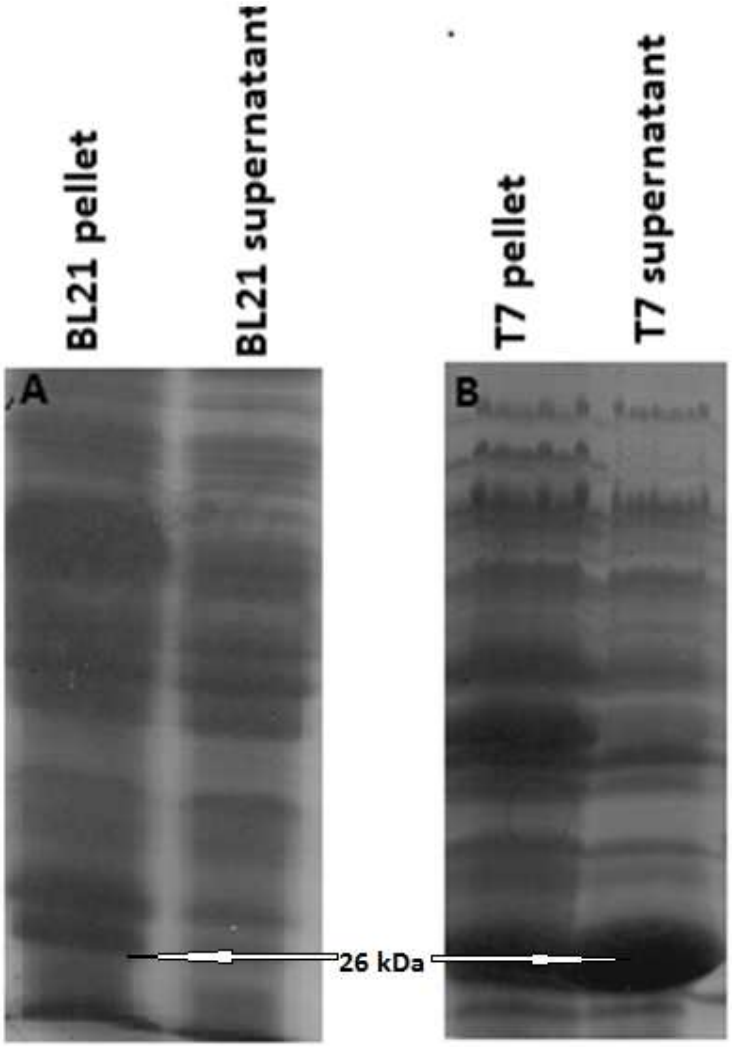


Figure 3-2: SDS-PAGE analysis of the F56S/R81A mutant protein overexpression.

(A) Expression of the F56S/R81A rGST M1 protein in *E. coli* BL21 (DE3) pLys cells. (B) Expression of the F56S/R81A rGST M1 protein in *E. coli* T7 Express *Iq* competent cells. The arrow indicates the expected molecular mass of the target protein.

inserted into the lac operon of the *E. coli* chromosome and is expressed under the control of the lac promoter. Thus, the T7 Express *I*q cells provided controlled induction of the polymerase and subsequently, inducible control of the transcription of genes downstream of the T7 promoter thereby providing an advantage over the BL21 (DE3) pLysS strain.

The appropriate *E. coli* cell strains were transformed with the pET-20b(+) plasmid containing the open reading frame encoding the appropriate proteins. The purification of recombinant wild-type rGST M1-1 and the F56S/R81A mutant was archived by *S*-hexylglutathione affinity chromatography (Hornby *et al.*, 2000) and SP-Sepharose cation exchange chromatography (Thompson et al., 2006), respectively, and both proteins were found to elute as single peaks from their respective columns (Figure 3-3). Protein purity was assessed by SDS-PAGE (Figure 3-4A). SDS-PAGE analysis shows both wild-type rGST M1-1 and the F56S/R81A mutant to migrate as a single band on the gel (Figure 3-4A) which corresponds to an apparent subunit molecular mass of approximately 27 kDa (Figure 3-4B).

3.3. Protein size determination

SE-HPLC was used to investigate the oligomeric state of wild-type rGST M1-1 and the F56S/R81A mutant (Figure 3-5). SE-HPLC analysis revealed protein samples to have some contaminants. The F56S/R81A rGST mutant elution profile (Figure 3-5B) shows two small peaks towards the end indicating that the contaminants are smaller in size than the target protein, whereas, on the other hand, the elution profile of wild-type enzyme (Figure 3-5A) shows a small peak just before the target protein elute, indicating that the contaminant has a lager molecular mass than target protein. The apparent molecular mass for wild-type rGST M1-1 and F56S/R81A rGST M1 is calculated from the calibration curve of low molecular mass standard proteins (see Section 2.6.) using linear regression analysis. The relative molecular mass of wild-type rGST M1-1 was estimated to be 53 kDa (Figure 3-5C) which compares well with previously published data (Ji *et al.*, 1992), and thus confirms the dimeric status of the wild-type enzyme.

The relative molecular mass of the F56S/R81A rGST M1-1 mutant was estimated to be aproximately 24 kDa, almost half that of the estimated wild-type enzymes molecular mass.

These findings suggest that disruption of intersubunit motifs in class Mu GST results in a stable monomeric species.

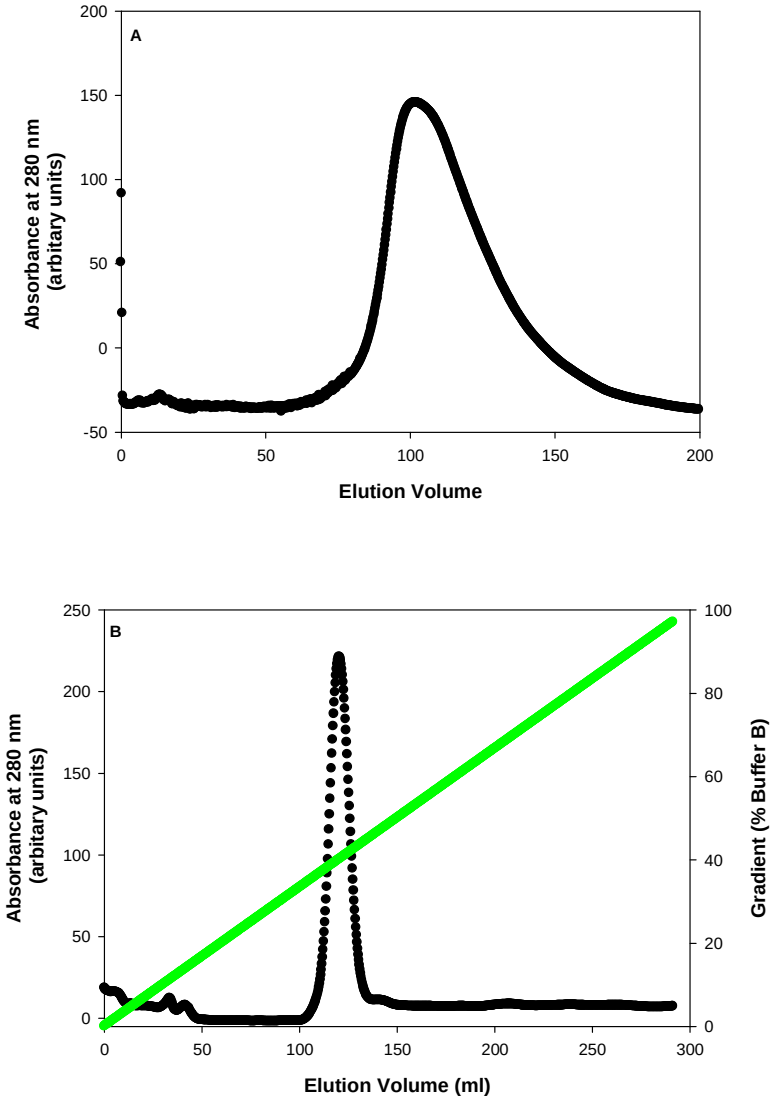


Figure 3-3: Elution profiles of wild-type rGST M1-1 and F56S/R81A rGST M1.

(A) Elution profile of wild-type rGST M1-1 purified by *S*-hexylglutathione affinity chromatography. Bound rGST M1-1 was eluted from the column with 50 mM glycine-NaOH buffer, pH 10. (B) Elution profile of F56S/R81A rGST M1 purified by SP-Sepharose cation exchange chromatography. Bound F56S/R81A was eluted from the column using a 0 - 300 mM NaCl gradient (buffer B, green line).

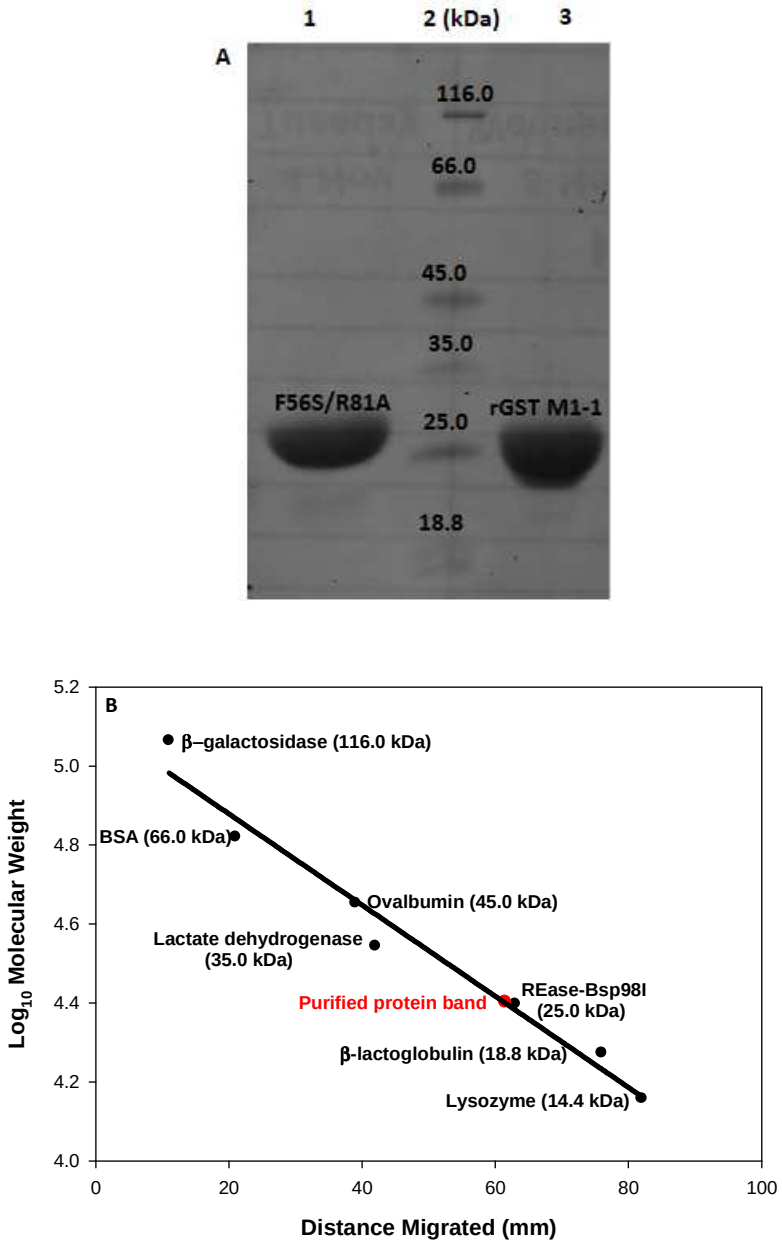


Figure 3-4: SDS-PAGE analysis of purified proteins.

(A) Lane 2 represents the molecular weight markers: β-galactosidase (116.0 kDa), bovine serum albumin (66.0 kDa), ovalbumin (45.0 kDa), lactate dehydrogenase (35.0 kDa), restriction endonuclease Bsp981 (25.0 kDa), β-lactoglobulin (18.8 kDa) and lysozyme (14.4 kDa). Lane 1 and 3 represents the F56S/R81A mutant and wild-type GST M1-1, respectively. SDS-PAGE analysis shows both proteins to be pure. (B) SDS-PAGE molecular weight marker calibration curve. The subunit molecular weight for both class mu GST forms is ~27 kDa (red). The correlation coefficient is 0.97.

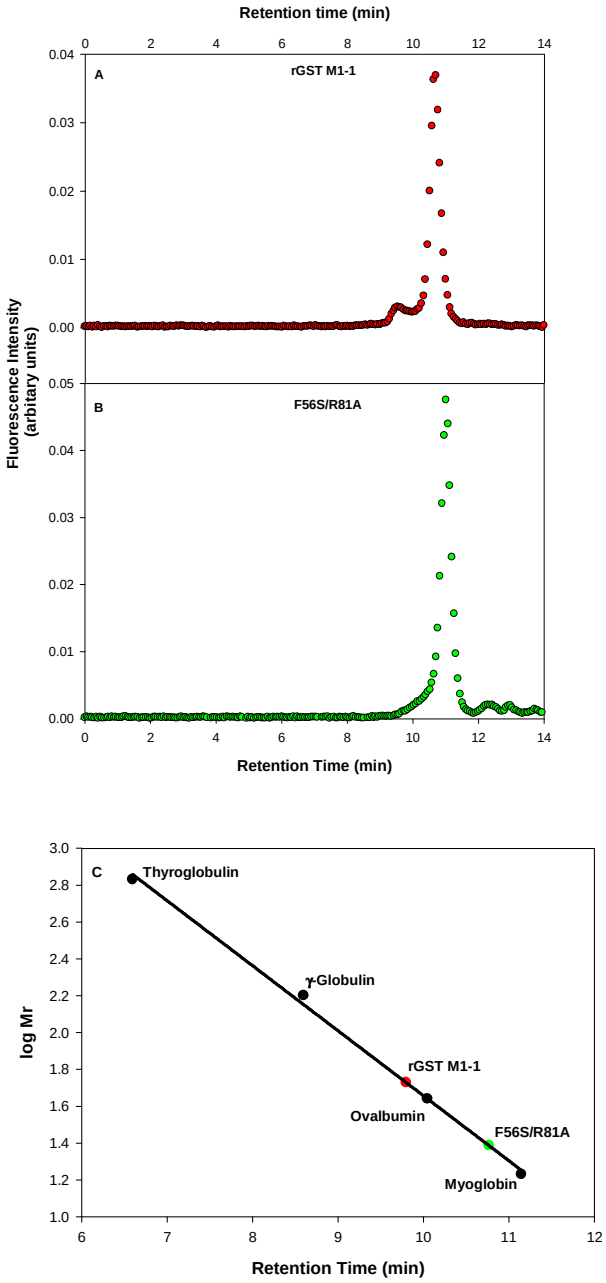


Figure 3-5: SE-HPLC elution profile of rGST M1-1 and F56S/R81A monomeric mutant.

(A) and **(B)** Elution of the proteins was monitored by fluorescence with excitation set at 295 nm and fluorescence emission monitored at 340 nm for both protein forms. The flow rate was 1 ml.min^{-1} at an isocratic pressure of 61 bar. The retention times for wild-type rGST M1-1 and F56S/R81A are 10.5 min and 11.02 min, respectively. **(C)** Calibration curve of protein standards: thyroglobulin (670 kDa); γ-globulin (158 kDa); ovalbumin (44 kDa) and myoglobin (17 kDa). Red and green filled circles indicate the retention times of wild-type rGST M1-1 and the F56S/R81A mutant, respectively. The correlation coefficient is 0.99.

3.4. Functional and Structural Characterisation

3.4.1. Enzymatic activity

Cytosolic GSTs catalyse the nucleophilic aromatic substitution reaction of CDNB and GSH resulting in the formation of 1-(*S*-glutathionyl)-2,4-dintrobenzene (also known as *S*-2,4-dinitrophenyl glutathione) which absorbs light maximally at 340 nm (Habig and Jakoby, 1981). The CDNB nucleophilic aromatic substitution reaction occurs through the formation of 1-chloro-1-(S-glutathionyl)-2,4-dinitrocyclohexadienate (Meisenheimer complex transition state) (Graminski *et al.*,1989). This Meisenheimer complex transiton state is stabilised at the GST active site (which is formed partially at the dimer interface) for efficient nucleophilic aromatic substitution reaction (Bico *et al.*, 1994; Graminski *et al.*, 1989). The enzyme activity of rGST M1-1 (Figure 3-6A) and the F56S/R81A mutant (Figure 3-6B) was measured spectrophotometrically at 340 nm by monitoring the formation of *S*-2,4- dinitrophenyl glutathione as a function of amount of enzyme (Habig and Jakoby, 1981). The specific activity of wild-type rGST M1-1 and the F56S/R81 mutant was determined from the slope of a linear fit through the experimental data shown in Figure 3-6, and is calculated to be 1.2 $\mu mol.min^{-1}.mg^{-1}$ of protein and 0.03 $\mu mol.min^{-1}.mg^{-1}$ of protein, respectively, indicating that the F56S/R81A mutant displays only 2.7 % of the dimeric wild-type enzymes activity.

3.4.2. Spectroscopic properties of rGST M1-1 and rGST M1.

3.4.2.1. Far-UV circular dichroism

Far-UV CD spectra of proteins serve as an ideal probe for the secondary structural elements content of proteins. The CD spectra of proteins in the far-UV region (190-250 nm) have features that are distinctive of the secondary structural elements adopted by the protein since the spectra are influenced by amide interactions in the polypeptide backbone (Woody, 1995). Predominantly α-helical proteins display characteristic minima at 208 and 222 nm, as well as a stronger positive band near 190 nm (Woody, 1995).

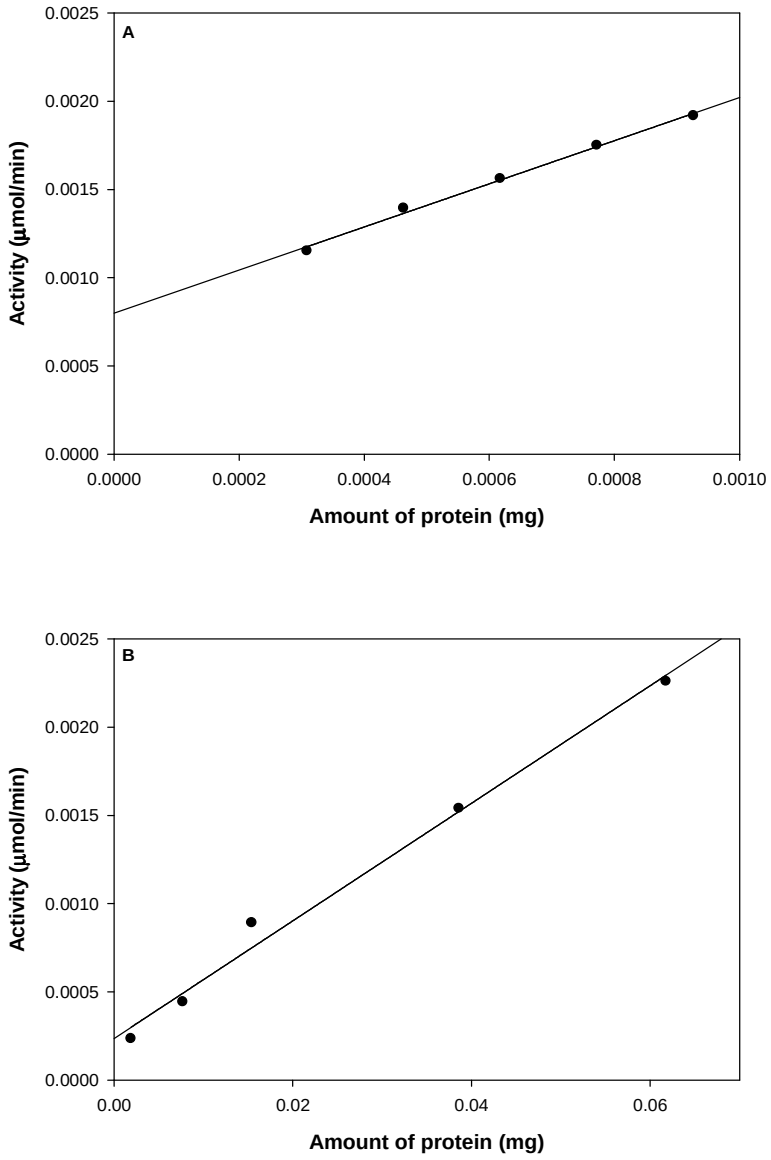


Figure 3-6: Specific activity of wild-type rGST M1-1 and F56S/R81A mutant.

Represented in A is the enzyme activity data for rGST M1-1 while B represents data for the F56S/R81 mutant. The solid lines represent a linear fit to the experimental data. The specific activity for each enzyme form is determined from the slope of the fit. The correlation coefficient is 0.99 for both fits.

Far-UV CD spectra of the wild-type rGST M1-1 and F56S/R81A in their native conformations are shown in Figure 3-7. The CD spectra of both wild-type and mutant class mu GST protein forms show two minima around 208 nm and 222 nm, which is characteristic for predominantly α-helical proteins (Woody, 1995). Such spectra have been reported for other GST classes including Alpha (Wallace *et al*., 1998) and Pi (Dirr and Reinemer, 1991) and Sigma (Stevens *et al*., 1998). In certain cases, it is impossible to record clear spectra below 210 nm as a result of noise contributions by buffers containing high salt concenrations which, in turn, lead to reduced signal to noise ratio. The far-UV CD spectrum of native F56S/R81A differs from that of the wild-type enzyme, exhibiting reduced intensity bands at both minima, that is, the variant protein's CD signal is higher than that of the wild-type protein (Figure 3-7). This observation suggests that the mutation results in the disruption of secondary structural elements. Secondary structural modification can occur upon the assembly or disassembly of a quaternary structure (Hennessey *et al.*, 1982). Therefore, the prevention of protein association by disrupting the two main dimer association motifs in rGST M1-1, may account for the observed difference in CD signals of the mutant relative to the wild-type protein. In addition, amide H/D exchange studies carried out by Thompson *et al.*(2006) on the F56S and F56S/R81A variants of rGST M1-1, indicate a change in structure and dynamics at the junction between α4 and α5 helices.

3.4.2.2. Intrinsic tryptophan fluorescence

The fluorescence of the indole ring side chain of tryptophan residues is sensitive to the polarity of their microenvironment, thus making intrinsic tryptophan fluorescence a useful technique for measuring the local environment of tryptophan residues in a protein (Lakowicz, 1999). The positioning of tryptophan residues within the protein molecule provides a useful means of probing for local or global conformational changes of a protein.

Each subunit of rGST M1-1 contains four tryptophan residues. Two of these tryptophan residues are located in the domain 1 at positions 7 and 45, whereas the other two are located in the C-terminal domain at positions 146 and 214 (Ji *et al.*, 1992). 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 (Ji *et al*., 1992). Thus, intrinsic tryptophan fluorescence can be used to probe the tertiary structural changes in wild-type rGST M1-1 and the F56S/R81 mutant protein. Shown in Figure 3-8 are the fluorescence emission spectra of the

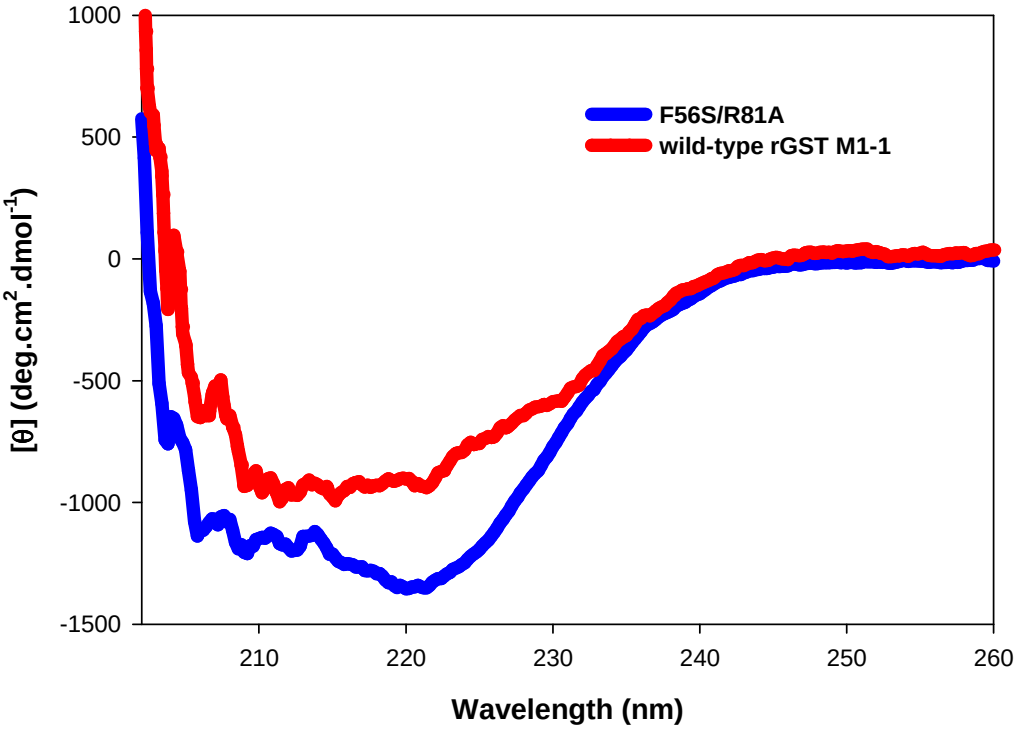


Figure 3-7: Far-UV circular dichroism spectra of native wild-type rGST M1-1 and F56S/R81A rGST M1.

Spectra were collected using final protein concentrations of 2µM and 4µM for rGST M1-1 and rGST M1, respectively, in 5 mM sodium phosphate buffer, pH 6.5. Each spectrum represents an average of 20 accumulations.

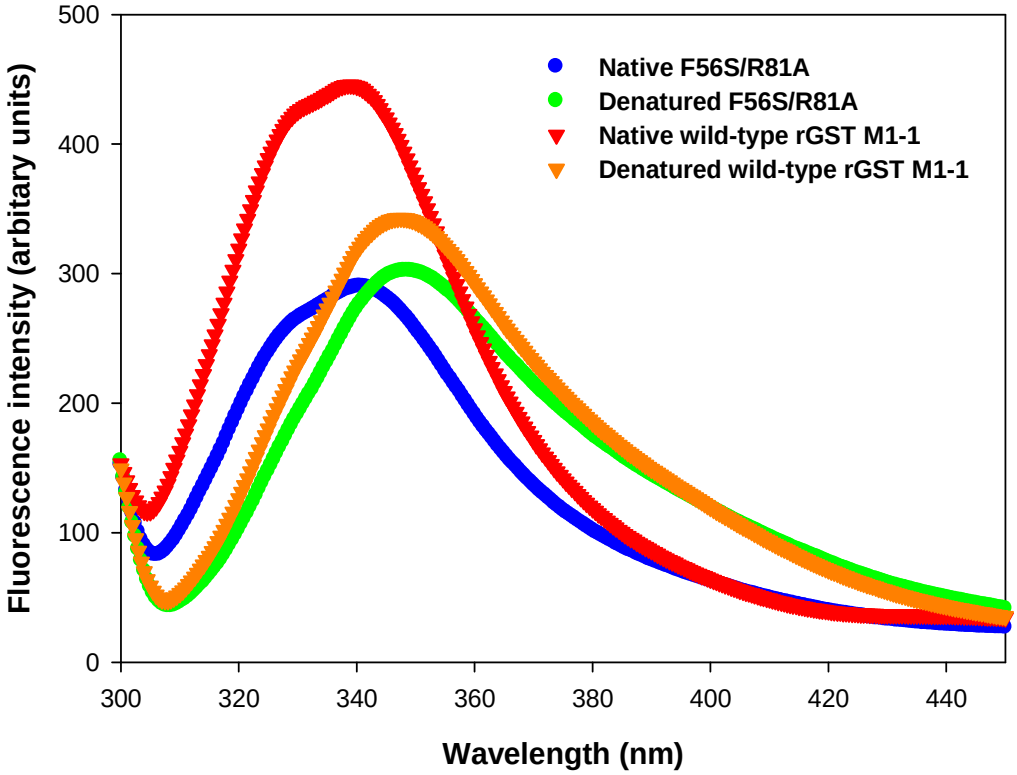


Figure 3-8: Fluorescence emission spectra of wild-type GST M1-1 and F56S/R81A monomeric mutant.

Tryptophan residues were selectively excited at 295 nm and the maximum emission wavelength for native 2 µM wild-type GST M1-1 and 4 µM F56S/R81A are shown to be 339 nm and 340 nm, respectively. Experiments were performed in 20 mM sodium phosphate buffer, pH 6.5. For the urea induced (8 M urea) unfolded GST M1-1 and F56S/R81A the maximum emission wavelengths are 347 nm and 348 nm, respectively.

native and unfolded species of wild-type rGST M1-1 and the F56S/R81A mutant generated by selective excitation of the tryptophan residues at 295 nm. The fluorescence emission maxima for the native conformations of the dimeric wild-type protein and the mutant are 339 nm and 340 nm, respectively (Figure 3-8), which suggests that the tryptophan residues in the native state of the two class mu GST protein forms are located in a hydrophobic environment. The mutation seems to have resulted in the quenching of tryptophan fluorescence as the mutant protein displays significant reduction of fluorescence intensity in the native state (Figure 3-8). Unfolded wild-type and F56S/R81A GST M1 exhibit red-shifted emission maxima at 347 nm and 348 nm, respectively (Figure 3-8).

3.4.2.3. ANS binding studies.

ANS is an amphipathic anionic dye which can be used as a probe for the detection of non-polar surface on proteins (Abdalla *et al.*, 2002). The fluorescence properties of this molecule are largely influenced by the polarity of its surrounding environment. The fluorescence emission maximum of ANS shifts to a lower wavelength when bound to hydrophobic surface, and this bathochromic shift is accompanied by an increase in fluorescence intensity (Slavik, 1982). Thus ANS was used in this study as a probe to measure the overall surface hydrophobicity of the mutant protein relative to the wild-type. Figure 3-9 shows the fluorescence emission spectra of ANS bound to wild-type rGST M1-1 and the F56S/R81A monomeric mutant. ANS is observed to have an emission maximum at 518 nm in the absence of protein (Figure 3-9). A blue-shift in ANS fluorescence emission maximum is observed in the presence of both wild-type and mutant protein. The emission maxima of ANS in the presence of native rGST M1-1 and F56S/R81A mutant are 486 nm and 490 nm, respectively (Figure 3-9). The results also show an increase in ANS fluorescence intensity when ANS is excited in the presence of protein, which indicates that there has been a change in the microenvironment of ANS, suggesting that ANS binds the two protein forms. In addition, the fluorescence intensity of rGST M1-1 bound ANS is lower than that of ANS bound to the F56S/R81A mutant (Figure 3-9).

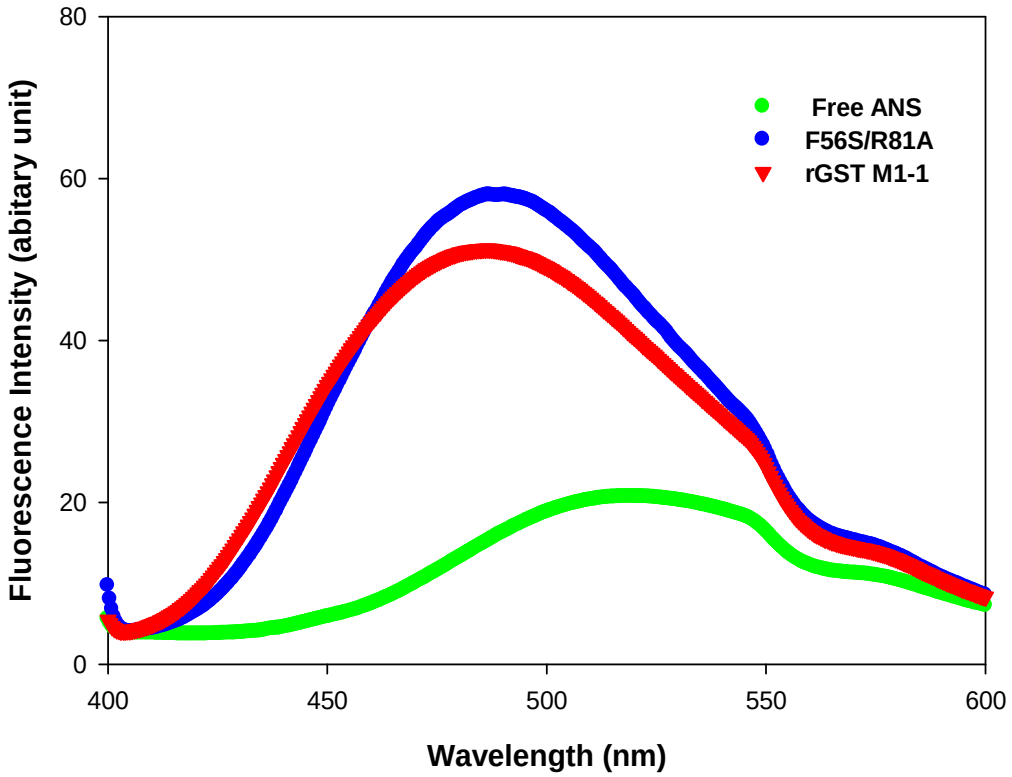


Figure 3-9: Fluorescence emission spectra of ANS bound to wild-type rGST M1-1 and the F56S/R81A rGST M1 mutant.

Experimental data was obtained by incubating 2 μM GST M1-1 and 4 μM GST M1 in 20 mM sodium phosphate buffer, pH 6.5 with ANS at a final concentration of 200 μM. ANS was then selectively excited at 390 nm, and emission spectra were measured from 395 to 600 nm. The maximum emission wavelengths for ANS bound to GST M1-1 and GST M1 were 486 nm and 490 nm, respectively.

3.5. Equilibrium Unfolding of the F56S/R81A rGST M1 Monomeric Mutant.

3.5.1. Recovery of native structure.

The reversibility of protein unfolding process and the recovery of the native fold of a protein need to be established in order to allow for the determination of the thermodynamic parameters (Pace, 1986). The recovery of urea induced denatured F56S/R81A was established using intrinsic tryptophan fluorescence as a tertiary structural probe (Figure 3-10). Following a ten-fold dilution of 4 µM rGST M1 unfolded in 8 M urea to 0.8 M urea, the protein is observed to regain approximately 81 % of its native structure, suggesting that the unfolded protein can refold and regain the same structure as that of the native protein in 0.8 M urea (Figure 3-10). Therefore, urea-induced equilibrium unfolding transitions of F56S/R81A can be analysed to determine the thermodynamic parameters of the unfolding process which will serve as an indication of the mutants conformational stability

3.5.2. Urea-induced unfolding of the F56S/R81A mutant protein.

Chemical induced denaturation curves of a protein serve as a convenient method for estimating the conformational stability of a protein. Using the chaotropic agent, urea, as a chemical denaturant, the conformational stability of the F56S/R81A mutant was determined by setting up a series of equilibrium unfolding reactions (0 - 8 M urea). Each unfolding reaction sample was allowed to reach equilibrium by incubation at 20 °C for at least 2 hours. The presence of aggregates along an unfolding process affects the equilibrium state of the system, that is, the process is no longer considered to be reversible. Thus, in order to determine the thermodynamic parameters of the unfolding process the effects associated with protein aggregation need to be minimized. To detect protein aggregation, Rayleigh scatter of 4 µM F56S/R81A was collected at a range of urea concentrations by setting the excitation wavelength at 295 nm and recoding the emission at the same wavelength (Figure 3-11B). The protein is shown not to form aggregates throughout the entire urea concentration range (0 - 8 M urea). The unfolding process was then followed by far-UV circular dichroism and intrinsic

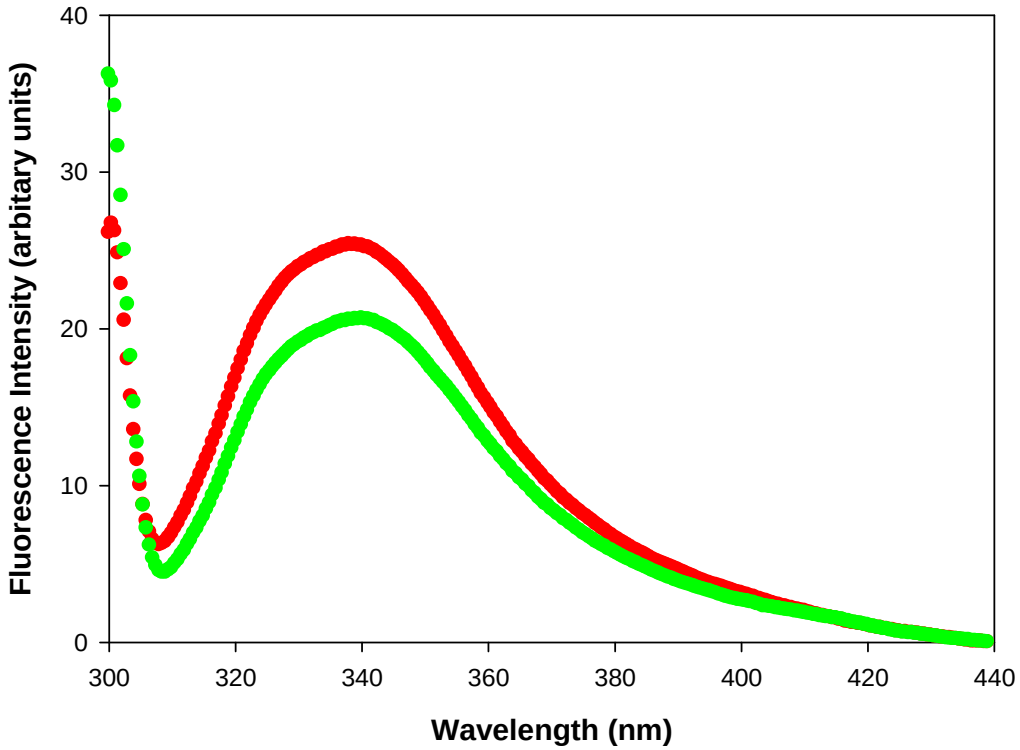


Figure 3-10: Recovery of the F56S/R81A rGST M1.

Shown here is the fluorescence emission spectrum of native (red) and refolded (green) F56S/R81A generated by exciting samples at 295nm. The maximum emission wavelength is 340 nm for the native and refolded protein. The residual concentration of urea for the refolded form was 0.8 M.

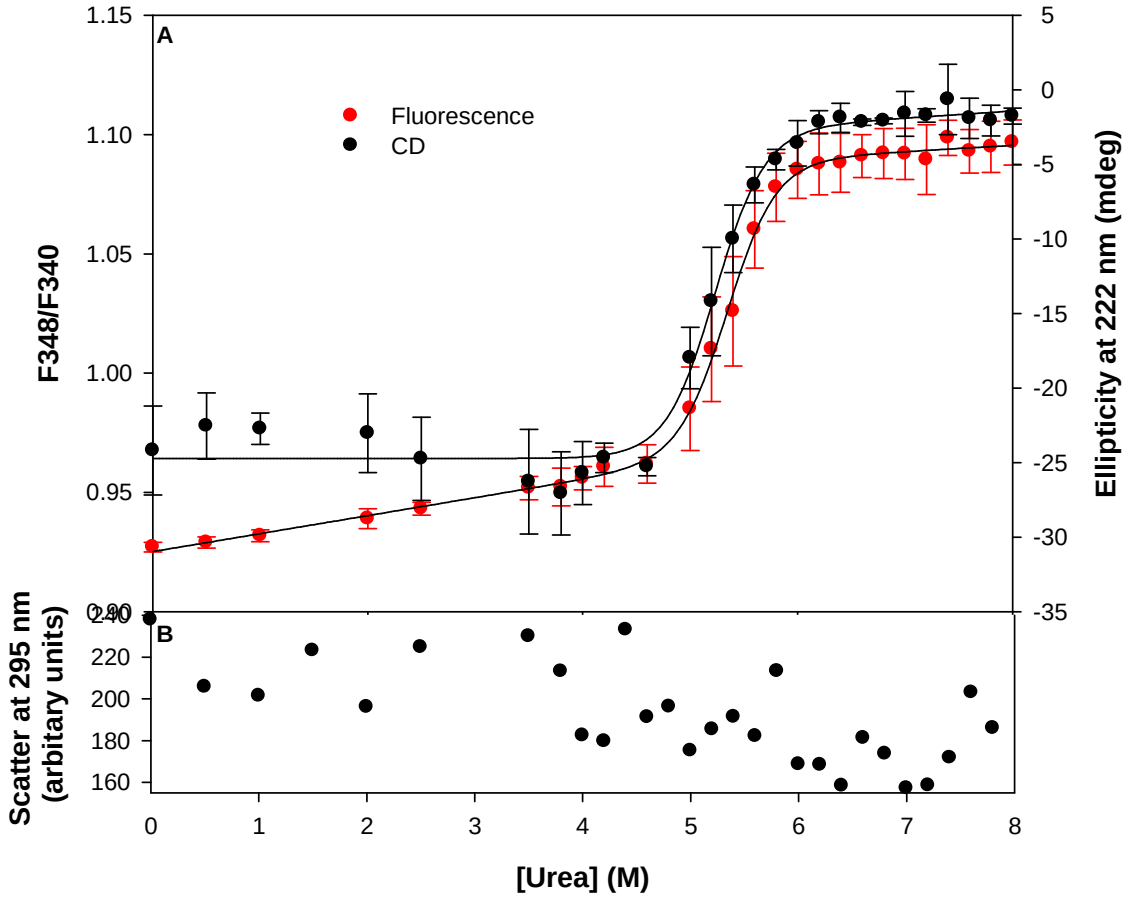


Figure 3-11: Urea-induced equilibrium unfolding of F56S/R81A.

(A) Urea-induced unfolding curve of 4 µM F56S/R81A mutant monitored by tryptophan fluorescence and far-UV circular dichroism**.** Experiments were performed in 20 mM sodium phosphate buffer, pH6.5. Tryptophan residues were selectively excited at 295 nm. Experimental data was fitted according to a 2-state model (solid lines). (B) Light scatter of 4 µM F56S/R81A rGST M1 as a function of urea concentration**.** Light scatter was measured by setting both the excitation and emission wavelengths to 295 nm. All experiments were performed in 20 mM sodium phosphate buffer, pH 6.5.

tryptophan fluorescence of the reaction samples (Figure 3-11A). For the fluorescence data, the ratio of the emission intensity at 348 nm to 340 nm is plotted for each urea concentration. The fluorescence and CD transition curves are found to be monophasic and sigmoidal, suggesting a two-state transition model (N⇔U), where N represents the native conformation, and U the unfolded one. The transition region is found to be in the 4.2 - 6 M urea concentration range for both fluorescence and CD data (Figure 3-11A) .The thermodynamic parameters were calculated by fitting the data to a two-state model for a monomer, both CD and fluorescence parameters are in agreement with one another (Table 3-1).

3.5.3. Urea-induced unfolding in the presence of ANS.

The binding of amphipathic dye, ANS, is frequently used to monitor the conformational changes induced in proteins during the unfolding process and to detect the presence of any unfolding intermediate states (Semistonov *et al.*, 1991). Previous studies have reported ANS to display a higher emission when bound to a molten globule state than when bond to a protein in its native state (Ptitsyn, 1995). Figure 3-12 illustrates the effects of urea on the binding of ANS to F56S/R81A. The ANS results show a monophasic sigmoidal transition, with the transition region lying between 3 - 5.8 M urea concentrations (Figure 3-12), which indicates that there are no detectable unfolding intermediates along the unfolding pathway. This further supports the idea that the F56S/R81A mutant follows a two-state unfolding model, with the only “stable” conformations being the native and the unfolded state.

Table 3-1: Equilibrium unfolding thermodynamic parameters.

Urea-induced equilibrium unfolding of F56S/R81A was monitored by fluorescence and CD. The thermodynamic parameters were calculated by fitting the unfolding data shown in Figure 10 to a 2-state model for a monomer (N⇔U).

Parameter	Tryptophan Fluorescence	Circular Dichroism	Average
ΔG(H2O) (kcal.mol^{-1})	**12.78 ± 1.95**	**13.85 ± 2.04**	**13.32**
***m*-value (kcal.mol^{-1}.M^{-1})**	**2.39 ± 0.37**	**2.57 ± 0.38**	**2.48**
C_m (M)	**5.3**	**5.4**	**5.35**

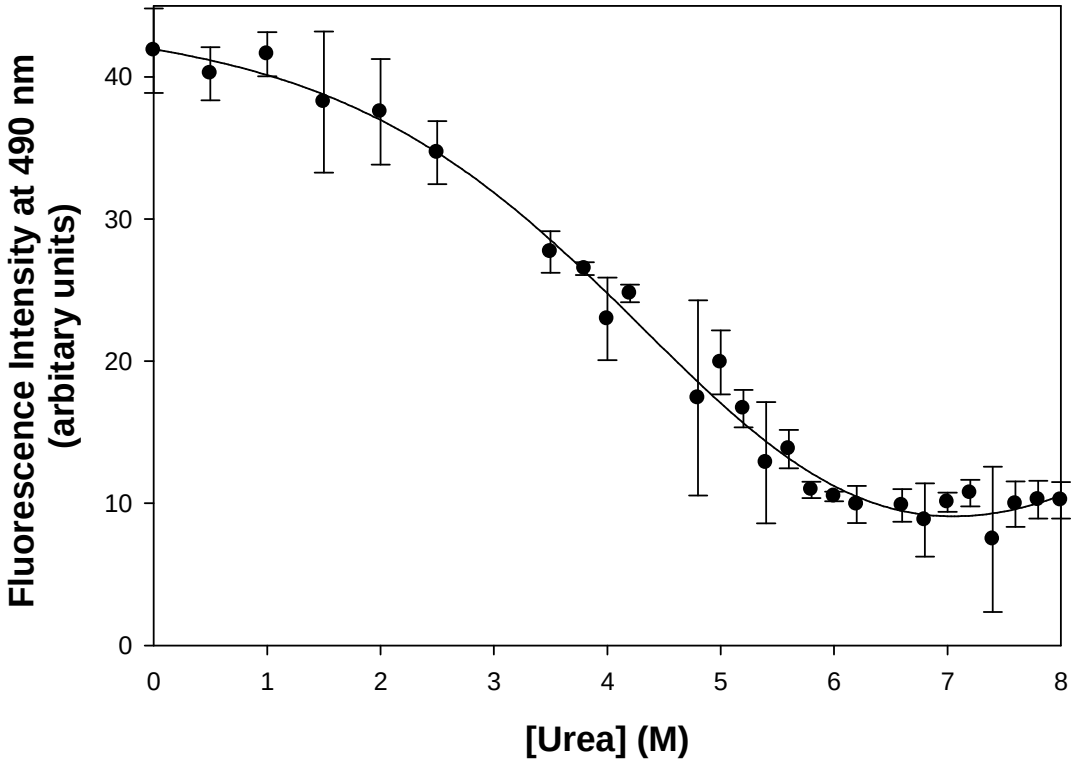


Figure 3-12: rGST M1 urea induced unfolding monitored by following ANS fluorescence.

Experiments were performed in 20 mM sodium phosphate buffer, pH6.5 with ANS at a final concentration of 200 µM. ANS was selectively excited at 390 nm, and a final protein concentration of 4 µM was used for all samples.

CHAPTER 4
DISCUSSION

The monomer-dimer transition of GST proteins has important physiological and pharmacological implications, however, studies performed in the past on the monomer-dimer equilibrium and the catalytic competence of the GST monomers have however reached conflicting conclusions. For instance, Adler and colleagues (1999) showed that class Pi GST P1-1 inhibits the activity of the pro-apoptotic enzyme c-Jun N-terminal kinase 1 (JNK) (thereby enhancing the proliferative nature of cancer cells) via formation of a JNK-GST complex, where GST P1-1 seems to be bound in a monomeric form (Adler *et al.*, 1999; Yin *et al.*, 2000). Contrary to these findings, through the use of molecular modelling Gildenhuys *et al*. (2010) have suggested that a dimeric form of GSTP1-1 can bind JNK. According to the available evidence with regard to the stability of the monomeric and dimeric forms of GSTP1-1 and the modality of the GST-JNK interaction, formation of a complex between GSTP1-1 and JNK most likely involves the dimeric form of the GST and not its monomer.

4.1. Quaternary Interactions are Important for the Catalytic Function of rGST M1-1

As stated in section 3.2.3., the F56S/R81A mutant displays only 2.7 % of the dimeric wild-type enzymes activity. It has been observed that GST lock-and key motif variants including the F52A (Vargo *et al.*, 2004) and F52S (Sayed *et al.*, 2000) of hGST A1-1, the F56R, F56S and F56E of rGST M1-1 (Hornby *et al.*, 2002), and the Y50S and Y50R of hGST P1-1 show a dramatic decrease in catalytic activity relative to the wild-type enzyme. Previous rGST M1-1 catalytic activity studies have indicated that there is a link between the structure of the dimer interface and the G-site (Hornby *et al*., 2002). The Phe-56 residue, which acts as the key of the lock-and-key motif in rGST M1-1, is located on a loop (referred to as the 56-loop) connecting α-helix 2 and β-strand 3 (Figure 4-1) (Ji *et al.*, 1992). The 56-loop is in van der Waals contact with Tyr-6 (Figure 4-1), which has been found to interact directly with the sulphur of GSH, thus substitution of Phe-56 to serine might have lead to some change in the structural conformation and/or orientation of the 56-loop and possibly contribute to the observed decrease in activity.

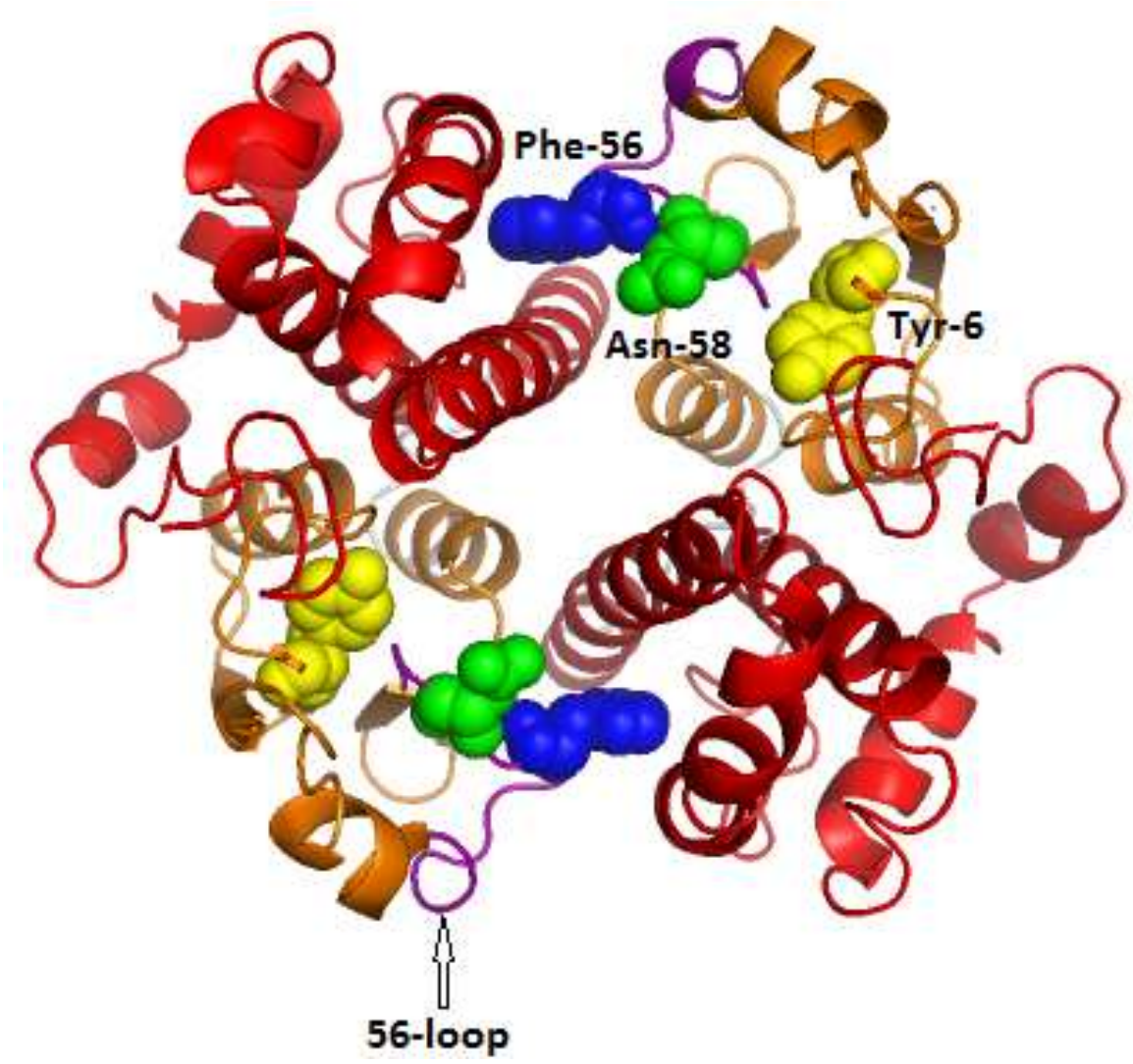


Figure 4-1: Important residues at the G-site of rGST M1-1.

The N-terminal domain is represented in orange, while the C-terminal domain is in red. Phe-56 (blue) and Asn-58 (green) are both located on the 56-loop (purple) which is in van der Waals contact with Tyr-6 (yellow) located at the G-site. Both Asn-58 and Tyr-6 interact directly with bound GSH. This figure was generated using PyMOL™ v0.99 (DeLano Scientific, 2006) (PDB code 6gst).

Another residue which has been reported to interact with bound GSH is Asn-58 (Figure 4-1) (Ji *et al.* 1992). As one might guess, this residue is in close proximity with Phe-56 residue. Having said this, it is possible that the substitution of Phe-56 results in decreased activity of the mutant protein due to the loss of this Asn-58/Phe-56 positioning, which will in turn result in the loss of the Asn-58/GSH interaction, thus altering enzymatic activity.

Unlike as is the case for class Mu GST, instead of a phenylalanine at position 56 of the amino acid sequence acting as the key of the lock-and-key motif, in class Pi GST, a tyrosine residue at position 50 of its amino acid sequence acts as the key of this motif. Substitutions at this position with either hydrophilic residues (Y50R and Y50S in Pi; F56R, F56S, and F56E in Mu) or bulky hydrophobic residues (Y50L in Pi) has been found to decrease the catalytic activity of the respective enzymes (Stenberg *et al.*, 2000; Hornby *et al.*, 2002) since such mutations alter the structural integrity of the dimer interface, and as mentioned before, part of the enzymes active site is located at the interface.

On the other hand, substitutions with aromatic residues at this position, Y50F in class Pi GST (Stenberg *et al.*, 2000) and F52Y in class Alpha GST (Sayed *et al.*, 2000), seems to favour catalytic activity. The presence of a tyrosine residue in the key position as opposed to a phenylalanine residue introduces additional electrostatic interactions at the interface, thus as is the case for the Y50F hGST P1-1 variant, the removal of the hydroxyl group of tyrosine by mutation to phenylalanine results in a less conformationally restricted enzyme which in turn exhibits enhanced catalytic activity (Sternberg *et al.*, 2000).

Light scattering and sedimentation equilibrium data has indicated the existence of relevant amounts of monomeric class Pi (Huang *et al.*, 2008), Alpha (Vargo *et al.*, 2004), and Mu (Hearne and Colman, 2006) GSTs when these enzymes are present at nanomolar levels. Furthermore, it was predicted, from the dissociation constants calculated for the homodimeric proteins (GST P1-1, 5.1 µM; GST A1-1, 0.34 µM, and GST M1-1, 0.16 µM), that at the concentration used for the activity assay (20-40 nM) these enzymes exist mainly as monomers. These groups also used site-directed mutagenesis and high KBr concentrations to force the dissociation process (Vargo *et al.*, 2004; Hearne and Colman, 2006; Huang *et al.*, 2008). Under these conditions, sedimentation equilibrium and light scattering data suggests that the monomeric form of GST P1-1(Huang *et al.*, 2008) and GST M1-1(Hearne and Colman, 2006) are catalytically competent.

In contrast to the above mentioned findings, Fabrini *et al.* (2009) conducted a study on the dimer-monomer equilibrium of wild-type GST isoenzymes (Alpha, Pi, and Mu) which demonstrated that the dimeric structure is in fact highly stable in solution, with dissociation constant values confined at or below the nanomolar range, which are far lower than the previously reported micromolar range K_d values (Vargo *et al.*, 2004; Hearne and Colman, 2006; Huang *et al.*, 2008). This K_d value shift for the monomer-dimer equilibrium of class Alpha, Pi, and Mu GST isoenzymes invalidates the idea that a fully catalytically active GST monomer exists (Fabrini *et al.*, 2009).

4.2. Loss of Quaternary Contacts Results in Enhanced Exposure of Hydrophobic Surface Area

The mechanism by which ANS binds to GST proteins and other proteins has been investigated in previous studies. Such studies have revealed that ion par interactions govern the binding of ANS to bovine serum albumin (BSA), forming between the sulfonate moiety of ANS and the cationic moieties of the protein molecule (Matulis and Lovrien, 1998). This would suggest that the extent to which the anionic ANS binds to a protein molecule is dependent on the amino acid composition of the binding site as well as the pH of the environment. Through ligand displacement studies and molecular modelling, ANS has been shown to bind to the H-site of the dimeric hGST A1-1 (Dirr *et al.*, 2005) and rGST M1-1 (Kinsley *et al.*, 2008). Ralat and Colman (2004) have also reported ANS to have a high affinity-binding site at or near the H-site of hGST P1-1. With that been said, ANS was used to probe for surface hydrophobicity of F56S/R81A rGST M1 which may arise due to subunit dissociation.

A blue-shift in ANS fluorescence emission maximum is observed in the presence of both wild-type rGST M1-1 (486 nm) and the F56S/R81A variant protein (490 nm). The extent of the blue shift is indicative of the polarity of the microenvironment of the ANS molecules, that is, the greater the extent of the blue-shift, the more hydrophobic is the ANS microenvironment (Lakowicz, 1999). For example, ANS has been reported to have a fluorescence emission maximum at 545 nm when free in water and that of 454 nm when bound to the highly hydrophobic site in apomyoglobin (Stryer, 1965). In addition, the spectral blue-shift exhibited by ANS in the presence of both protein forms is accompanied by

an enhancement in ANS fluorescence intensity, with F56S/R81A bound ANS having a higher intensity than when bound to wild-type rGST M1-1 (Figure 3-9). Water has the ability to quench ANS fluorescence, thus the fluorescence intensity of protein bound ANS is sensitive to its accessibility to water (Kirk *et al.,* 1996). The enhancement in the fluorescence of F56S/R81A bound ANS suggests that ANS is less accessible to solvent when bound to wild-type rGST M1-1, and since ANS has been shown to bind the H-site of wild-type class Mu GST M1-1 (Kinsley *et al.,* 2008), this may possibly be due to ANS being less solvent accessible since the disruption of the intersubunit contacts exposes more hydrophobic patches on the double mutant protein variant, which would otherwise be buried in the dimeric wild-type protein. In a previous study by Hornby *et al.* (2002) it was found that, much like the F56S/R81A, the extent of solvent accessibility to the ANS binding site in the two hydrophobic lock-and-key motif mutants, F56S and F56R rGST M1-1 proteins, was reduced compared to the wild-type enzyme.

4.3. Quaternary Interactions Increase the Stability of the Individual Subunits

Previous studies have revealed that interactions across the dimer interface of GST proteins play an important role in stabilising the tertiary structure of the individual subunits since GSTs are obligate functional dimeric proteins (Dirr, 2001). However, the effect of these interactions in the stabilisation of the individual subunits seem to vary from class to class, giving rise to number of unfolding pathways ranging from simple two-state transitions (Kaplan *et al.,* 1997; Wallace *et al.,* 1998) to complex multistate transitions that include thermodynamically stable intermediates which may either be monomeric or dimeric (Stevens *et al.,* 1998; Sacchetta *et al.,* 1999; Hornby *et al.,* 2000; Lian *et al.,* 2006; Abdalla and Harned, 2006; Wang *et al.,* 2009).

To determine the thermodynamic parameters of the F56S/R81A rGST M1 unfolding process, a valid model describing the urea-induced equilibrium unfolding for this protein must be employed. The observation that the F56S/R81A urea-induced unfolding curves, monitored by both far-UV CD and intrinsic tryptophan fluorescence, display single sigmoidal transitions which overlap (Figure 3-10) suggests that there are no thermodynamically stable intermediate states which may form along the unfolding process, thus the F56S/R81A unfolding curves are

analysed by fitting the raw data to a two-state model. This model assumes that the only protein species present at significant levels within the transition region are the folded native state (N) and the unfolded denatured state (U) (Pace, 1986). The determination of the Gibbs free energy change in the absence of denaturant ($\Delta G(H_2O)$) serves as means to measure a protein's conformational stability. The determined $\Delta G(H_2O)$ values obtained by following intrinsic tryptophan fluorescence (12.78 ± 1.95 kcal.mol^{-1}) and far-UV CD (13.85 ± 2.04 kcal.mol^{-1}) do not differ significantly from one another (Table 1). Unlike the F56S/R81A mutant, the dimeric wild-type enzyme is reported to unfold according to a three-state state model with a monomeric intermediate, i.e. $N_2 \Leftrightarrow 2I \Leftrightarrow 2U$, with an overall conformational stability of 19.6 kcal/mol (Hornby *et al.*, 2000). The fact that the monomeric F56S/R81A variant has a lower $\Delta G(H_2O)$ (13.3 kcal.mol^{-1}) value compared to the reported value of 16.5 kcal.mol^{-1} (Hornby *et al.*, 2000) for wild-type enzymes second transiton (i.e. I⇔U transition, the unfolding of the monomeric intermediate state) is indicative of reduced conformational stability.

An increase in a protein's solvent accessible surface area due to unfolding is related to the *m*-value of the denaturation curve (Myers *et al.*, 1995). A protein's theoretical *m*-value can thus be estimated by relating the amount of change in buried surface area to the number of amino acid residues that make up the protein molecule using the equations devised by Myers *et al.* (1995):

$$\Delta ASA = (n \times 93) - 907 \quad (17)$$

where n is the number of residues that make up the protein molecule. Using this equation, the change in solvent accessible surface area for wild-type rGST M1-1 and the F56S/R81A variant is calculated to be 39455 Å^2 and 19274 Å^2, respectively. The theoretical *m*-value of the protein can then be determined based on the calculated change in solvent accessible surface area as follows (Myers *et al.* (1995)):

$$m = 374 + (0.11 \times \Delta ASA) \quad (18)$$

Using the above equation, the theoretical *m*-value for wild-type rGST M1-1 and the F56S/R81A mutant is calculated to be 4.71 kcal.mol^{-1}.M^{-1} urea and 2.49 kcal.mol^{-1}.M^{-1} urea, respectively. Previous studies report the experimental *m*-value for wild-type rGST M1-1 to be 3.38 kcal.mol^{-1}.M^{-1} urea (Hornby *et al.*, 2000), which is significantly lower than the theoretical value. On the other hand, the experimental *m*-values obtained for the F56S/R81A mutant protein in this study using far-UV CD (2.57 ± 0.38 kcal.mol^{-1}.M^{-1} urea) and tryptophan fluorescence (2.39 ± 0.37 kcal.mol^{-1}.M^{-1} urea) data do not differ significantly

from the theoretical value, which is in contrast to the previously reported experimental *m*-value of 2.0 kcal.mol^{-1}.M^{-1} urea (Thompson *et al.*, 2006).

GSTs have been shown to be completely unfolded by urea (Wallace *et al.*, 1998), thus, the observed decrease in *m*-value of the F56S/R81A mutant (2.48 kcal.mol^{-1}.M^{-1}) relative to the I⇔U transition of the wild-type enzyme (3.45 kcal.mol^{-1}.M^{-1}) (Hornby *et al.*, 2000) is probably due to differences in the solvent exposed surface area of the protein in the folded and/or the unfolded state (Myers *et al.*, 1995; Shortle, 1995) rather than the possibility that the denaturant interacts differently with the denatured state of the two protein forms (Arakawa and Timasheff, 1984). The possibility of the F56S/R81A mutant following any other alternative unfolding pathway, rather than a 2-state unfolding pathway, is highly unlikely since no intermediate state was detected from urea induced equilibrium unfolding experiments conducted in the presence of ANS (Figure 3-12).

The present study was set to investigate the role of subunit-subunit interactions in the stabilisation of the individual subunits of the dimeric rGST M1-1 protein. From the evidence presented herein, it is apparent that the interactions at the dimer interface are essential for stabilising the individual subunits of the protein and maintaining its catalytic functionality since part of the active site is at the dimer interface (Ji *et al.*, 1992). Urea-induced equilibrium unfolding studies suggest that the loss of quaternary contacts results in reduced stability of the GST subunits, that is, the change in Gibbs free energy of unfolding for the F56S/R81A mutant is lower (13.85 kcal.mol^{-1}) compared to that of the second transition(I⇔U) of the wild-type enzyme (16.50 kcal.mol^{-1}) (Hornby *et al.*, 2000) although this is not clear judging from the changes in the unfolding transition mid-points (C_m-value) (5.35 for the mutant and 5.2 for the wild-type). Furthermore, ANS binding studies indicated an increase in the accessible hydrophobic surface area of the mutant relative to the wild-type protein. Thus, we conclude that quaternary contacts are essential for the stability and functionality of the individual subunits in dimeric rGST M1-1.

CHAPTER 5
REFERENCES

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

Abdalla, A. M. and Hamed, R. R. (2006) Multiple unfolding states of glutathione transferase from Physa acuta (Gastropoda [correction of Gastropada]: Physidae). *Biochem. Biophys. Res. Commun.* **340**, 625-632

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. *Biochem. J.* **285**, 241-245

Adler, V., Yin, Z., Fuchs, S. Y., Benezra, M., Rosario, L., Tew, K. D., Pincus, M. R., Sardana, M., Henderson, C. J., Wolf, C. R., Davis, R. J. and Ronai, Z. (1999) Regulation of JNK signalling by GSTp. *EMBO J.* **18**, 1321-1334

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

Anfinsen, C. B. (1973) Principles that govern the folding of protein chains. *Science* **181**, 223-30

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. (1991) Glutathione S-transferases: Reaction mechanism, structure, and function. *Chem. Res. Toxicol.* **4**, 131-140

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

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 Jermin, 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., Jermin, L.S., Schulte, G. K., Danley, D. E., Hoth, L. R. and Grilfor, M. C. (2000) Identification, characterization and crystal structure of the omega class glutathione transferase. *J. Biol. Chem.* **275**, 24798-24806

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

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

Buetler, T. M. And Eaton, D. L. (1992) Glutathione *S*-transferases: amino acid sequence comparison, classification and phylogenetic relationship. *J. Environ. Sci Health. Part C: Environ. Carcinog. Ecotoxicol. Rev.* **10**, 181-203

Caccuri, A. M., Antonini, G., Ascenzi, P., Nicotra, M., Nuccetelli, M., Mazzetti, A. P., Federici, G., Lo Bello, M., and Ricci, G. (1999) Temperature adaptation of glutathione *S*-transferase P1-1.Acase for homotropic regulation of substrate binding. *J. Biol. Chem.* **274**, 19276-19280

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

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

Clackson, T. and Wells, J. A. (1995) A hot spot of binding energy in a hormone-receptor interface. *Science* **267**, 383-386

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

Creighton T. E. (1990) Protein folding. *Biochem. J.* **270**, 1-16

Cunningham, B. C. and Wells, J. A. (1991) Rational design of receptor-specific variants of human growth hormone. *Proc. Natl Acad. Sci. USA*, **88**, 3407-3411

Dauber, P. and Hagler, A.T. (1980). Crystal packing, hydrogen bonding, and the effect of crystal forces on molecular conformation. *Accts. Chem. Res.* **13**, 105−112

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

Dill, K.A. and Chan, H.S. (1997) From Levinthal to pathways to funnels. *Nat. Struct. Biol.* **4**, 10-19

Dill, K. A., Ozkan, S. B., Shell, M. S. and Weikl, T. R. (2008) The protein folding problem. *Annu Rev Biophys* **37**, 289-316

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

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 stabilisation and dynamics of the C-terminal region of class alpha glutathione transferases. *J. Biol. Chem.* **280**, 19480-19487

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. and Reinemer, P. (1991) Equilibrium unfolding of class pi glutathione Stransferase. *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

Ehrlich, L., Reczko, M., Bohr, H. and Wade, R. C. (1998) Prediction of protein hydration sites from sequence by modular neural networks. *Protein Eng*. **11**, 11-19

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

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

Fabrini, R., De Luca, A., Stella, L., Mei, G., Orioni, B., Ciccone, S., Federici, G., Lo Bello, M., and Ricci, G. (2009) Monomer-dimer equilibrium in glutathione transferases: A critical re-examination. *Biochemistry* **48**, 10473–10482

Facchiano, A. M., Colonna, G. and Ragone, R. (1998) Helix stabilizing factors and stabilization of thermophilic proteins: an X-ray based study. *Protein Eng.* **11**, 753-760

Fanucchi, S., Adamson, R. J. and Dirr, H. W. (2008) Formation of an unfolding intermediate state of soluble chloride intracellular channel protein CLIC1 at acidic pH. *Biochemistry* **47**, 11674-11681

Fariselli, P., Pazos, F., Valencia, A. and Casadio, R. (2002) Prediction of protein–protein sites in hetero- complexes with neural networks. *Eur. J. Biochem*. **269**, 1356-1361

Fersht, A. (1999) Kinetics of Protein Folding. In: Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding. W.H. Freeman, New York.

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

Gallet, X., Charloteaux, B., Thomas, A. and Brasseur, R. (2000) A fast method to predict protein interaction sites from sequences. *J. Mol. Biol*. **302**, 917-926

Gildenhuys, S., Wallace, L. A. and Dirr, H. W. (2008) Stability and unfolding of reduced Escherichia coli glutaredoxin 2: A monomeric structural homologue of the glutathione transferase family. *Biochemistry* **47**, 10801-10808

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

Gittelman, M. S. and Matthews, C. R. (1990) Folding and stability of trp aporepressor from Escherichia coli. *Biochemistry* **29**, 7011-7020.

Graminski, G. F., Zhang, P., Sesay, M. A., Ammon, H. L. and Armstrong, R. N. (1989) Formation of the 1-(*S*-glutathionyl)-2,4,6-trinitrocyclohexadienate anion at the active site of glutathione *S*-transferase: Evidence for enzymatic stabilization of σ-complex intermediates in nucleophilic aromatic substitution reactions. *Biochemistry* **28**, 6252-6258.

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

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

Hagler, A,T., Dauber, P. and Lifson, S. (1979). Consistent force field studies of intermolecular forces in hydrogen bonded crystals. III. The C=O...H−O hydrogen bond and the analysis of the energetics and packing of carboxylic acids. *J. Am. Chem. Soc*. **101,** 5131−41

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

Hayes, J. D., Flanagan, J. U. and Jowsey, I. R. (2005) Glutathione transferases. *Annu. Rev. Pharmacol. Toxicol.* **45**, 51-88

Hearne, J. L. and Colman, R. F. (2006) Catalytically active monomer of class Mu glutathione transferase from rat. *Biochemistry* **45**, 5974-5984

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.

Hobart, S. A., Meinhold, D. W., Osuna, R., and Colon, W. (2002) From two-state to three- state: the effect of the P61A mutation on the dynamics and stability of the factor for inversion stimulation results in an altered equilibrium denaturation mechanism. *Biochemistry* **41**, 13744-13754

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 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

Huang, Y., Misquitta, S., Blond, S. Y., Adams, E. and Colman, R. F. (2008) Catalytically active monomer of glutathione *S*-transferase π and key residues involved in the electrostatic interaction between subunits. *J.B.C.* 283, 32880-32888

IUPAC-IUB Joint Commission on Biochemical Nomenclature (JCBN). (1984) Nomenclature and symbolism for amino acids and peptides. Recommendations 1983. *Eur. J. Biochem.* **138**, 9–37

Jaenicke, R. and Rudolph, R. (1986) Refolding and association of oligomeric proteins. *Methods Enzymol* **131**, 218-250

Jaenicke, R. (1987) Folding and association of proteins. *Progr. Biophys. Mol. Biol.* **49**, 117–237

Jaenicke, R. (1998) What ultrastable proteins teach us about protein stabilization. *Biochemistry* **63**, 312–321

Jaenicke, R. (1999) Folding and stability of domain proteins. *Progr. Biophys. Mol. Biol.* **71**, 155–241

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

Janin, J. (1995) Principles of protein-protein recognition from structure to thermodynamics. *Biochemistry.* **77**, 497-505

Janin, J. (1997) Specific versus non-specific contacts in protein crystals. *Nature Struct. Biol.* **4**, 973-974

Janin, J., Henrick, K., Moult, J., Eyck, L. T., Sternberg, M. J. and Vajda, S. (2003) CAPRI: a critical assessment of predicted interactions. *Proteins: Struct. Funct. Genet.* **52**, 2-9

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

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. (1997) Analysis of protein– protein interaction sites using surface patches. *J. Mol. Biol.* **272**, 121-132

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

Kauzmann, W. (1959) Some factors in the interpretation of protein denaturation. *Adv. Prot. Chem.* **14**, 1-63

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

Kinsley, N., Sayed, Y., Mosebi, S., Armstrong, R. N. and Dirr, H. W. (2008) Characterization of the binding of 8-anilinonaphthalene sulfonate to rat class mu GST M1-1. *Biophys. Chem.* **137**, 100-104

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

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, Protein fluorescence, pp 445-465. Plenum Publishers, USA

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

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

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

Liang, H. and Terwilliger, T.C. (1991). Reversible denaturation of the gene V protein of bacteriophage f1. *Biochemistry* **30**, 2772-2782

Lichtarge, O. and Sowa, M. E. (2002) Evolutionary predictions of binding surfaces and interactions. *Curr. Opin. Struct. Biol.* **12**, 21-27

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

Mallam, A. L. and Jackson, S. E. (2005) Folding studies on a knotted protein. *J. Mol. Biol.* **346**, 1409-1421

Mannervik, B., Alin, P., Guthernberg, C. Jansson, H., Tahir, M. K., Warholm, M. And Jornvall, H. (1985) Indentification of the three classes of Cytosolic glutathione transferases common to several mammalian species: Correlation between structural data and enzymatic properties. *Proc. Natl. Acad. Sci. U.S.A.* **82**, 7202-7206

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

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

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

Milla, M. E. and Sauer, R. T. (1994) P22 Arc repressor: folding kinetics of a single-domain, dimeric protein. *Biochemistry* **33**, 1125-1133

Murphy, K. P. (1999). Predicting binding energetics from structure: looking beyond DeltaG degrees. *Med Res Rev* **19**, 333-9.

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

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

Nölting, B., Golbik, R. and Fersht, A.R. (1995) Submillisecond events in protein folding. *Proc. Natl. Acad. Sci. USA* **92**, 10668-10672

Ofran, Y. and Rost, B. (2003) Predicted protein-protein interaction sites from local sequence information. *FEBS Lett.* **544**, 236-239

Pace, C. N. (1990). Conformational stability of globular proteins. *Trends Biochem Sci* **15**, 14-7

Park, C. and Marqusee, S. (2004) Analysis of the stability of multimeric proteins by effective Delta-G and effective m-values. *Protein Sci.* **13**, 2553-2558

Park, Y. C. and Bedouelle, H. (1998) Dimeric tyrosyl-tRNA synthetase from Bacillus stearothermophilus unfolds through a monomeric intermediate. A quantitative analysis under equilibrium conditions. *J. Biol. Chem.* **273**, 18052-18059

Pemble, S. E., Wardle, A. F. and Taylor, J. B. (1996) Glutathione *S*-transferase class Kappa: Characterisation by 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. The calculations of partial specific volumes, neutron scattering matchpoints and 280-nm absorption coefficients for proteins and glycoproteins from amino acid sequences. *Eur. J. Biochem.* **157**, 169-180

Privalov, P. L. and Makhatadze, G. I. (1993) Contribution of hydration to protein folding thermodynamics. II. The entropy and Gibbs energy of hydration. *J Mol Biol* **232**, 660-79.

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

Ragone, R. (1997) Protein association and stability of thermophilic enzymes. *Protein Pept. Lett.* **4**, 277–279

Ragone, R. (2000) How the protein concentration affects unfolding curves of oligomers. *Biopolymers* **53**, 221-225

Ragone, R. and Colonna, G. (1995) Do globular proteins require some structural peculiarity to best function at high temperatures? *J. Am. Chem. Soc.* **117**, 16–20

Ralat, L. A. and Colman, R. F. (2004) Glutathione *S*-transferase Pi has at least three distinguishable xenobiotic substrate sites close to its glutathione-binding site. *J. Biol. Chem.* **279**, 50204-50213

Ralat, L. A., Manevich, Y., Fisher, A. B., and Colman, R. F. (2006) Direct evidence for the formation of a complex between 1-cysteine peroxiredoxin and glutathione *S*-transferase pi with activity changes in both enzymes. *Biochemistry* **45**, 360–372.

Ralat, L. A.,Misquitta,S.A.,Manevich,Y., Fisher,A.B., andColman, R. F. (2008) Characterization of the complex of glutathione *S*-transferase pi and 1-cysteine peroxiredoxin. *Arch. Biochem. Biophys.* **474**, 109–118.

Rumfeldt, J. A. O., Galvagnion, C., Vassall, K. A. and Meiering, E. M. (2008) Conformational stability and folding mechanisms of dimeric proteins. *Progress in Biophysics and Molecular Biology* **98**, 61-84

Rushmore, T. H. and Pickett, C. B. (1993) Glutathione *S*-transferases, structure, regulation, and therapeutic implications. *J. Biol. Chem.* **268**, 11475–11478

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

Sanger, F., Nicklen, S. and Coulson, A. R. (1977) DNA sequencing with chain-terminating inhibitors. DNA polymerase/nucleotide sequences/bacteriophage φX174. *Proc. Natl. Acad. Sci. USA* **74**, 5463-5467

Sayed, Y., Wallace, L. A. and Dirr, H. W. (2000) The hydrophobic lock-and-key intersubunit motif of glutathione transferase A1-1: implications for catalysis, ligandin function and stability. *FEBS Lett.* **465**, 169-172

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

Sheehan, D., Meade, G., Foley, V. M. and Dowd, C. A. (2001) Structure, function and evolution of glutathione transferases: implications for classification of nonmammalian 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

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

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

Tanford C. (1968) Protein denaturation. *Adv. Prot. Chem.* **23**,122-282

Thompson, L. C., Walters, J., Burke, J., Parsons, J. F., Armstrong, R. N., and Dirr, H. W. (2006) Double mutation at the subunit interface of Glutathione Transferase rGST M1-1 results in a stable, folded monomer. *Biochemistry.* **45**, 2267-2273

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

Wang, X. Y., Zhang, Z. R. and Perrett, S. (2009) Characterization of the activity and folding of the glutathione transferase from Escherichia coli and the roles of residues Cys(10) and His(106). *Biochem. J.* **417**, 55-64

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

Weber, G. and Young, L. B. (1964). Fragmentation of Bovine Serum Albumin by Pepsin. I. the Origin of the Acid Expansion of the Albumin Molecule. *J Biol Chem* **239**, 1415-1423

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

Wongsantichon, J. and Ketterman,A. J. (2006) An intersubunit lock- and-key 'clasp' motif in the dimer interface of delta class glutathione transferase. *Biochem. J.* **394**, 135-144

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

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

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