TURUN
YLIOPISTO
UNIVERSITY
OF TURKU

STRUCTURAL STUDIES OF GLUTATHIONE TRANSFERASES TOWARDS THE DEVELOPMENT OF NOVEL APPLICATIONS AND DRUG DESIGN

Nirmal Poudel

TURUN YLIOPISTON JULKAISUJA – ANNALES UNIVERSITATIS TURKUENSIS
SARJA – SER. AI OSA – TOM. 671 | ASTRONOMICA – CHEMICA – PHYSICA – MATHEMATICA | TURKU 2022

TURUN
YLIOPISTO
UNIVERSITY
OF TURKU

STRUCTURAL STUDIES OF GLUTATHIONE TRANSFERASES TOWARDS THE DEVELOPMENT OF NOVEL APPLICATIONS AND DRUG DESIGN

Nirmal Poudel

TURUN YLIOPISTON JULKAISUJA – ANNALES UNIVERSITATIS TURKUENSIS
SARJA – SER. AI OSA – TOM. 671 | ASTRONOMICA – CHEMICA – PHYSICA – MATHEMATICA | TURKU 2022

University of Turku

Faculty of Technology
Department of Life Technologies
Biochemistry
Doctoral programme in Technology

Supervised by

Adj. Prof. Anastassios Papageorgiou
Turku Bioscience Centre, University of
Turku and Åbo Akademi University
Turku, Finland

Reviewed by

Docent Rajaram Venkatesan
Faculty of Biochemistry and
Molecular Medicine
University of Oulu
Oulu, Finland

Docent Evangelia Chrysina
Institute of Chemical Biology
National Hellenic Research Foundation
Athens, Greece

Opponent

Professor Arwen Pearson
Universität Hamburg
Institute for Nanostructure and
Solid State Physics
Hamburg, Germany

ISBN 978-951-29-8963-8 (Print)
ISBN 978-951-29-8964-5 (PDF/Online)
ISSN 0082-7002 (Print)
ISSN 2343-3175 (Online)
Painosalama, Turku, Finland 2022

To my beloved family…

UNIVERSITY OF TURKU
Faculty of Technology
Department of Life Technologies
Biochemistry
NIRMAL POUDEL: Structural Studies of Glutathione Transferases towards the Development of Novel Applications and Drug Design
Doctoral Dissertation, 148 pp.
Doctoral Programme in Technology
August 2022

ABSTRACT

Glutathione transferases (GSTs) represent a ubiquitous large family of multifunctional enzymes which protect cellular macromolecules from reactive electrophiles. GSTs catalyse the nucleophilic addition of glutathione to electrophilic groups of a large variety of hydrophobic molecules, thereby increasing their solubility to facilitate their excretion from the cell. GSTs play a key role in detoxification, drug metabolism, and multiple-herbicide resistance in weed species. However, the structural basis for the activity of several classes of GSTs remains unknown. Multiple-herbicide resistance (MHR) is a global threat to weed control in cereal crops. MHR weeds express specific phi class glutathione transferases (MHR-GSTF) that confer resistance against multiple herbicides. MHR-GSTFs, therefore, represent a promising target against MHR weeds. On the other hand, the mammalian GSTs, such as the mu-class GSTs, are linked to the development of resistance to a variety of anti-cancer drugs, resulting in the failure of the treatment. The mu-class GSTs are also associated with Parkinson´s disease as well as other illnesses related to oxidative stress. The present work investigates the structure of the MHR-GSTFs from different grass weeds and crops, in particular *Alopecurus myosuroides* and *Lolium rigidum*. The work also presents structural and functional insights into the mu-class GSTs from *Camelus dromedarius* and *Homo sapiens*.

The crystal structures of MHR-GSTF from *Alopecurus myosuroides* (*Am*GSTF) and *Lolium rigidum* (*Lr*GSTF) were determined by molecular replacement at 1.33 Å and 1.90 Å resolution, respectively. The structure of *Am*GSTF was resolved with a bound glutathione sulfenic acid (GSOH) and succinic acid (SIN) at the enzyme´s active site whereas the *Lr*GSTF structure was determined in complex with the inhibitor S-(4-nitrobenzyl) glutathione. Both enzymes showed conserved structural features compared with other phi class glutathione transferases. However, differences were observed at the C-terminal α8-helix and the H-site α4-helix that may affect the substrate specificity. Moreover, the structural analysis of GSTFs revealed an induced-fit mechanism and a decisive role of conserved Tyr^{118} and Phe^{122} in ligand binding. The results presented here provide new knowledge on the enzymology of phi class glutathione transferases and may be used to derive strategies to combat MHR weeds.

The structures of a mu-class GST from *Camelus dromedarius* (*Cd*GSTM1-1) with a bound substrate (GSH) or the reaction product, S-p-nitrobenzyl-glutathione (Nb-GSH) were determined by X-ray crystallography at 2.55 Å and 2.05 Å resolution respectively. The H-site of *Cd*GSTM1-1 is variable and lined by Met^{35}, Arg^{43}, Tyr^{116}, Phe^{209}, Leu^{210}, and Met^{212}, which govern the recognition and binding of substrate in the active site. A noticeable 4 Å move of the β2-α2 loop region upon Nb-GSH binding presents snapshots of an induced-fit mechanism that facilitates the binding of various substrates. The studies will improve our understanding of camelid GSTs detoxification mechanisms and their contribution to abiotic stress adaptation in the desert environment. Besides, the structure of ligand-free *Homo sapiens* mu-class GST (hGSTM1-1) was determined at 1.59 Å resolution. The high-resolution hGSTM1-1 structure allowed the study of the induced-fit mechanism operated by hGSTM1-1 and the binding of Nb-GSH in detail.

TURUN YLIOPISTO
Teknillinen tiedekunta
Bioteknologian laitos
Biokemia
NIRMAL POUDEL: Glutationitransferaasien rakennetutkimuksilla kohti uusien sovellusten kehitystä ja lääkeaineiden suunnittelua
Väitöskirja, 148 s.
Teknologian tohtoriohjelma
Elokuu 2022

TIIVISTELMÄ

Glutationitransferaasit (GST:t) ovat tärkeä ryhmä entsyymejä, joilla on useita biokemiallisia funktioita ja ne suojaavat solun makromolekyylejä reaktiivisilta elektrofiileiltä. GST:t katalysoivat glutationin nukleofiilistä lisäystä useiden hydrofobisten molekyylien elektrofiilisiin ryhmiin, mikä lisää niiden liukoisuutta helpottaen niiden poistamista soluista. GST:eilla on keskeinen rooli rikkakasvilajien myrkkyjen poistamisessa, lääkeaineenvaihdunnassa ja useiden rikkakasvien torjunta-aineiden vastustuskyvyssä. Useiden GST-luokkien toiminnan rakenteellinen perusta on kuitenkin tuntematon. Moniherbisidiresistenssi (MHR) on maailmanlaajuinen uhka viljakasvien rikkakasvien torjumiselle. MHR-rikkakasvit ekspressoivat spesifisiä phi-luokan glutationitransferaaseja (MHR-GSTF), jotka antavat vastustuskyvyn useita herbisidejä vastaan. Siksi MHR-GSTF:t edustavat lupaavaa kohdetta MHR-rikkakasveja vastaan. Toisaalta nisäkkäiden GST:t, kuten mu-luokan GST:t, liittyvät resistenssin kehittymiseen useille syöpälääkkeille, mikä johtaa hoidon epäonnistumiseen. Mu-luokan GST:t liittyvät myös Parkinsonin tautiin sekä muihin oksidatiiviseen stressiin liittyviin sairauksiin. Tässä työssä tutkittiin eri ruohorikkakasvien ja -kasvien, erityisesti *Alopecurus myosuroidesin* ja *Lolium rigidumin*, MHR-GSTF:ien rakennetta. Tutkimuksessa käsiteltiin myös rakenteellisia ja toiminnallisia eroja *Camelus dromedariuksen* ja *Homo sapiensin* mu-luokan GST:ien välillä.

Alopecurus myosuroidesin (*Am*GSTF) ja *Lolium rigidumin* (*Lr*GSTF) MHR-GSTF:n kide-rakenteet määritettiin templaattirakenteen avulla 1,33 Å:n ja 1,90 Å:n resoluutiolla. *Am*GSTF:n rakenteessa oli aktiiviseen kohtaan sitoutuneena glutationisulfeenihappo (GSOH) ja meripihka-happo (SIN), kun taas *Lr*GSTF:n rakenne määritettiin kompleksina inhibiittorin S-(4-nitroben-tsyyli)glutationin kanssa. Molemmilla entsyymeillä oli konservoituneita rakenteellisia piirteitä verrattuna muihin phi-luokan glutationitransferaaseihin. Kuitenkin eroja havaittiin C-termi-naalisessa α8-heliksissä ja H-alueen α4-heliksissä, jotka voivat vaikuttaa substraattispesifi-syyteen. Lisäksi GSTF:n rakenneanalyysi paljasti indusoidun konformaatiomuutoksen ja konser-voituneiden aminohappojen Tyr^{118}:n ja Phe^{122}:n ratkaisevan roolin ligandin sitoutumisessa. Tässä esitetyt tulokset tarjoavat uutta tietoa phi-luokan glutationitransferaasi-entsyymien toiminnasta, ja tuloksia voidaan hyödyntää uusien MHR-rikkakasvien torjunta-aineiden kehittämiksessä.

Camelus dromedariuksen (*Cd*GSTM1-1) mu-luokan GST:n rakenteet sidotun substraatin (GSH) tai reaktiotuotteen S-p-nitrobentsyyliglutationin (Nb-GSH) kanssa määritettiin röntgenkristallografialla 2,55 Å ja 2,05 Å resoluutiolla. CdGSTM1-1:n H-kohta on aminohapposekvenssiltään vaihteleva ja sitä reunustavat Met^{35}, Arg^{43}, Tyr^{116}, Phe^{209}, Leu^{210} ja Met^{212} säätelevät substraatin tunnistamista ja sitoutumista aktiivisessa kohdassa. Huomattava 4 Å:n liike β2-α2-silmukka-alueella Nb-GSH:n sitoutumisen yhteydessä esittää tilannekuvia indusoiduista konformaatiomuutoksista, jotka vaikuttavat erilaisten substraattien sitoutumiskykyyn. Tutkimukset parantavat ymmärrystämme kamelien GST:n detoksifikaatio-mekanismeista ja niiden vaikutuksesta abioottiseen stressiin sopeutumiseen aavikko-ympäristössä. Lisäksi ligandittoman *Homo sapiens* mu-luokan GST:n (hGSTM1-1) rakenne määritettiin 1,59 Å:n resoluutiolla. Korkean resoluution hGSTM1-1-rakenne mahdollisti hGSTM1-1:n indusoidun konformaatiomuutoksen (engl. induced-fit mechanism) ja Nb-GSH:n sitoutumisen yksityiskohtaisen tutkimuksen.

Table of Contents

Abbreviations

2, 4 –D	2, 4-Dichlorophenoxyacetic acid
*Am*GSTF	Phi class GST from *Alopecurus myosuroides*
ABC	ATP binding cassette
CDNB	1-chloro-2, 4-dinitrobenzene
CSO	S-hydroxycysteine
*Cd*GSTM1	Mu class GST from *Camelus dromedarius*
DESY	Deutsches Elektronen-Synchrotron
DDT	Dichlorodiphenyltrichloroethane
GSH	Glutathione
GST	Glutathione transferase
GSOH	Glutathione sulfenic acid
GSTU	Tau class glutathione transferase
GS8	S-Hydroxy-Glutathione
LB	Luria-Bertani
ROS	Reactive oxygen species
MAPEG	Membrane-associated proteins in eicosanoid and glutathione metabolism
MGST	Microsomal glutathione transferase
MHR	Multiple herbicide resistance
MDR	Multiple drug resistance
MR	Molecular replacement
Nb-GSH	S-(p-nitrobenzyl)-glutathione
DHAR	Dehydroascorbate reductase
*Lr*GSTF	Phi class GST from *Lolium rigidum*
hGSTM1-1	Mu class GST from *Homo sapiens*
PDB	Protein Data Bank
PAH	Polyaromatic hydrocarbon
RMSD	Root mean square deviation
SSE	Secondary structure elements
SIN	Succinic acid
T_m	Melting temperature

Amino acids

Ala	Alanine	A
Arg	Arginine	R
Asn	Asparagine	N
Asp	Aspartate	D
Cys	Cysteine	C
Gln	Glutamine	Q
Glu	Glutamate	E
Gly	Glycine	G
His	Histidine	H
Ile	Isoleucine	I
Leu	Leucine	L
Lys	Lysine	K
Met	Methionine	M
Phe	Phenylalanine	F
Pro	Proline	P
Ser	Serine	S
Thr	Threonine	T
Trp	Tryptophan	W
Tyr	Tyrosine	Y
Val	Valine	V

List of Original Publications

This dissertation is based on the following original publications, which are referred to in the text by their Roman numerals:

I Georgakis N, **Poudel N**, Papageorgiou AC, Labrou NE. Comparative structural and functional analysis of phi class glutathione transferases involved in multiple-herbicide resistance of grass weeds and crops. Plant Physiol Biochem. 2020 Apr;149:266-276. doi: 10.1016/j.plaphy.2020.02.012. Epub 2020 Feb 20. PMID: 32088578.

II Georgakis N, **Poudel N**, Vlachakis D, Papageorgiou AC, Labrou NE. Phi class glutathione transferases as molecular targets towards multiple-herbicide resistance: Inhibition analysis and pharmacophore design. Plant Physiol Biochem. 2021 Jan;158:342-352. doi: 10.1016/j.plaphy.2020.11.018. Epub 2020 Nov 17. PMID: 33257232.

III Perperopoulou F*, **Poudel N***, Papageorgiou AC, Ataya FS, Labrou NE. Structural and functional characterization of *Camelus dromedarius* glutathione transferase M1-1. Life (Basel). 2022 Jan 12;12(1):106. doi: 10.3390/life12010106. PMID: 35054499.

IV Bodourian CS*, **Poudel N***, Papageorgiou AC, Antoniadi M, Georgakis ND, Abe H; Labrou NE. Ligandability assessment of human glutathione transferase M1-1 using pesticides as chemical probes. Int. J. Mol. Sci. 2022, *23*, 3606. https://doi.org/10.3390/ijms23073606.

*Joint first authors

Original Publication not Included in the Thesis

I Chronopoulou EG, Mutabdzija L, **Poudel N**, Papageorgiou AC, Labrou NE (2022). A key role in catalysis and thermostability of conserved residues at helix H5 in human glutathione transferase A1-1. (Submitted)

1 Introduction

1.1 Glutathione transferase overview

Historically, glutathione transferases (E.C. 2.5.1.18) have been called glutathione S-transferases and, hence, they are known by the acronym GST. In 1961, the first GST-catalyzed reaction, the addition of glutathione (GSH) to 1,2 dichloro-4-nitrobenzene in cytosolic extracts of the liver, was identified [1, 2]. The discovery triggered a new era of research in the genetics and enzymology of GSTs. The three-dimensional structures of several cytosolic GSTs were determined to understand the functional and structural diversity of the enzyme[3]. The first determined crystal structure was that of a porcine GST in 1991 [4]. The structural information of the GST enzymes was crucial in deciphering some of the caveats in the evolution of the characteristic structural fold in GST and the molecular basis of the detoxification reactions [3].

GSTs are multifunctional enzymes involved in the Phase II detoxification process. They catalyze the nucleophilic addition of glutathione (γ-glutamyl-L-cysteinyl-glycine; GSH) to a range of endogenous and exogenous compounds consisting of an electrophilic carbon, nitrogen, or sulfur atom, thereby tagging the complexes for excretion from the cell (Fig. 1a, b). The non-polar exogenous compounds targeted by GSTs could include chemical carcinogens, environmental pollutants, and in some cases anti-tumor agents. Also, GSTs are involved in the inactivation of endogenous α, β- unsaturated aldehydes, hydroperoxides quinones, and epoxides formed as secondary metabolites during oxidative stress in a cell [5]. As a result, possible attacks on the cellular macromolecules from these electrophilic xenobiotics are prevented. The action of GSTs in these compounds yields more water-soluble products, hence facilitating their elimination from the cell. This detoxification mechanism is one of the vital defense armors and metabolic capabilities of living organisms [6]. The formation of reactive oxygen species (ROS) during oxidative stress is a potentially life-threatening condition, which, if left unchecked, could lead to lethal consequences like membrane dysfunction, DNA damage, and protein inactivation. Thus, to maintain a proper cellular balance against oxidative stress, GSTs work in an integrated fashion with glutathione biosynthesis, glutathione peroxidases, and glutathione S-conjugate efflux pumps [5].

Shortly after their discovery in animals, GSTs were documented in plants in 1970. The GST activity from maize was found to conjugate a potent herbicide, namely chloro-S-triazine atrazine, with GSH, thus protecting crops from herbicide damage [7]. Interestingly, GSTs are abundant in most of the living organisms. To add to their catalytic diversity, apart from GSH conjugation, GSTs possess multiple roles, such as biosynthesis of leukotrienes, prostaglandins, testosterone and progesterone, tyrosine catabolism, peroxide breakdown, dehydroascorbate reduction, apoptosis and other functions [8]. Recent advances have also outlined their role in the regulation of mitogen-activated protein kinases, as facilitators in S-glutathionylation reactions in certain proteins, and as a participant in the process of cocaine addiction [9]. GSTs also exhibit a ligand binding (‘ligandin’) function [10]. The ligandin function facilitates the binding of numerous hydrophobic and amphipathic compounds, such as xenobiotics and hormones, in a non-substrate manner into a distinct ligand-binding site (L-site) for their efficient transport to other parts of the cell or disposal. Thus, apart from their involvement in these important biological processes, GSTs have in fact surpassed their classic role in metabolism to a more advanced role in drug resistance, as prodrugs in cancer and neurodegenerative diseases [11]. Consequently, the research interest in GSTs has increased in recent years.

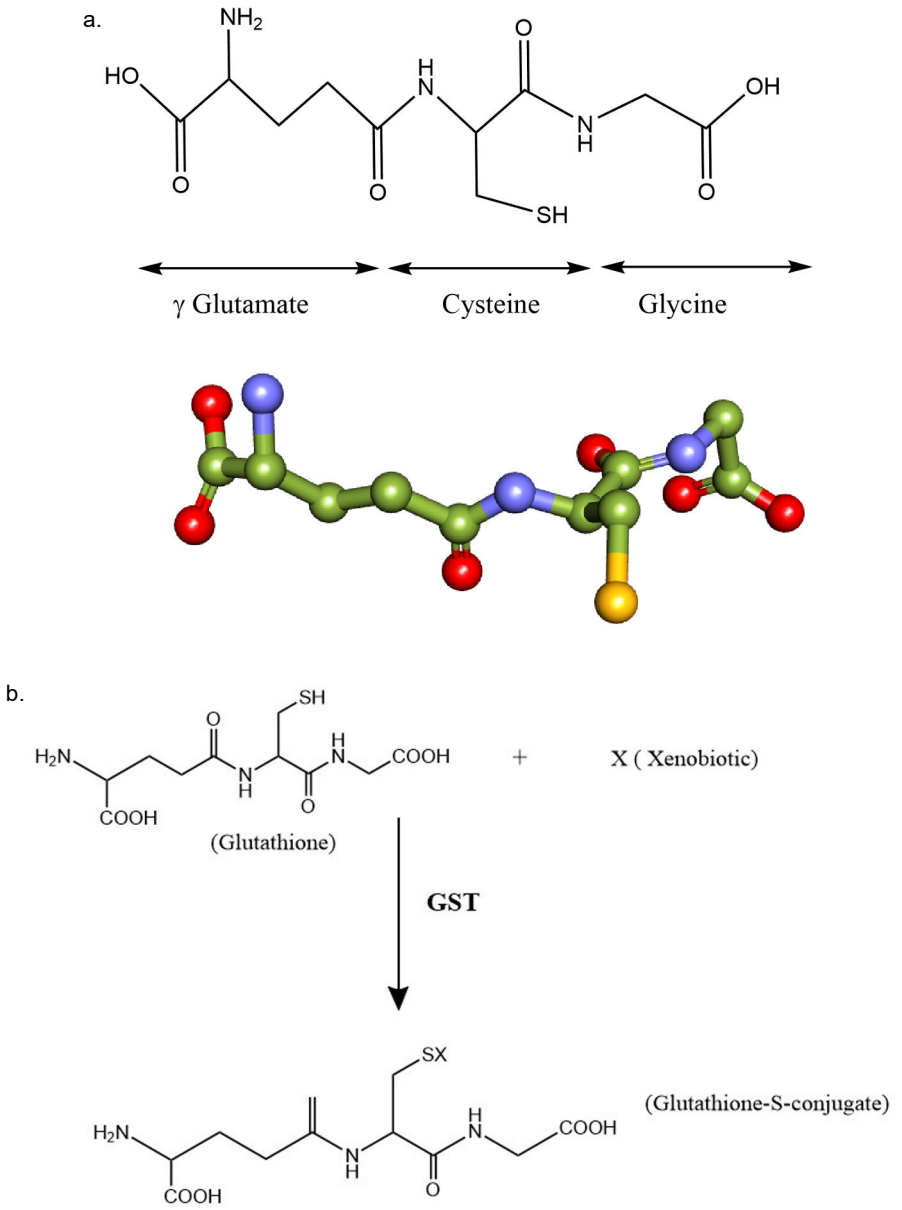


Figure 1. **a.** General structure of glutathione (GSH). The γ-glutamate, cysteine, and glycine parts of the tripeptide GSH are denoted by arrows. The nucleophilic sulfhydryl/thiol group is colored yellow in the 3D structure. **b**. General conjugation reaction catalyzed by glutathione transferase resulting in the formation of a glutathione-S-conjugate. The figures were created in ChemDraw (PerkinElmer Informatics).

1.2 GST classification

Three broad superfamilies of GSTs have been identified based on their sequence, structure similarities, cellular localizations, and immunological characteristics. These are i) the cytosolic GSTs, ii) the microsomal GSTs now designated as membrane-associated proteins in eicosanoid and glutathione metabolism (MAPEGs), and iii) the mitochondrial GSTs.

1.2.1 Cytosolic GSTs

Cytosolic GSTs comprise the largest superfamily. The currently recognized classes of cytosolic GSTs include alpha, beta, delta, epsilon, theta, mu, nu, pi, sigma, tau, phi, and omega. In humans, seven GST classes are present: α (alpha), ζ (zeta), θ (theta), μ (mu), π (pi), σ (sigma) and ω (omega) [12]. These classes are mostly abbreviated in Roman capitals A, Z, T, M, P, S, and O, respectively [13]. The GSTs within each class possess a sequence identity of more than 40% whereas between classes the sequence identity is below 25%. Several classes of currently recognized cytosolic GSTs in mammals, plants, fungi, bacteria, and insects are summarized in Table 1.

1.2.2 Microsomal GSTs

This class of GSTs comprises integral membrane proteins that are not related to other major classes of GSTs. The eukaryotic MAPEG members are represented by six families corresponding to the microsomal glutathione transferases (MGST) 1, 2, and 3, leukotriene C4 synthase (LTC4), 5-lipoxygenase activating protein (FLAP), and prostaglandin E synthase. MGST1 was the first MAPEG structure to be solved at 3.2 Å by electron crystallography [14]. The MGST 1-3 families have roles in drug metabolism whereas the other three families are involved in the leukotrienes and prostaglandin E synthesis [15].

1.2.3 Mitochondrial GSTs

This class of GSTs is known as the kappa class. It is one of the most ancient protein families with orthologues in the bacteria and eukaryotes [16]. The kappa class GST was originally isolated from the mitochondrial matrix of the rat liver [17]. Kappa class GSTs have some substrate specificities that are similar to the cytosolic GSTs but have different evolutionary pathways than other soluble GSTs. Thus, the kappa class GSTs are structurally and functionally different from the cytosolic and MAPEG GST classes.

Table 1. Classification of cytosolic GSTs in different organisms and their associated functions [18]. The asterisk (*) mark above certain classes indicate organism-specific class.

ORGANISM	CLASS	FUNCTION
Mammalian	*Alpha**	Isomerase activities, drug metabolism, peroxidase activity, detoxification
	*Mu**	Drug metabolism
	*Pi**	Drug metabolism
	Theta	Prevention of hepatocarcinogenesis, metabolism of industrial compounds
	Zeta	α–halo acids catalysis
	Omega	Oxidative stress
	Sigma	Prostaglandin synthesis
Bacteria	*Beta*	Organic compounds catabolism
	Chi	
Plants	*Phi**	Detoxification, oxidative stress, signaling, ligandin functions, intermediary metabolism
	*Tau**	
	Theta	
	Zeta	
	*Lambda**	
	DHAR	
Fungi	*Alpha*	
	Mu	
	Gamma	
Insects	*Delta**	Detoxification of environmental xenobiotics
	*Epsilon**	Detoxification of insecticides, peroxidase activity, oxidative stress
	Theta	Unknown
	Sigma	Probably against oxidative stress products
	Zeta	Tyrosine degradation
	Omega	Probably against oxidative stress

1.3 Plant GST classes - The phi and tau class of GSTs

Plant GSTs are extensively studied due to their role in herbicide detoxification[19]. The plant GST family comprises six members namely the phi, tau, theta, zeta, lambda, and dehydroascorbate reductase (DHAR). The phi (GSTF) and tau (GSTU) classes are both unique to plants and are involved in some key roles in plants such as stress tolerance, secondary metabolism, herbicide resistance in weeds, and herbicide tolerance in crops [20]. GSTFs were previously placed in the theta class and were known as Type I GSTs whereas the tau class GSTs were formerly called Type III GSTs. That is because the theta class mostly displayed lower sequence identities than other classes, thus many plant's GSTs were initially allocated to this class [21].

Both the tau and phi GSTs have a wide range of substrate specificities that align well with their function in abiotic stress tolerance. They are highly active toward several electrophilic molecules that include organic thiocyanates, halides and hydroperoxides, and α and β- unsaturated compounds [22]. Further, a few examples of herbicides detoxified by these two GSTs classes include triazines, thio-carbamates, chloroacetanilides, diphenyl-ethers, and aryloxyphenoxypropionates [18]. Both phi and tau classes of GSTs demonstrate the glutathione transferases class specificity in substrate preference. For example, GSTFs favor the binding of chloroacetanilide and thiocarbamate herbicides [23] whereas GSTUs are more preferential towards the diphenyl ethers and aryloxyphenoxypropionates herbicide detoxification [20]. These herbicides are typically used in agricultural fields co-formulated with herbicide safeners. Herbicide safeners can be defined as agrochemicals that promote herbicide tolerance in cereal crops by activating detoxification pathways [24] but without compromising weed control efficacy. The herbicide safeners have also been found to induce both the phi and tau class GSTs [25, 26]

In addition, the tau class GSTs are implicated in the transportation of anthocyanins, flavonoids, phenolics, and hormones such as auxin and cytokinin, suggesting a possible role in cell signaling [19]. On the other hand, the phi and tau classes along with the theta classes show GSH-dependent peroxidase activity (GPx, EC 1.11.1.9); it reduces organic hydroperoxides to monohydroxyalchols [19, 27]. The peroxides are usually created in plants during photosynthesis, pathogen invasion, or toxin detoxification and are harmful [28]. Illustrations of different reactions catalyzed by GSTs are shown in Fig. 2.

Interestingly, the generic peroxide and herbicide detoxification ability of plant GSTs has been used to design stress and herbicide-tolerant transgenic crops [18]. For example, a transgenic tobacco plant expressing *Zm*GSTU1 and GST I gene from maize displayed significantly higher resistance and tolerance, respectively, to certain herbicides that are otherwise inhibitory to the wild-type GSTs [29, 30].

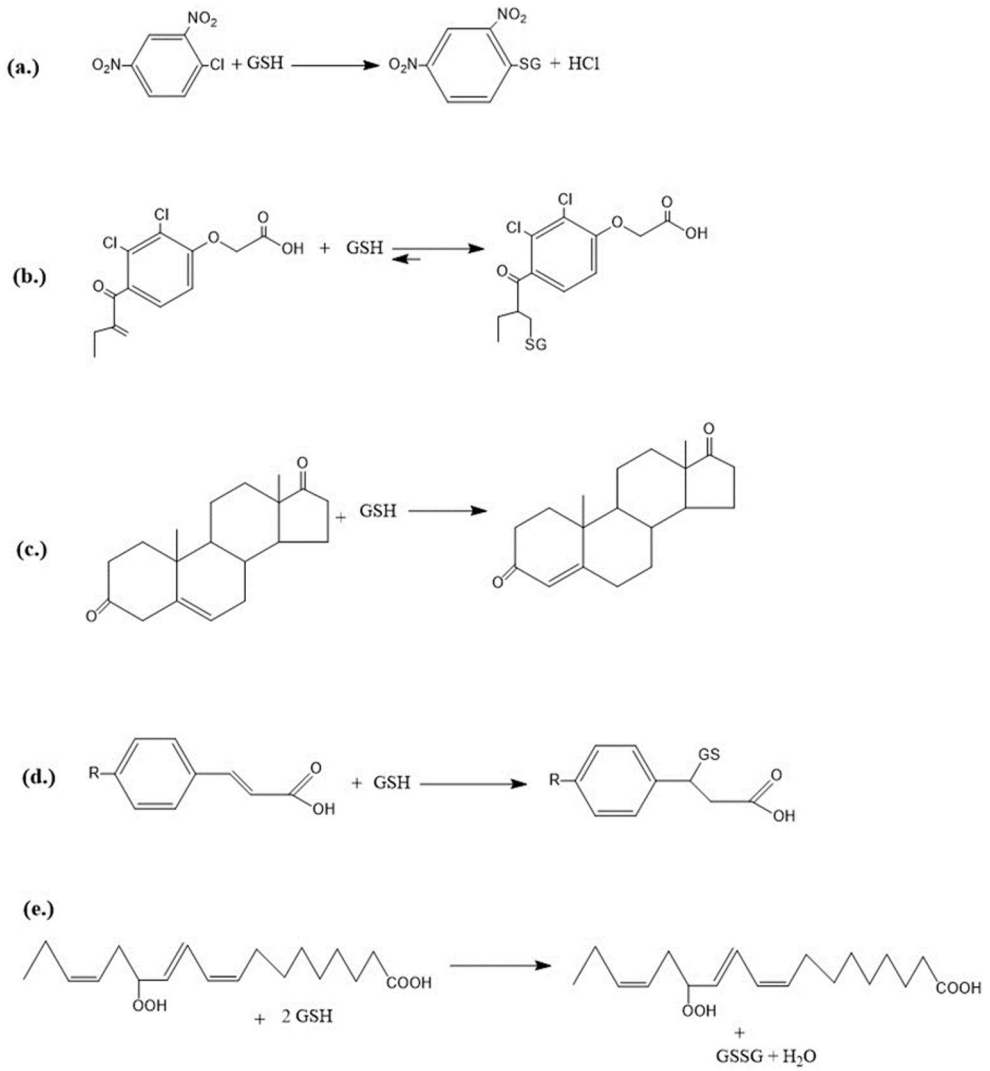


Figure 2. Examples of GST-catalyzed reactions. **(a)**. Nucleophilic aromatic substitution reaction with 1-chloro-2,4-dinitrobenzene (CDNB). CDNB is widely used as a model substrate to elucidate GST activity [31] **(b**). Michael-type addition reaction with ethacrynic acid **(c).** Isomerization of the double bond of Δ^5-androstene-3,17-dione into Δ^4-androstene-3,17-dione, a testosterone precursor **(d).** Nucleophilic addition reaction between cinnamic acid and GSH **(e)**. Peroxidase activity of GST in the reduction of fatty acid hydroperoxides to hydroxy derivatives.

1.4 Mammalian GSTs – The alpha, mu, and pi class of GSTs

The mammalian GSTs are intensively studied due to their clinical relevance in cancer, toxicology, and drug metabolism. There is evidence that certain polymorphic

changes in GST genes have been linked with the increased susceptibility to cancer development and anticancer drug resistance in humans [32]. On the other hand, mammals are also exposed to many lethal agents like environmental contaminants, xenobiotics, and reactive oxygen species (ROS) in their lifetime. These exposures have consequences on the genomic and proteome stability of mammals [33]. Thus, effective detoxification is important for natural selection. The human GST family is classified into seven families of catalytically active enzymes termed as alpha (GSTA), mu (GSTM), pi (GSTP), sigma (GSTS), theta (GSTT), zeta (GSTZ), and omega (GSTO). Alpha, mu, and pi GST classes are unique to mammals.

1.4.1 The alpha class

The Alpha class was one of the first GSTs classes to be biochemically characterized and comprises enzymes that were initially termed “ligandins” based on their ability to bind hydrophobic ligands [34]. The alpha class GSTs are dimers that are expressed in several tissues in varying amounts. The human short arm chromosome 6 comprises a cluster of five alpha class genes. The alpha class GSTs are named GSTA1-1, GSTA2-2, GSTA3-3, GSTA4-4, and GSTA5-5 [35]. It has been shown that the Alpha class plays a significant role in hydrophobic ligand-binding, such as bilirubin and steroid hormones, and is involved in isomerization reactions in the synthesis of testosterone and progesterone [36].

1.4.2 The mu class

There are a total of five mu class GST genes in humans located in the short arm of chromosome 1. The mu-class GSTs are dimeric proteins and are mostly tissue-specific in their expression [37]. For example, GSTM1-1 is more strongly expressed in the liver than in other tissues [38], GSTM2-2 is predominantly expressed in muscle, and GSTM3-3 is expressed mostly in the testis and the brain [39]. In one study it was shown that deficiency of GSTM1-1 contributes drastically to survival after chemotherapy for leukemia in children [40, 41]

1.4.3 The pi class

Humans possess only one functional pi class GST gene, namely *GSTP1*, which is present in chromosome locus 11q13 [42]. The GSTP1-1 enzyme has been the subject of intensive investigation due to its potential role in drug metabolism and cancer susceptibility. GSTP1-1 is also found expressed in many forms of tumors and is involved in the drug resistance process in cancer [43]. Polymorphisms in genes that encode GSTP1 alter susceptibility to chemotherapy-induced carcinogenesis. It was

found that GSTP1 poses an increased risk of acute myeloid leukemia (AML) after cytotoxic chemotherapy but not after the radiotherapy [44].

1.4.4 The theta and omega class

The mammalian theta class GSTs are also found responsible for cancer susceptibility [45], organ transplant rejection, and autoimmune hepatitis [46]. While the sigma class and zeta class have a single GST class and are known as hemopoietic prostaglandin D synthase (HPGDS) and maleylacetoacetate isomerase (MAAI), respectively [47]. Lastly, the mammalian omega class is unique in the sense that the GSTO class has cysteine as catalytic residues unlike tyrosine and serine in other mammalian GSTs. Thus, this class is capable of catalyzing reduction reactions that are not substrates for other GSTs. The GSTO1 is determined to be a factor in Alzheimer's, Parkinson´s, and amyotrophic lateral sclerosis diseases [48]. Other research has also shown that GSTO1 polymorphism is linked to vascular dementia and stroke [49]. Thus, it is reasonable to conclude that mammalian GSTs possess a plethora of potential targets for medical research.

Representatives of each human GST class and respective chromosome locus are summarized in Table 2 below [36].

Table 2. Representative human GSTs and their respective chromosome locus.

GSTs	CHROMOSOME	REFERENCES
CLASS ALPHA	6p12.2	
GSTA1-1		[50]
GSTA2-2		[51]
GSTA3-3		[51]
GSTA4-4		[52]
CLASS MU	1p13.3	
GSTM1-1		[53]
GSTM2-2		[54, 55]
CLASS PI	11q13.3	
GSTP1-1		[56]
CLASS THETA	22q11.23	
GSTT1-1		[57]
GSTT2-2		[58]
CLASS SIGMA	4q22.3	
GSTS1-1		[59]
CLASS ZETA	14q24.3	
GSTZ1-1		[60]
CLASS OMEGA	10q25.1	
GSTO1-1		[61]
GSTO2-2		[62]

1.5 GST nomenclature

The large size of the GST family outlines the necessity of an unambiguous system of protein nomenclature. Many forms of soluble GSTs do exist and the inconsistencies in the GST nomenclatures confuse the interpretation of any research. To simplify future assignments in GSTs research and correlate the GSTs naming to their evolutionary background, a new gene and enzyme nomenclature for GSTs was proposed already in the year 2000 [19]. Such a classification system already exists in naming mammalian GSTs and the same has been extended to the plant GSTs [13]. Thus, in the new method, e.g., *Xy*GSTM1-1 would indicate a homodimer that is encoded by the *XyGstM1* gene whereas *Xy*GSTM1-2 would indicate a heterodimer encoded by *XyGstM1* and *XyGstM2* genes (Fig. 3).

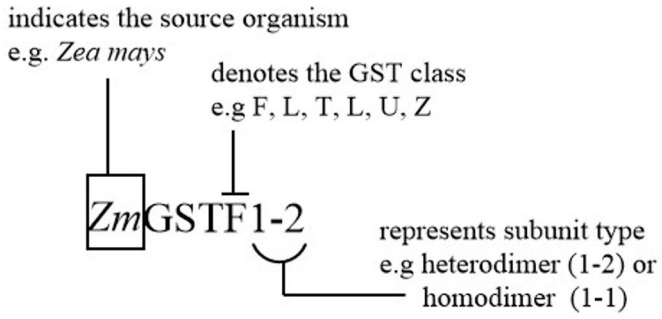


Figure 3. A representative plant GST nomenclature system [19].

1.6 Role of GSTs in detoxification

One of the prerequisites for cellular survival is the robust mechanism to manage reactive oxygen species (ROS), endogenous phytochemicals, and exogenous toxins. The exogenous toxins include environmental additives, which range from several insecticides and pesticides used in agriculture to improve the yield to polyaromatic hydrocarbons (PAHs) and other herbicides. GSH is the most abundant form of organic sulfur in plants apart from those incorporated into the proteins [63]. It has been estimated that the concentration of GSH in plants just exceeds 1mM in the cytoplasm. GSH is used not only by GSTs but also by many other enzymes, such as formaldehyde dehydrogenases, glyoxalases, and glutathione peroxidases, in living organisms [63]. However, the most noticeable use of GSH is in glutathione transferases with a collective constitution of over 1% of the soluble protein in maize leaves [64]. The key role of plant GSTs in herbicide detoxification is one of the reasons for their extensive studies. Thus, several plant GSTs that contribute to the herbicide tolerance in major crops have been well characterized [64] because of their agronomic values.

1.6.1 The detoxification process

The detoxification of xenobiotics in plants generally proceeds in three phases, known as transformation (Phase I), conjugation (Phase II), and compartmentation and storage (Phase III). A further possibility of degradation in the vacuole is characterized as the phase IV step. In addition, there are further three possible tracks in phase III to determine the fate of conjugated xenobiotics. The xenobiotics could be either exported into cell vacuoles in a process called vacuolar sequestration or they could be exported into the extracellular space or even deposited into lignin or other cell-wall components [65]. The detoxification process is superficially similar in animals where xenobiotics metabolism occurs in the liver and instead of compartmentation there is an active excretion process in phase III via urine and feces. Thus, the term "green livers" is coined for plants acting as a global sink for environmental pollutants [66] in parallel with the functionality of an animal´s liver. The schematic representation of the four-stage detoxification reaction is shown below (Fig. 4).

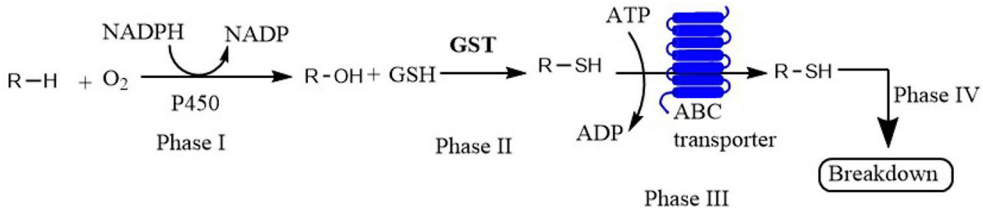


Figure 4. Simplified four-phase detoxification mechanism, namely the P450 monooxygenase activity (I), the GST-conjugation reaction (II), the ATP-dependent transport (III), followed by (IV), the breakdown in the vacuole.

1.6.1.1 Phase I detoxification

In this phase, the xenobiotics are activated by the introduction of certain reactive centers or functional groups so that they could more easily be acted upon by phase II enzymes. For example, oxidations reactions are the most common phase I reactions carried out by cytochrome P450 monooxygenases (P450s) or other mixed-function oxidases [67]. The phase I activation reactions are also catalyzed by other enzymes, such as esterases and amidases, but the majority are catalyzed by the cytochrome P450 system. The P450 proteins contain heme as a cofactor, possess several isoforms, and are found in the endoplasmic reticulum. More than 60 genes of P450 proteins have been identified in *Arabidopsis*, thus the P450s can recognize a wide array of the substrate and catalyze diverse reactions [68]. In cases where xenobiotics already possess reactive centers essential for phase II reactions, the detoxification process could proceed without the phase I stage. Some of the other categories of reactions catalyzed by P450s include hydroxylations, isomerizations, dealkylations, epoxidations, deaminations, and decarboxylations [69, 70]. The

reactions yield oxygenated products which are generally more reactive, thus appropriate for downstream detoxification phases [71]. The role of P450s in non-target site resistance alone and in coordination with phase II detoxifying enzymes, such as GSTs, has been well-established [72].

1.6.1.2 Phase II detoxification

This stage involves the conjugation of a hydrophilic molecule such as glucose, malonate, or glutathione to the activated xenobiotics from phase I using thiols or sugars to form water-soluble conjugates [73]. The reaction may be catalyzed by respective glucosyl-, malonyl- or glutathione transferases. The latter, however, is the most common phase II conjugation enzyme. The conjugation reaction leads to the formation of inactive R-SH products (Fig. 5), which are then exported from the cytosol into the vacuole by ABC transporters in Phase III [7]. In contrast to the phototoxic compound generated by phase I reaction, the product of phase II is either not toxic or less toxic than its parent compound [73]. The over-expression or increased catalytic activity of GSTs involved in the phase II detoxification process could lead to multi-herbicide resistance which is even more complicated and diverse than those mediated by P450s [29, 74]. Some cases of herbicide resistance mediated by P450s and GSTs are mentioned in Table 3.

1.6.1.3 Phase III detoxification

The accumulation of GSH-conjugates in the cytosol is potentially harmful, for example, the GST activity could be hampered, or the conjugates may be converted into other toxic metabolites. Thus, the GSH-conjugates need to be exported from the cytosol. In the phase III detoxification stage, the inactivated and conjugated molecule is transported into the vacuole or the apoplast. The most common group of transporters used in the stage are the ABC transporters [75], which facilitate the transport across the plasma membrane or tonoplast. The ABC transporters have one or two ATP binding cassettes for the active transport using ATP hydrolysis. In plants, the first ABC transporter was identified in the transportation of GSH conjugated chemicals in 1993 [76]. Interestingly, there are also some studies suggesting ABC transporter-mediated herbicide resistance [77, 78]. Besides, there have been instances where ABC transporters were found upregulated upon herbicide safener application in crops. The result was the synchronized overexpression of GSH-conjugates, ABC transporters, and GST enzymes [79].

1.6.1.4 Phase IV detoxification

The last step of the detoxification step involves the further degradation of the conjugated molecules in the vacuole or extracellular space. The overall detoxification process is shown in Fig. 5.

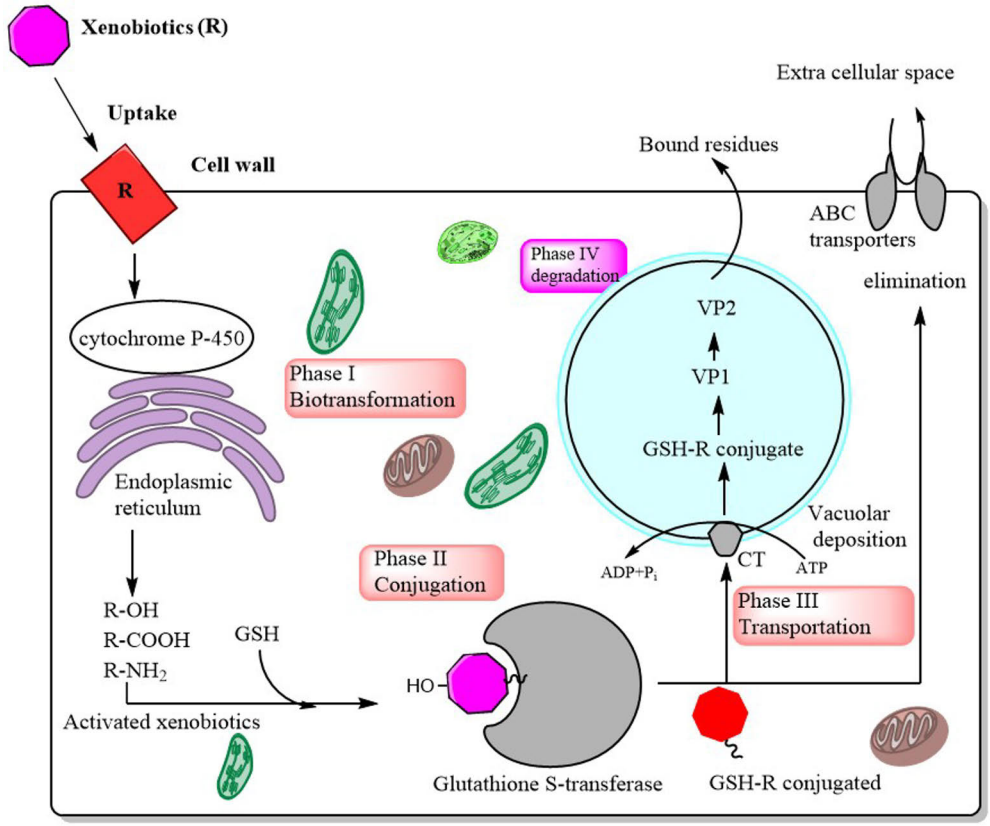


Figure 5. Schematic representation of phase I, II, III, and IV in xenobiotic (R) detoxification in plant cells. The detoxification phases occur in different organelles and compartments in plants. Phase I involves the functionalization of xenobiotics, phase II is the conjugation step to glutathione via GSTs activity, phase III is the elimination or compartmentation of the conjugated products from the cell and in the final phase IV, the further degradation of the conjugated product occurs. Abbreviations: CT, glutathione conjugate transporter; GSH, glutathione; VP1, VP2, vacuolar peptidases including the vacuolar dipeptidase and carboxypeptidase. The diagram was created in ChemDraw (PerkinElmer Informatics).

1.7 Multiple herbicide resistance (MHR) in plants

The feeding of 7.9 billion people in the world today is a great challenge and agriculture is one of the main food providers for the world's population. One of the major threats that agriculture faces today is the growth of grass weeds, which compete with crops for water and nutrients. Thus, herbicides together with insecticides used in farmland make a major contribution to global food production. However, the application of herbicides over a prolonged time in an extensive area not only triggers herbicide resistance in the weed species but also poses a negative impact on the local ecology. Thus, it seems very likely that herbicides cannot ever fully succeed to eliminate the weeds, just as the antibiotics are not able to eliminate

the bacteria or the insecticides to eradicate the pests [80]. However, their effectiveness could be customized to obtain optimal weed control.

Herbicide resistance is an evolutionary process that is governed by several factors such as the dose and frequency of use, the mechanism of herbicide action, the biology of weed species (e.g., cross-pollination or self-pollution), and dominance, frequency, and several resistance genes in the weed species [80]. Interestingly, the first reported case of herbicide resistance was not against the most widely used herbicide 2,4-Dichlorophenoxyacetic acid (2, 4-D) but against two relatively new herbicides called atrazine and simazine [81]. The reason for this pattern of resistance was later addressed by modeling studies [82]. But these early cases of resistance were not taken seriously and were circumvented with blame on bad farming practices. Also, in those early studies, any research and laboratory data about the potential resistance or new strains were kept secret to avoid mass panic. Thus, valuable early scientific evidence on the herbicide resistance materials was eventually destroyed [82]. It is now argued that with a suitable quarantine or confinement practice in those early days when herbicide resistance first appeared, the problem would not have been so widespread and deadly now.

Although the research on herbicide resistance has an immense agricultural value, it has also contributed to other areas of science, such as enzymology, pharmacology, ecology, and sustainability. Notably, the work on the X-ray crystallography structures of atrazine-resistant and atrazine-susceptible target proteins was awarded the Nobel prize in medicine with inferences in the drug binding [83]. The extent of the herbicide resistance problem can be observed in a continuously updated and dedicated "international herbicide-resistant weed database" website [84]. Many weed species are resistant to a particular herbicide and each weed species in turn is resistant to several herbicidal modes of action (Fig. 6a and b).

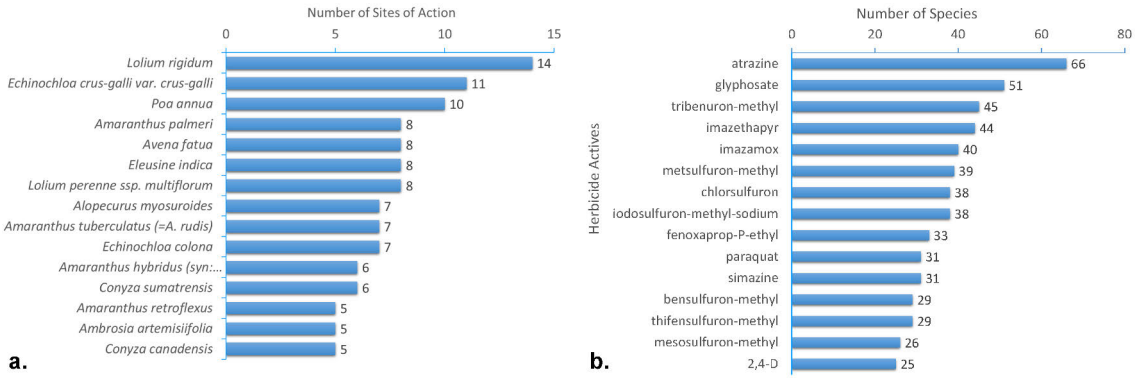


Figure 6. **a.** Top fifteen weed species resistant to multiple herbicides site of action. **b.** Representative resistant species to top fifteen individual active herbicides. The diagram has been adapted from www.weedscience.org (accessed 31st Aug 2021) [84].

1.7.1 Modes of herbicide resistance

Most herbicides act by inhibiting certain plant enzymes at the target site which are vital for metabolism. As herbicides enter the plants, they are translocated and reach their destination at a lethal dose [80]. Thus, the herbicide resistance achieved by a plant can be classified either as target-site resistance or non-target site resistance. Prior studies also have highlighted that the type of resistance achieved depends upon the dose of herbicide used. For example, a high-dose herbicide triggers target-site resistance whereas a low dose application promotes non-target site resistance [85, 86].

1.7.1.1 Target-site resistance

The target-site resistance (TSR) is a process in which a herbicide reaches its target in the cell in a sufficient quantity but certain modifications or mutations in the target site diminish the herbicide´s effectiveness. Such modifications could be done by gene mutations concerning one of the amino acids in the active site, thereby preventing the herbicide to bind its target. One example of target-site resistance is the resistance to Photosystem II (PSII) Inhibiting herbicide, triazine. The use of triazine flourished since the early 1950s and its wide uncontrolled application led to the herbicide resistance evolution. Triazine competes with other molecules in the cell to bind with PSII and its incapability to bind to the preferred location leads to carbohydrate starvation and oxidative stress [87]. The most remarkable feature of the PSII triazine target-site resistance is the global evolution of a single resistance mutation leading to the herbicide resistance [88].

1.7.1.2 Non-target site resistance

In the non-target site resistance, the amount of active herbicide reaching the target is minimal. Such decreased delivery may be due to the reduced penetration of herbicides in the plants in the prior steps, lower herbicide translocation rates, or increased rates of herbicide sequestration and metabolism [80]. Some common examples of non-target site resistance include mutations in Phase I, II, and III enzymes and transporters gene families, namely, P450, GSTs, and ABC transporters respectively. P450s are also involved in the hydroxylation or dealkylation of herbicides, the process commonly called transformation or activation of herbicides for the following conjugation stages and ultimately excretion from the active cell [69]. The resistance caused by P450s-catalyzed enhanced rates of herbicide metabolism was first identified in the 1980s in *Alopecurus myosuroides* (black-grass) and *Lolium rigidum* (ryegrass). The resistance was observed across herbicides with wide modes of action and surprisingly across the herbicide groups used for the very first time [89, 90].

Finally, it can be concluded that even for strong herbicides for which the resistance evolution is difficult, if the selection pressure is extensive, insistent, and

powerful, then resistance mechanisms will evolve in large weed populations sooner or later [80].

1.7.2 Role of GSTs in MHR

Multiple herbicide resistance (MHR) is a global problem that poses risk to sustainable agricultural practice**.** MHR was first spotted in *A. myosuroides* in Essex, England in 1982 which was soon followed by a series of independent outbreaks across the European continent [91]. MHR in plants aligns with that of multiple-drug resistance (MDR) in animals. Specifically, MDR in humans is associated with overexpression of pi class glutathione transferase (GSTP1) whereas MHR with the overexpression of phi class GSTs (GSTF1) [92]. In contrast to the TSR, MHR weeds utilize a central defense mechanism that metabolizes herbicides with different modes of action and poses a greater threat to agriculture. MHR is peculiar with its enhanced ability to detoxify an array of herbicides and is also termed as the metabolism-based resistance. Also, the non-target site resistance such as the MHR is achieved by the chemical modifications of the herbicide and/or the excretion or compartmentation of the herbicide [93]. Once the resistance is established in a weed population, it vigorously spreads to other populations via pollen and might also get transmitted to other species via hybridization [94, 95]. Some examples of non-target site herbicide resistance mediated by GSTs and P450s are summarized in Table 3.

Table 3. Non-target site herbicide resistance governed by P450s and GSTs in different weed species.

GENE FAMILY	SCIENTIFIC NAME	COMMON NAME	HERBICIDE	REFERENCES
P450	*Lolium rigidum*	Rigid ryegrass	Atrazine, simazine, diuron	[96]
P450	*Alopecurus myosuroides*	Blackgrass	Chlorotoluron	[97]
P450	*Lolium rigidum*	Rigid ryegrass	Diclofop-methyl	[98]
P450	*Lolium rigidum*	Rigid ryegrass	Chlorsulfuron	[99]
GST	*Alopecurus myosuroides*	Blackgrass	Clodinafop, haloxyfop, Aryloxyphenoxypropionates (fops), flupyrsulfuron, ureas	[72]
GST	*Alopecurus myosuroides*	Blackgrass	Fenoxaprop-ethyl and slightly resistant to cholorotoluron.	[91, 100]
GST	*Lolium multiflorum*	Italian ryegrass	Isoproturon, cycloxydim, Diclofop-methyl, tralkoxydim, fluazifop-P-butyl and tralkoxydim	[101]
GST	*Abutilon theopharasti*	Velvetleaf	Atrazine	[102–104]

Data suggest that *Alopecurus myosuroides* and *Lolium rigidum* are the most challenging weeds for Europe while Australia and the USA lead the race for the presence of the highest number of herbicide-resistant weeds [84]. More specifically, *A. myosuroides* and *L. rigidum* are found to be resistant to seven and fourteen modes of action of herbicides, respectively (Fig.6a). Both of these weeds develop resistance that has been correlated with their gene expression [105, 106], more specifically, GSTF associated genes [100, 107]. In grass weeds, MHR is associated with the increased levels of herbicide-detoxifying enzymes, including cytochrome P450 mixed-function oxidases (CYPs), UDP-glucose dependent glycosyltransferases (UGTs), GSTs, and membrane-associated ABC proteins [108–115]. More conveniently, the xenobiotics detoxifying enzymes and transporters are termed collectively as "xenome" [28]. Thus, in *A.myosuroides*, the upregulation of the xenome is correlated with the resistance to several herbicides [100, 107]. Previous studies have also revealed a higher expression phi class GSTs (GSTF1) in independent MHR black-grass population [100, 107] and named as *Am*GSTF1 [28].

Recent results have also demonstrated that not only has *Am*GSTF1 a pivotal role in MHR in black-grass but also another orthologous, *Lr*GSTF1, has enhanced activity under herbicide incorporation. The studies also imply that other weeds might possess similar functions under stress conditions. Also, the key role of *Am*GSTF1 has been demonstrated in a study that showed that the GSTP1- and MDR-inhibiting pharmacophore compound (e.g., 4-chloro-7-nitro-benoxadiazole) inhibited *Am*GSTF1 activity and helped reinstate the herbicide control in MHR blackgrass [92]. The results suggested the need for further research in the GSTF family to identify potential targets for chemical intervention in resistant weed management.

On the other hand, with the limited influx of new herbicide compounds, the identification of co-active compounds, especially GST inhibitors that could be used conjointly with currently available herbicides may certainly alleviate the MHR-GSTF challenge in the fields today. Interestingly, such combinatorial compounds have already been used in the fields. For example, piperonyl butoxide, which inhibits the detoxification enzymes cytochrome P450 monooxygenases and esterases, is used to counteract metabolic-based insecticide resistance in *Trialeurodes vaporariorum* (glasshouse whitefly) [116]. Another example is the *Am*GSTF inhibiting compound 4-chloro-nitro-benoxadiazole as discussed in the earlier paragraph. Thus, the urgency posed by MHR weeds could be minimized with the application of GST inhibitors alongside herbicides in the agricultural fields. Some examples of GST-mediated herbicide conjugation reactions are given in Fig. 7.

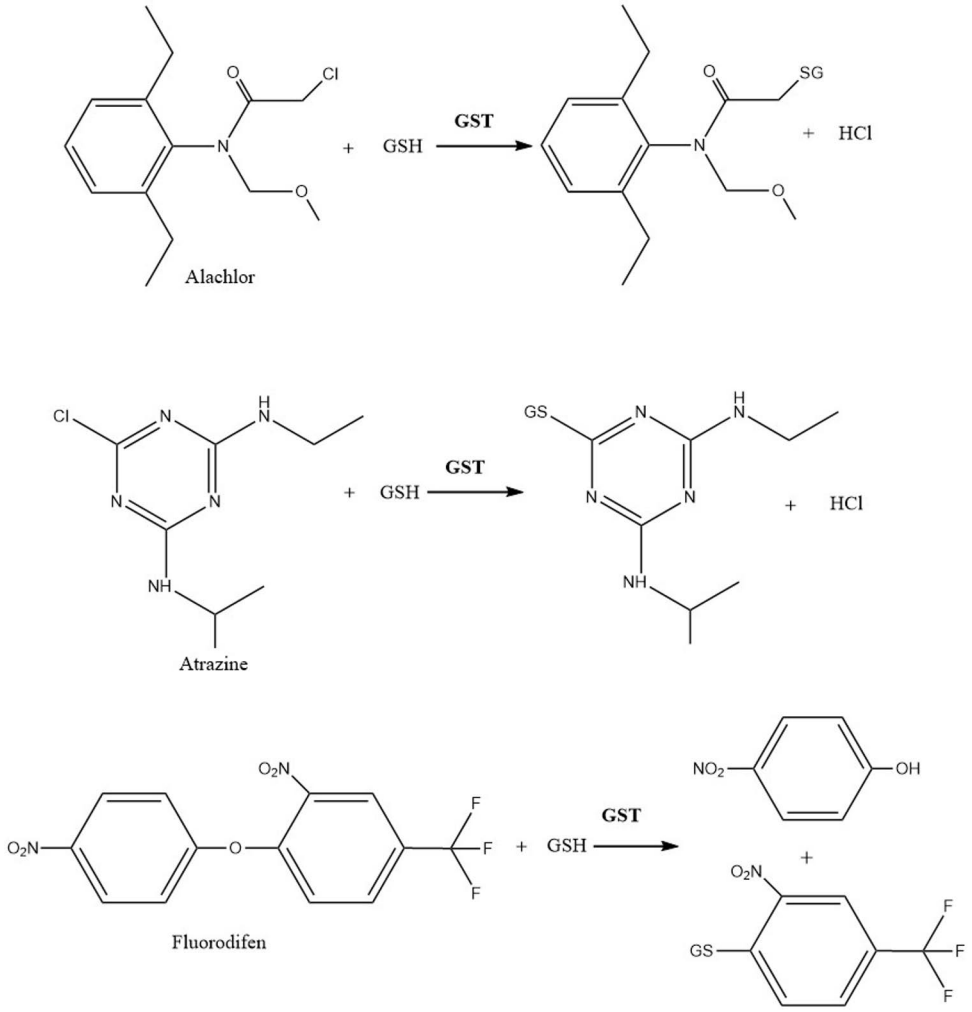


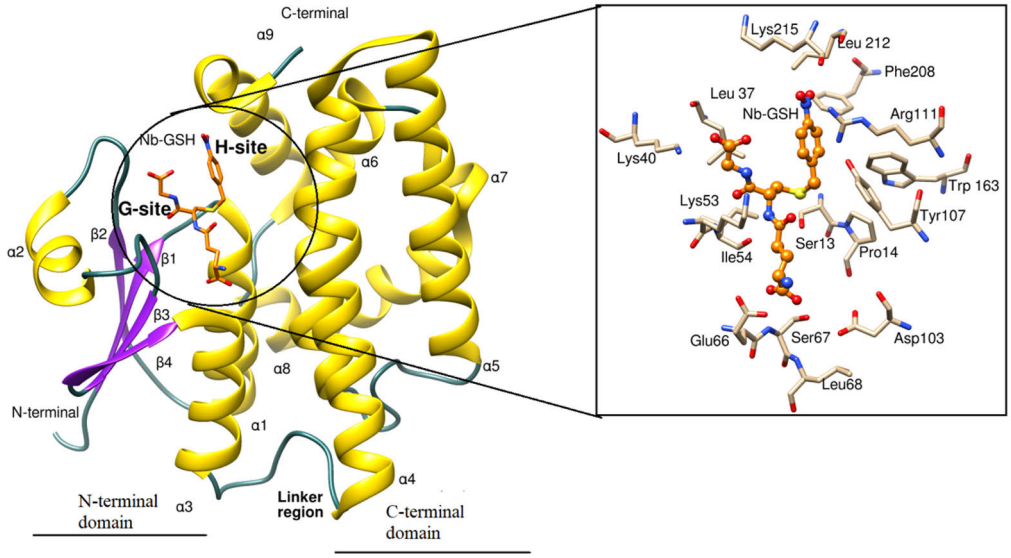
Figure 7. Plant GSTs act on some common herbicides, such as alachlor (chloroacetanilide), atrazine (chlorotriazine), and fluorodifen (diphenylether). Alachlor and atrazine detoxifications are examples of halide substitution reactions whereas fluorodifen detoxification is a GST-mediated substitution reaction resulting in herbicide cleavage. All three herbicides are widely used to control several weeds in maize and legume plants [117–119].

1.8 General structure of GSTs

Glutathione transferases are dimeric enzymes. They are either homodimers of a single gene product or heterodimers of subunits encoded by different genes. Each soluble GST dimer consists of two subunits, each approximately 26 kDa. The subunits that make up the dimer are related by a two-fold symmetry. The X-ray

crystallography has shown similar structures for three phi-class GSTs from maize [120, 121] and *Arabidopsis* [122] which only share 20% sequence identity.

Each GST monomer consists of two domains, a α/β domain that includes helices α1-α3 and a largely α-helical domain comprising helices α4-α9. The former is called the N-terminal thioredoxin-like domain and it houses a highly conserved region. The glutathione-binding site named as G-site is located towards the top of the α/β-domain. The N-terminal thioredoxin-like domain consists of β1α1β2 (α2) β3β4α3 secondary structural elements (SSE). The SSEs are arranged in such a way that the four β-strands lie in the center flanked by two α1 and α3 helices, while the irregular α2 helical segment orients itself away from the β-sheets to the outside (Fig.8 a, b, and c). This N-terminal domain constitutes roughly one-third of the protein. The second domain is termed the C-terminal domain and houses the binding site for hydrophobic co-substrates (H-site). The H-site in the cytosolic GSTs is in the cleft between the N and C-terminal domains and residues from either domain contribute to the interaction with the bound substrate. The H-site is a highly variable region between GST classes and is responsible for hydrophobic substrate specificity. The H-site in plant GSTs has a larger cleft than mammalian GSTs, thus enabling the binding of much larger and more diverse substrates [27, 123] . Some of the substrates that bind to the mammalian H-site include organic halides, arene oxides, epoxides, organic thiocyanates, organic nitrate esters, and several drugs [124]. The H-site was first identified in the crystal structure of a human pi-class GST, hGSTP1-1, which was co-crystallized with S-hexyl glutathione at 2.8 Å [125].



a.

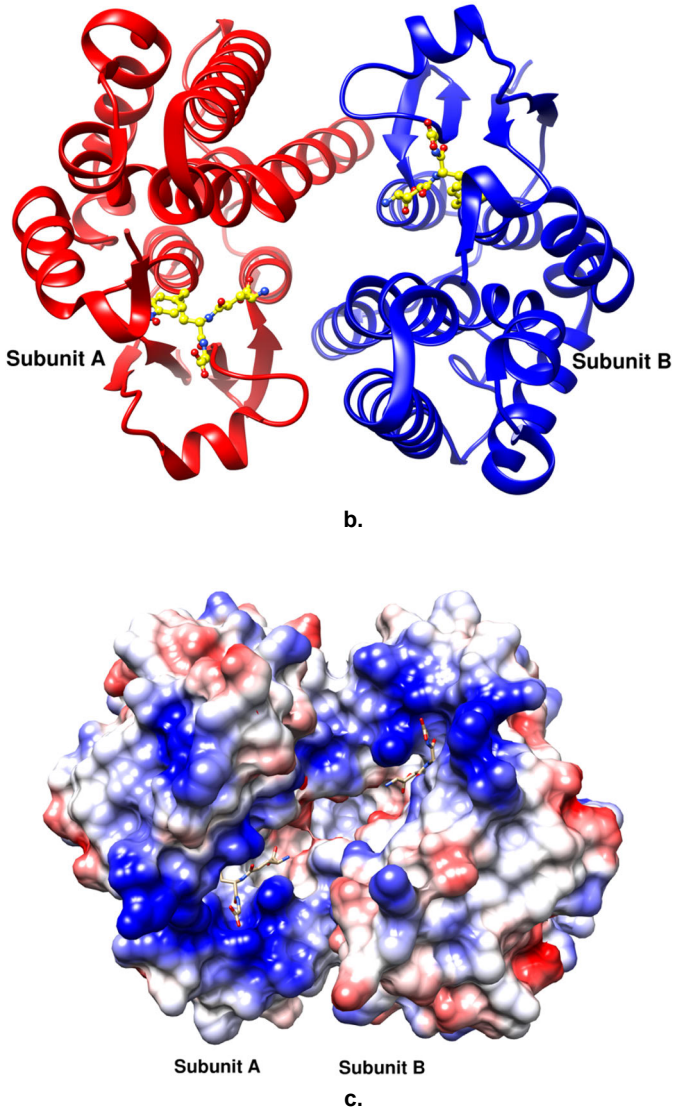


Figure 8. **a.** A ribbon representation of a tau class GST monomer from *Glycine max*, *Gm*GSTU4-4 (PDB id 2vo4). The N-terminal domain and C-terminal domain are denoted by lines below the structure. The α-helix, β-strands, and loops are colored in yellow, purple, and cyan, respectively. A single molecule of Nb-GSH bound to the active site of the enzyme is shown in stick representation (orange). The active site is signified with a circle and the interaction of Nb-GSH with residues in the active site is highlighted in a box. Also, the G- and H-sites, the linker region, and SSEs are labeled. **b.** Dimeric structure of the enzyme is shown in ribbon representation, each chain is colored differently, and bound Nb-GSH is colored yellow. **c.** A molecular surface representation of the two subunits and the surface colored according to the electrostatic potential. The positive and negative potentials are generated in blue and red colors respectively. The bound ligand, Nb-GSH, is shown in the active-site cavity in yellow.

Furthermore, the C-terminal domain and the N-terminal domain are connected by a short connector sequence of approximately ten residues. Comparative analysis of all the 3D representative plant GST structures reveals that all classes of plant GSTs exhibit a similar protein fold (Fig.9). Major differences, however, occur in the connector region between the N-terminal and the C-terminal domain, the α2-β2 loop region, the bend region in α4 and α6 helix, and the length of helix α9 (Fig. 8). Also, the inter-subunit hydrophobic interface is mostly conserved. The active site of GSTF and GSTU have a more conserved architecture with a common purpose of activating the sulfur atom of GSH to form a reactive thiolate species. Also, the hydrophobic substrate is oriented in such a way that its electrophilic center is accessible for the nucleophilic substitution/addition reaction. Some of the known crystal structures of phi and tau plant GSTs are given in Table 4.

Table 4. Some plant GST crystal structures in PDB from *Arabidopsis thaliana* (*At*), *Zea mays* (*Zm*), and *Glycine max* (*Gm*).

GSTs	CLASS	PDB ID	LIGAND	RESOLUTION	REFERENCES
***At*GSTF2-2**	Phi	1gnw	2 S-hexyl GSH	2.2 Å	[122]
***Zm*GSTF1-1**	Phi	1axd	Lactoyl GSH	2.5 Å	[120]
***Zm*GSTF1-1**	Phi	1bye	Atrazine GSH	2.8 Å	[126]
***Gm*GSTU4-4**	Tau	2vo4	S-(p-nitro-benzyl)- GSH	1.75 Å	[27]
***Gm*GSTU4-4**	Tau	3fhs	GSH	2.7 Å	[123]

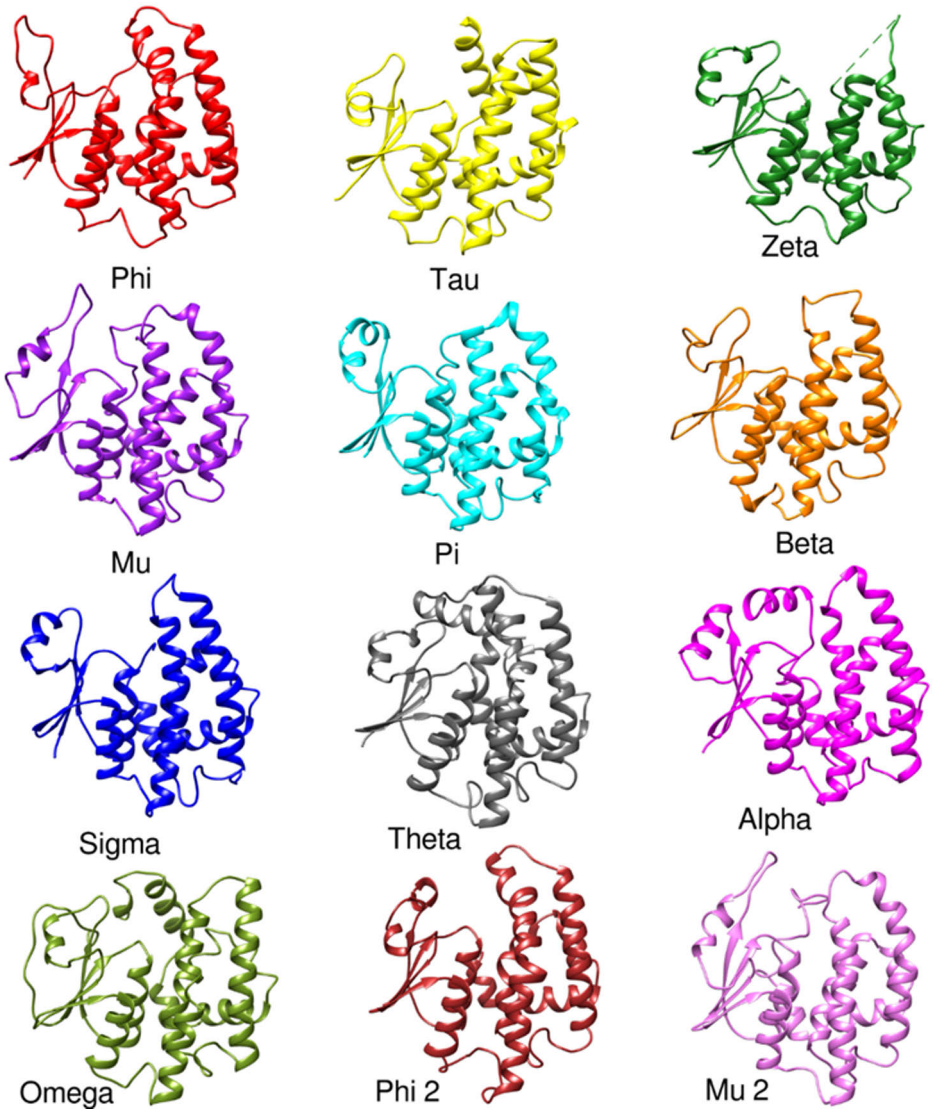


Figure 9. An overview of different enzyme folds in several classes of the GST superfamily. To facilitate a comparative view, the monomers were superimposed onto the Phi class enzyme *Zm*GSTF3-3 (PDB id 1aw9) with UCSF Chimera [127].

In the active site, a conserved Tyr residue is present in the mammalian alpha, mu, and pi classes GSTs, a conserved active site Ser residue is identified in the case of plant tau, phi, theta, and zeta classes, and a conserved Cys residue in the case of lambda GSTs and the DHARs [128]. The role of a conserved Ser residue in the active site is to activate GSH by acting as a hydrogen bond donor to the thiol group of the GSH [123] while the Cys-containing GSTs participate in deglutathionylation

activity. The catalytic Cys, then, performs a nucleophilic attack on glutathionylated molecules resulting in protein-glutathione adducts [128].

On the other hand, the dimer interface is large, and the buried surface area (BSA) of the interface is between 2700 $Å^2$ and 3400 $Å^2$ [7]. The dimer interface can take different shapes depending upon the organism and class, for example in phi, alpha, mu, and pi class it is a hydrophobic ball-and- socket interface while in theta, sigma, and beta class the interface is hydrophilic [3]. The interaction between the residues of both subunits is important for the dimer stability [27]. The interaction between each adjacent monomer is also minimal [120, 121]. Thus, the reason for dimerization remains unclear when each of the monomers can function independently [7]. Only subunits from the same class can dimerize and not from different GST classes owing to incompatibility of the interfacial residues [129].

Despite the traditional view that enzymes are supposed to be highly substrate-specific, GSTs are amongst the most promiscuous enzymes. Such promiscuous catalytic activity is one of the prerequisites for board substrate acceptance as in the case of the cellular detoxification process [130]. The functional promiscuity of GSTs is allied with their structural flexibility and heterogeneity. A comparison of several plant GSTs with substrate-bound and unbound forms have also demonstrated an induced substrate fitting mechanism in GSTs. For example, the crystal structure of GST-I from *Zea mays* has been determined at 2.5 Å resolution. The ligand-bound GST-I when compared with the apo structure of maize GST-III demonstrated a movement of a ten-residue loop upon binding of the ligand to the active site. [120, 123]. It has also been shown that heterodimer formation increases GST´s ability to uniquely recognize special phytochemicals and exhibit vital kinetic properties for the metabolite management [19]. Therefore, by understanding the structure-function relationship between GSTs and their substrates, it is possible to design new molecules that are more stable, specific, effective, and eco-friendly.

1.9 Catalytic mechanism

To interpret the role of GSTs in xenobiotics detoxification, it is essential to fully understand the catalytic mechanism and the influence of the enzymes' 3D structure on substrate selectivity. There are possibly three points to consider in understanding the catalytic mechanism of GSTs. First, the interaction of the enzyme with GSH and the chemical alternations thereafter. Second, the residues involved in binding and catalysis, and third, the binding of specific substrates in the active H-site.

The nucleophilic addition reaction catalyzed by GSTs can be divided into two stages. In the first stage, the substrate binds to the active site and triggers the activation of GSH by deprotonation to form a nucleophilic thiolate anion. The second stage involves a thiolate anion mediated nucleophilic attack at the substrate

electrophilic center [19]. The addition of thiols to the electrophile occurs via the activation of thiols to the thiolate anion. The thiolate anion is 10^9 times more reactive than its conjugate acid [131]. Thus, the first role of the enzyme is to remove a proton from the thiol of GSH and generate a strong nucleophile GS^-. This is achieved by considerably lowering the p*K*a of the bound thiol group of GSH so that its majority fraction is ionized at physiological pH. For this reason, the enzyme active site has the positive electrostatic potential with an objective to reduce the p*K*a value or to destabilize the thiol group [131]. The p*K*a value of the sulfhydryl group of GSH is reduced from 9.0 in an aqueous solution to approximately 6.5 when it is bound to the active site to promote the proton dissociation [132].

On the other hand, the active site residues involved in the catalysis vary amongst the GST classes. In the alpha, mu, pi, and sigma classes the active site residue is a Tyr (involved in the deprotonation of the bound thiol group of GSH). In the delta, epsilon, theta, phi, tau, and zeta GSTs, the active site residue is Ser, and in omega and beta class GSTs, the residue is Cys [18]. The Tyr also plays an additional role in correctly orientating the GSH in the active site and facilitating the safe passage of the departing thiol proton out of the enzyme active site [133]. In *Zea mays*, five residues located in the G-site (Ser^{11}, His^{40}, Lys^{41}, Gln^{53}, and Ser^{67}) were identified to contribute to the GSH binding and its activation. Similarly, hydrophobic residues in the H-site (Met^{10}, Trp^{12}, Phe^{35}, and Ile^{118}) are held responsible to maintain the active site architecture to contribute to the catalytic efficacy [132]. The specificity of GSH binding to the G-site is extremely high; only GSH and some closely related molecules can serve as substrates contrary to the broad specificities for the electrophilic substrates in the H-site [64].

1.10 The “ligandin” function of GSTs

In addition to their catalytic roles, GSTs are also involved in binding many lipophilic molecules (MW> 400 Da), such as bile acids, fatty acids, bilirubin, haemin, and certain drugs in animals. Also, in plants, GSTs bind several hydrophobic molecules to a ligand-binding site known as the L-site [130, 134]. Because of this property, GSTs were also originally termed “ligandins”. The localization of the L-site varies highly. For example, in phi class GSTs it is found located next to the G-site whereas in the tau-class the L-site is located in the hydrophobic surface pocket [122, 123]. In contrast, in the case of human pi-class glutathione transferases GST P1-1, the L-site is located in the electrophilic substrate-binding side [56]. The ligandin GSTs function is exploited in the transportation of anthocyanins, flavonoids, oxylipins, phenolics, and even hormones such as auxin and cytokinin with implications in cell signaling [123]. While the ligand binding to the L-site is non-competitive with the substrate binding to H-site but recently it has interestingly been shown that the ligand

binding promotes the glutathione-conjugating activity of *Arabidopsis thaliana* GSTF2 (*At*GSTF2) [135]. The same study also pointed out that, *At*GSTs are involved in more instances of signaling and transport rather than in GSH mediated conjugation reactions. Some of the identified ligands for *At*GSTF2 are camalexin, 1-acetyl-β-carboline, indole-3-aldehyde, lumichrome, harmane, norharmane and quercetin-3-O-rhamnoside [135]. Also, these observations highlight that plant GSTs have continuously evolved regulatory non-catalytic functions in addition to their ability to selectively bind the biologically active ligands. In some more examples as such, tau class GSTs from *Arabidopsis* (*At*GSTUs) and *Zea mays* (*Zm*GSTUs) have been shown to bind selectively to fatty acids and porphyrins respectively [136]. But it is surprising to note that to this date far less enthusiasm has been put into understanding the GSTs ligand binding properties, rather the effort has been converged to understand the GSH-dependent catalytic activities [137]. Thus, addressing the ligand binding properties of plant GSTs could not only provide additional information on diversification of the GST families and future agronomic values but also provide the basis for the design of new engineered GSTs with altered ligandin properties [138].

1.11 Applications of GSTs

Because of their role in xenobiotics detoxification and protection from oxidative stress, GSTs have found various applications in medicine, environment protection, and sustainable livelihood practice. Some of their applications are discussed below.

1.11.1 Role of GSTs in cancer and chemotherapeutics

Cancer as complex is the disease, the treatment is even more complicated by several factors. Some factors include the limited capability of chemotherapeutic agents to selectively target cancer cells and avoid side effects, chemoresistance attained by tumors, metastasis, and relapse. Glutathione transferases are responsible to detoxify several traditional anti-cancer drugs and are involved in cell signaling pathways that govern cell division. Unfortunately, cancer cells exploit the detoxifying power of GSTs to acquire drug resistance and avoid cell death. As a result, several GSTs members are found overexpressed in varieties of cancers. The overexpression of GSTs has been correlated with the increased detoxification of antineoplastic drugs [139], thus contributing to multi-drug resistance (MDR). Some of the common traditional chemotherapeutic drugs such as chlorambucil, cyclophosphamide, melphalan, thiotepa, cisplatin, and carmustine are directly conjugated with GSH and inactivated by GSTs [140].

Recently, GST isoenzyme's involvement in modulating cell signaling pathways that control cell proliferation and apoptosis has been shown [141].GSTs are involved

in promoting the MDR by the inhibition of c-Jun N-terminal kinase (JNK) signaling pathways. JNK is one of the key signaling cassettes of the mitogen-activated protein kinase (MAPK) signaling pathway. The module plays an important role in apoptosis, inflammation, cytokine production, and metabolism. Thus, inhibition of JNK pathways particularly by GSTP1 helps tumor cells evade apoptosis and metastasize [142]. The GSTP1:JNK1 protein-protein interaction is also one of the first discoveries in the GST-mediated MAPK regulation [141]. The human GSTM1/GSTP1 also has been continuously explored as a possible molecular target for chemotherapy as a plausible means to sensitize drug-resistance tumors that overexpress GSTs [43]. However, much needs to be explored regarding the role of GSTM1 isoenzymes in chemoresistance but some are already found to take part in the anti-cancer drugs detoxification [143]. In addition, the high expression of GSTP1 in cancer cells supplemented by its role as an enzyme detoxifying anti-cancer drugs and an inhibitor of apoptosis, GSTP1 is a promising cancer therapeutic target. While still much remains to be explored about these interactions in detail, GSTP1 inhibitors are certainly important therapeutic agents for cancer and other diseases associated with abnormal cell proliferation (Table 5).

Table 5. Examples of different classes of GST inhibitors. GSH analogs compete with GSH for binding to the enzyme. Natural products and small molecules directly bind to the active site of GSTs and inhibit their function.

Inhibitor class	Examples
Natural products	Aloe emodin and Benastatin A
Small molecules	Ethacrynic acid (EA), NBDHEX, Nitazoxanide (NTZ) and Haloenol lactone (HEL)
GSH analogs	TLK117, TLK 119

1.11.2 GSTs as prodrugs in cancer therapy

The GSH-conjugation function of GSTs can also be exploited in the activation of certain prodrugs used in the cancer therapy [144]. GSTs overexpression in cancer-specific cells is the key to the prodrugs bioactivation [145]. Prodrugs are not pharmacologically active *in vitro*, while their introduction *in vivo* followed by certain enzymatical modifications convert them into active drug molecules. The main advantage of prodrugs is that they minimize the damage to non-cancerous cells while maximizing their effect on the target cells. Several inactive cytotoxic agents are available which can be converted into active drugs. Examples include canfosfamide (TLK286), nitric oxide (NO) prodrugs, metformin analogs, and doxorubicin analogs (DOX). Thus, effective drug activity is ensured in the prodrug activated cancer cells

with overexpressed GSTs. The active drug concentration will be much lower in normal cells with normal GSTs levels, posing less damage [144, 146]. An example of prodrug activation by GST is shown below (Fig. 10).

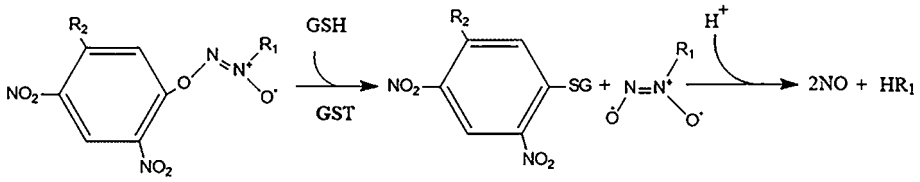


Figure 10. GSH conjugation reaction with nitric oxide releasing agent. Two molecules of nitric oxide are subsequently released along with the diazeniumdiolate anion. A high level of nitric oxide in cancer cells can induce apoptosis [147].

1.11.3 GSTs as biosensors

The ability of GSTs to catalyze GSH-conjugation reactions has been explored for the determination of biosensors for direct monitoring of environmental pollutants, such as herbicides and insecticides [148–151]. Studies in GSTs mostly relate to their biological relevance, but the enzyme can also be developed as an analytical probe [149]. Biosensors can offer a cheap, fast, and easy assessment of environmental pollutants. Biosensors can provide real-time qualitative and quantitative information with minimum sample preparation and generate data of superior accuracy [18]. Early detection of toxic compounds is critical to avoid their circulation in the ecosystems that pose threats to human health through food chains. With the developments in enzyme engineering, new mutants of GSTs with altered xenobiotic-binding properties can be studied and screened for bio-sensing and bio-scavenging [152]. The enzymatic biosensors cannot only be used in agriculture for sensing the contaminants in the soil but they can also be used in direct monitoring of chemotherapeutic drug concentration in humans.

Different GST isoenzymes have already been used as electrochemical or potentiometric biosensors to detect several herbicides, such as atrazine [149], malathion [148], molinate [153], α-endosulfan [108], insecticides DDT [154], dieldrin, and spiromesifen. In fact, due to the high usage of certain herbicides like alachlor, there is a significant deposition of the remnants of the toxic chemicals in food and water, which pose serious effects to human health worldwide [155]. To deal with this problem, efficient GST-based optic biosensors have been developed. The method utilized an engineered GST enzyme (*Gm*GSTsf) which was immobilized on a polyvinylidenefluoride membrane [156].

The GST-based biosensors also have applications in quantifying anticancer drugs in patients [157]. To optimize cancer therapy and minimize the risks associated

with drug overdose or underdose, it is essential to monitor the concentration of active cancer medications in patients. There are currently several HPLC-detector based and probe-based techniques available but these methods have their limitations [157]. For example, the HPLC-based techniques are more expensive, and the probe-based technique is not suitable for broad varieties of anti-cancer drugs. In that light, GST-based biosensors are more effective, they can bind a wide array of cancer drugs efficiently and provide a strong signal. Thus, it is important in the future to explore the structural and catalytic promiscuity of GSTs for the design of engineered and robust GST variants for developing biosensors that have wide applications from agriculture and environment to precision medicine.

1.11.4 Applications of GSTs in agriculture

In agriculture, GSTs have been studied in the development of transgenic plants with increased resistance to biotic and abiotic stresses [158–160]. Specific phi class glutathione transferases (GSTF) are overexpressed in multiple-herbicide resistant (MHR) weeds e.g. *Alopecurus myosuroides* GSTF (*Am*GSTF) and *Lolium rigidum* (*Lr*GSTF) and thus the GSTFs represent a promising target against the MHR weeds [92]. Furthermore, the ubiquitous nature of the enzyme and its ability to detoxify a wide range of pesticides, herbicides, and pollutants contribute to the development of herbicide-tolerant crops. For example, the overexpression of GSTs from maize and soybeans in wheat and tobacco to detoxify chloroacetanilides and diphenyl ether herbicides respectively [161, 162] ensured enhanced plant survival. Similar work has also been performed in other GSTs classes such as the lambda and zeta. Such as overexpression of zeta class GST has been performed to develop low-temperature tolerant rice species [163]. Also, the overexpression of lambda class GSTs could be a good strategy for increased stress tolerance in several crops [164]. Thus, the ability of GSTs to maintain tolerance to various stress conditions, for example cold, salinity, and drought, by detoxification of xenobiotic compounds and reactive oxygen species (ROS) could be exploited further in the development of new crop protection agents and tools to counteract herbicide resistance.

2 Aims of the thesis

The thesis aimed to provide structural and functional insights into i) plant phi class GSTs (GSTF) from *Alopecurus myosuroides* (blackgrass) and *Lolium rigidum* (ryegrass) ii) mammalian mu class GSTs (GSTM) from *Camelus dromedarius* (camel) and *Homo sapiens* (human).

The specific aims are given below:

I. To obtain structural and functional information of the plant phi class GSTs *Am*GSTF and *Lr*GSTF involved in herbicide resistance.

II. To obtain structural and functional information about the mammalian mu class GSTs, *Cd*GSTM1 and hGSTM1-1.

III. To explore the active sites of different GST enzyme classes and understand the recognition and binding mechanism of various substrates and inhibitors into the active site.

IV. To compare the structures of *Am*GSTF, *Lr*GSTF, *Cd*GSTM1, and hGSTM1-1 with other GSTs within the same class to better understand the catalysis and specificity of the enzymes.

3 Materials and Methods

The Materials and Methods are divided into five sections: expression and purification, crystallization, data collection and processing, structure determination, and analysis.

3.1 Expression and purification

*Am*GSTF and *Lr*GSTF were cloned and expressed in *E.coli* and purified by affinity chromatography [165]. In brief, *E. coli* transformants harboring the recombinant plasmid were plated on LB agar medium. The expression was induced by isopropyl-1-thio-β-galactopyranoside (IPTG) in appropriate medium for 24 h at 37 °C. The culture was then centrifuged, resuspended in lysis buffer, and sonicated. The resultant cell lysate was centrifuged and the enzyme supernatant was loaded on a Ni^{2+}-IDA-Sepharose affinity chromatography adsorbent for purification [132]. The enzymes were eluted with buffer containing imidazole. The enzyme purity was evaluated by 12% SDS-PAGE. All purification steps were carried out at 4 °C.

The recombinant *Cd*GSTM1-1 was expressed as described by Pouliou *et al*., 2017 [166]. *E. coli* BL21 (DE3) cells harboring recombinant plasmid were grown at 37 °C in LB medium containing ampicillin. The expression was induced by IPTG. The culture was then centrifuged, resuspended in potassium phosphate buffer, sonicated, and centrifuged. The protein supernatant was used for affinity chromatography using GSH-Sepharose matrix. The protein purity was determined by 12% (w/w) SDS-PAGE.

hGSTM1-1 was expressed and purified as described by Georgakis *et al*., 2017 [165]. Recombinant *E. coli* cells were grown in LB medium containing kanamycin. Affinity chromatography using glutathione immobilized on agarose was employed for purification. Protein purity was >98% as determined by 12% SDS-PAGE. Protein concentrations for all enzymes were determined with the Bradford method [167] using BSA as standard.

3.2 Crystallization

3.2.1 *Am*GSTF crystallization

Before setting up crystallization trials, the purified *Am*GSTF was concentrated to ~ 6mg/mL in 10 mM HEPES and 100 mM NaCl (pH 7.0) buffer solution. Crystallization trials were set up in commercial screens with 96-well plates using the sitting-drop vapor diffusion method at 20 ℃. The plates were inspected in an automated fashion at regular time intervals with a Formulatrix RockImager. Crystals were found in the PACT crystallization screen. The PACT suite consists of 96 solutions with different salt, buffer, and precipitant formulations [168]. Crystals of *Am*GSTF were found in conditions 2 and 38 of the PACT crystallization screen. Although data were collected from both the crystal conditions, only crystals (Fig.11a) from condition 2 (25% w/v PEG 1500, 0.1 M SPG, pH 5) gave good diffraction data that could be processed.

3.2.2 *Lr*GSTF crystallization

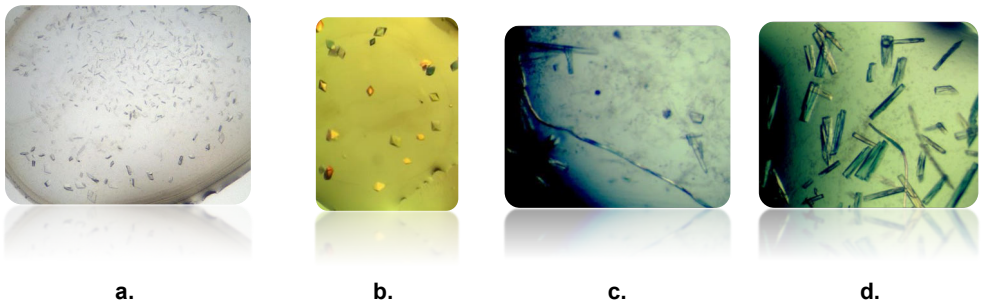
Before crystallization, *Lr*GSTF was concentrated to approximately 15 mg/mL in 10 mM HEPES buffer, 100 mM NaCl, pH 7.0. The crystallization trials were set up in five commercially available crystallization screens in 96-well plates at 20 ºC and inspected with a Formulatrix RockImager. Crystals in the presence of 10 mM S-(4-nitrobenzyl)glutathione (Nb-GSH) were found in the MIDAS*plus* screen (Molecular Dimensions), condition 21 (20% w/v Sokalan PA 25CL, 0.1 M HEPES buffer pH 7.0, 0.1 M Sodium tartrate) (Fig.11b). The MIDAS*plus* screen consists of new precipitants, such as Sokalan PA 25 CL, PPGBA 230, or Sokalan CP45 as alternatives to polyethylene glycols (PEGs) based screens [169]

3.2.3 *Cd*GSTM1 crystallization

Purified *Cd*GSTM1 protein was concentrated to 10 mg/mL in a buffer containing 10 mM HEPES and 150 mM NaCl (pH 7.0). Crystallization trials were set up with the hanging drop vapor diffusion method. Each drop consisted of 2.0 µL of protein solution mixed with 2 µL of the reservoir (Tris-HCl 100 mM, PEG 4000 24% (w/v), sodium formate 0.2 M, pH 8.6). The plates were left for equilibrium at 16 ℃. N-nitrobenzyl-glutathione (Nb-GSH) was used to a final concentration of 2.5 mM from a stock solution (100 mM in 0.1 M KNa phosphate buffer, pH 7.0). Similarly, under the same conditions, the crystallization of free *Cd*GSTM1 was carried out (Fig.11c).

3.2.4 hGSTM1-1 crystallization

Prior to crystallization, hGSTM1 protein was concentrated to 10mg/mL in buffer solution containing 10 mM HEPES (pH 7.5), 150 mM NaCl, NaN_3 0.002% w/v. Crystals were produced with the hanging-drop vapor diffusion method using a reservoir solution of HEPES 0.1 M (pH 7.2), PEG 6000 18% (w/v) and ammonium chloride 0.2 M (Fig.11d).



a. **b.** **c.** **d.**

Figure 11. Crystals of the GST enzymes used in the thesis **(a)** *Am*GSTF **(b)** *Lr*GSTF (**c)** *Cd*GSTM1 (**d)** hGSTM1-1.

3.3 Data collection and processing

3.3.1 *Am*GSTF data collection and processing

Data were collected on the P14 beamline (PETRA III, DESY, Hamburg) with an EIGER 16 M detector at cryogenic temperatures. Glycerol 5% v/v was added to the crystallization solution for cryoprotection. The total exposure time was 20s and 2500 diffraction frames were collected at 0.9762 Å wavelength (exposure time per frame 0.008 s) to 1.33 Å resolution. The data processing was carried out with EDNA [170] and scaling with AIMLESS [171]. The crystal was found to belong to the P1 space group with two molecules in the crystallographic asymmetric unit as judged by the Matthews coefficient [172].

3.3.2 *Lr*GSTF data collection and processing

Data were collected on the P14 beamline (PETRA III, c/o DESY, Hamburg, Germany) with an EIGER 16M detector. Data were collected to 1.90 Å resolution with a total exposure time of 28s at 0.9762 Å wavelength and an exposure time per frame of 0.021s (total number of frames 1333) under cryogenic temperatures with 20% glycerol as cryoprotectant. Data processing was carried out with XDS [173] followed by merging and scaling with AIMLESS [171] through the EDNA pipeline [170].

3.3.3 *Cd*GSTM1 data collection and processing

X-ray data collection was carried out on the P13 beamline at PETRA III (DESY, Hamburg) at cryogenic temperature (100K) from crystals frozen with liquid N_2 under cryogenic temperatures in the presence of 20 % v/v glycerol as cryoprotectant.

3.3.4 hGSTM1-1 data collection and processing

X-ray diffraction data were collected on the P13 beamline at PETRA III (DESY, Hamburg) at cryogenic temperature (100K) from flash cooled crystals with liquid N_2 in the presence of 20% (v/v) glycerol as cryoprotectant. A Dectris Pilatus 6M detector was used and the total exposure time was 120s for 3000 frames (0.04 s per frame). The data were processed with XDS and AIMLESS.

3.4 Structure determination

3.4.1 *Am*GSTF structure determination

The CCP4i2 suite [174] was used for the structure determination of *Am*GSTF. An automated search model was generated and automated molecular replacement (MR) was performed with MrBUMP [175] incorporated within the CCP4i2 suite. MrBUMP is an automated scheme for MR, which requires a proper input of sequence files and reflection data related to the structure. The program automatically searches for homologous structures, creates suitable search models from template structures, and performs molecular replacement. MrBUMP can also test the structure solution with some rounds of restrained refinement. Although automated MrBUMP uses several existing programs to iterate different steps, MR itself is performed in PHASER or MOLREP [176] and sequence alignment is performed in ClustalW. A total of six search models were generated and the best solution was derived from GST-III from *Zea mays* (PDB id 1aw9) template (sequence identity 48%). The best solution was used for automated model building with Buccaneer [177] in the CCP4i2 suite. The output from Buccaneer is a coordinate (PDB) file of the model which is visualized in COOT [178] for further manual building and refinement. Each step of model building was followed by manual inspection of $2|F_o|-|F_c|$ and $|F_o|-|F_c|$ electron density maps and refinement via REFMAC [179]. The final refined structure has R_{work}/R_{free} value of 0.131/0.163 (5% of the reflections set aside for R_{free} calculations). Manual building to the model included fixing the incorrect rotamers, sidechain flips in Asn and Gln residues, and addition of ligands and waters carried out in COOT. The addition of GSH to the structure was carried out based on $|F_o|-|F_c|$ electron density difference maps. Furthermore, an additional density was seen near GSH at

the active site and was subsequently modeled as succinic acid (SIN) from the SPG (succinic acid, sodium dihydrogen phosphate, and glycine) buffer solution. Cys120 was found in the oxidized state and modeled as S-hydroxycysteine. In addition, a globular density was observed next to the sulfur atom of GSH which was ultimately modeled as glutathione sulfenic acid (GSOH) (GS8 in PDB 3-letter code) in both the subunits. The summary of crystallographic data collection and final refinement statistics are shown in Table 6.

3.4.2 *Lr*GSTF structure determination

The structure was determined with molecular replacement using PHASER [180] as implemented in PHENIX v.1.17.1-3660 [181]. A homologous structure search was performed manually and a search model was constructed from *Zea mays* GST (PDB id 1bye, 62.4% seq. identity) with SCULPTOR [182]. SCULPTOR is used to truncate the search models and it typically operates by trimming loops and side chains from the search model with poor sequence alignment to the target sequence. According to the solvent content calculations, six molecules were expected in the asymmetric unit. All six molecules of *Lr*GSTF were located and the resulting solution was set for automated building followed by manual rebuilding in COOT and refinement in PHENIX. Among all the molecules, molecule C displayed electron density for all the residues, including the six His residues from the purification tag at the C-terminal. The final R_{work}/R_{free} of 0.165/0.203 was achieved with the iterative refinement process and rebuilding (Table 6). Based on the $|F_o|-|F_c|$ electron density difference map, a molecule of S-(p-nitrobenzyl)-glutathione (Nb-GSH) was added to five of the *Lr*GSTF subunits while in the sixth subunit (subunit D), a molecule of GSH was added.

3.4.3 *Cd*GSTM1 structure determination

The *Cd*GSTM1 structure was solved with molecular replacement using the ligand-free human glutathione transferase M1a-1a (PDB id 1gtu) as a search model. MrBUMP was employed resulting in an initial solution with R_{work}/R_{free} of 0.29/0.33 after 30 cycles of restrained refinement in REFMAC. The refinement was finally carried out in PHENIX using simulated annealing at 1000 K and maximum likelihood as a target function. Refinement was alternated with model building and validation steps. The protocol yielded final structures with R_{work}/R_{free} of 0.195/0.252 for *Cd*GSTM1- Nb-GSH and 0.206/0.273 for the *Cd*GSTM1-GSH complex (Table 6).

3.4.4 hGSTM1-1 structure determination

The structure was determined by molecular replacement using the structure of ligand free hGSTM1-1 (PDB id 1gtu) as a search model in PHASER [180]. Refinement was carried out with PHENIX v. 1.17.1-3660 [181] using simulated annealing and maximum likelihood as a target function. Structure visualization and rebuilding were carried out with COOT [178]. Structure validation was performed using COOT and PHENIX validation tools (Table 6).

3.5 Structure validation and analysis

Structure validation was initially carried out with validation tools in COOT [178]. The MolProbity package [183] implemented in the PHENIX suite was also used to identify and resolve problematic regions such as steric clashes, basic geometry, and phi/chi angles in peptide bonds. In addition, the MolProbity web server (http://molprobity.biochem.duke.edu/) for structural deposition was used in the final steps to analyze the structure quality and refinement statistics.

Furthermore, the PDBePISA web server (https://www.ebi.ac.uk/pdbe/pisa/) was used to analyze enzyme interfaces and assemblies [184]. The PDBeFold (https://www.ebi.ac.uk/msd-srv/ssm/) webserver was used in the investigation of new GST structures for similarity against the entire PDB archive and to generate a 3D structure alignment [185]. The PDBsum [186] webserver was used to examine the topology of the proteins, particularly the secondary structure elements (SSEs) and the contributing residues, protein-substrate interactions, and cavities and grooves in the protein structure. Multiple sequence alignment was performed in ClustalW [187]. Structure-based superpositions and structure-based multiple sequence alignment were performed in UCSF Chimera [127] with the embedded MatchMaker tool. The sequence alignment outputs were formatted and colored using ESPript (https://espript.ibcp.fr/) for the visualization [188]. All the final structures were manually analyzed and figures were generated in UCSF Chimera [127], PyMOL [189], and Discovery Studio (BIOVIA, Dassault Systèmes).

Table 6. Data collection and refinement statistics.

Data Collection	***Am*GSTF**	***Lr*GSTF**	***Cd*GSTM1-NbGSH**	***Cd*GSTM1-GSH**	**hGSTM1-1**
Resolution range (Å)	41.3–1.33 (1.37–1.33)	66.68–1.90 (1.93–1.90)	66.68–2.05 (2.11–2.05)	95.9–2.55 (2.64–2.55)	52.8–1.59 (1.65–1.59)
Space group	P1	$P2_1$	$P2_1$	$P2_12_12_1$	$P2_1$
Unit cell dimensions a, b, c (Å)	50.3, 50.7, 51.6	92.6,98.5, 95.9	50.4, 177.4, 59.4	50.8, 150.2, 191.8	49.1, 211.2, 49.4
α, β, γ (°)	97.7,110.2,110.9	90,109.1,90	90, 115.0, 90	90,90,90	90,116.2,90
No. of measurements	631790 (4413)	486200 (18253)	192763 (11609)	323629 (30311)	310250 (18421)
Unique reflections	93011 (1536)	126068 (5801)	55058 (4146)	49030 (4404)	112127 (8772)
Completeness (%)	94.9 (45.8)	98.6 (91.5)	93.5 (90.5)	100 (100)	93.36 (73.05)
Mean I/sigma (I)	17.3 (1.7)	7.4 (1.4)	5.0 (1.4)	10.3 (0.9)	10.6 (2.0)
Wilson B-factor ($Å^2$)	16.0	21.15	33.18	49.03	20.74
R_{meas}	0.054 (0.69)	0.113 (0.862)	0.175 (1.593)	0.188 (2.369)	0.064 (0.466)
$CC_{1/2}$	1.0 (0.70)	0.997 (0.843)	0.984 (0.305)	0.997 (0.348)	1.0 (0.78)
Refinement					
Reflections (work/test)	88547/4463	125811/1829	54874/2627	48915/2393	112117/5544
R_{work}/R_{free}	0.131/0.163	0.165/0.203	0.195/0.252	0.206/0.273	0.1983/0.2358
No. of non H-atoms	4008	11820	7762	7564	8454
Macromolecules	3486	10508	7232	7232	7258
Ligands	70	239	127	81	7
Solvent	438	1073	403	251	1149
RMS in bonds (Å)	0.012	0.006	0.008	0.008	0.007
RMS in angles (°)	1.53	0.81	0.92	0.96	0.90
Ramachandran favored (%)	97.4	96.84	96.1	95.9	98.50
outliers (%)	0.5	0.39	0.5	0.7	0.12
Average B-factor ($Å^2$)	22.1	36.28	55.6	59.1	25.39
PDB Ids	**6riv**	**6zb6**	**7opy**	**7opz**	**7beu**

Statistics for highest-resolution shell are shown in parenthesis.

4 Results and Discussion

4.1 Structural studies of *Alopecurus myosuroides* GSTF

4.1.1 Crystal structure of *Am*GSTF

The quaternary structure of the enzyme is a homodimer. The homodimer assembly consists of two identical subunits, namely A and B, which are related by a 2-fold symmetry (Fig.12a and b). Each subunit comprises 217 amino acids and the root mean square deviation (RMSD) between the subunits is 0.22 Å. The RMSD value suggests that the A and B subunits are almost similar though subtle differences exist. The *Am*GSTF structure displays the characteristic canonical GST fold, which consists of two distinct domains: a small N-terminal domain (residues 1-79) and a larger C-terminal domain (residues 91-214). The N-terminal domain adopts a typical thioredoxin fold, in which four-stranded β-sheets (β1, β2, β3, β4) are flanked by two large α-helices, α1 (residues 12-26) and α3 (residues 67-79). The N-terminal domain also consists of a much smaller and irregular α2-helix (residues 44-47) which runs slightly outside and above the four β-strands. The β-strands are arranged in the order β1 (residues 4-8), β2 (residues 28-38), β3 (residues 56-60), and β4 (residues 63-66) with the β3 strand oriented antiparallel to other strands. On the other hand, the C-terminal domain is an exclusive α-helical domain and comprises six α- helices, namely α4 (residues 92-128), α5 (residues 133-156), α6 (residues 167-181), α6′ (residues 185- 190), α7 (residues 192- 202) and α8 (residues 204-214). Helix α4 is the longest helix that runs across the entire enzyme monomer and displays a distinctive kink in middle, possibly an evolutionary feature to cover the active site of the enzyme. In addition, the α4 helix almost runs parallel with the α5 helix. All helices are arranged in a right-hand spiral architecture. The C-terminal domain is connected to the N-terminal domain between the α3 and α4 helices with a short linker sequence (residues 80-91). The molecular surface view of the *Am*GSTF structure revealed that the C- and N-terminal domains form a cleft in the middle of the molecule that serves as the active site (Fig. 12c). The active site can accommodate glutathione and various xenobiotic substrates.

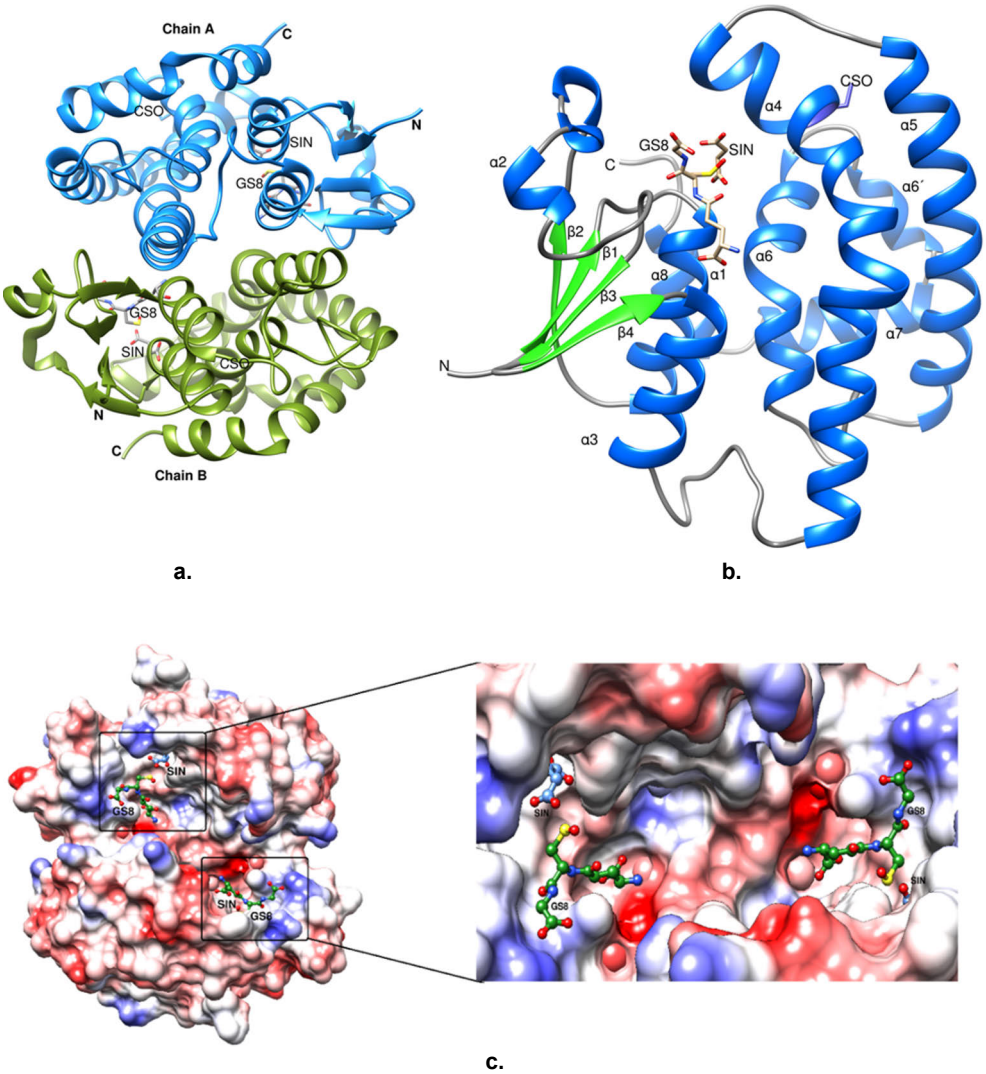


Figure 12. a. Ribbon representation of subunits A and B of the *Am*GSTF dimer. Subunits A and B are colored light blue and olive, respectively. **b.** Subunit A of the dimer. The subunit is colored according to the secondary structure elements. The α-helices are colored light blue, β-strands are colored green, and the loops are shown in grey color. Bound GS8 and SIN molecules are shown in stick representation, in the active site, and colored by elements. **c.** The molecular surface of an *Am*GSTF dimer is colored according to electrostatic potential (blue = positive and red = negative). The close-up view of the active site cleft with bound ligands (ball and stick representation) is shown on the right.

4.1.2 Subunit-subunit interface analysis

The enzyme interface analysis revealed that, of the total 217 residues in each subunit, 194 residues from subunit A and 191 residues from subunit B are present at the

enzyme´s surface. Out of the total number of residues on the surface, only 32 residues from subunit A and 33 residues from subunit B are involved in interface interactions, which is 15% of the total number of residues in the enzyme assembly. The *Am*GSTF interface area is 1407.5 $Å^2$ and the solvation free energy gained upon the interface formation is -20.9 kcal/M, suggesting good enzyme stability. The interface area is slightly bigger than other Phi class GSTs (1axd, 1371.6 $Å^2$; 6f01, 1297.8 $Å^2$; 5ey, 1320 $Å^2$; 5f06, 1377.0 $Å^2$; 5f07, 1334.5 $Å^2$). There are predominantly two types of interactions identified in the subunit interface: salt bridges and hydrogen bonds. Both types of these interactions are responsible for the stabilization of the protein-protein interface. A total of 10 salt bridges and 5 hydrogen bonds are identified between the *Am*GSTF subunits. The most prominent salt bridge interaction at the interface occurs between Asp^{99} and Arg^{77}. This interaction probably administers the stability, structural communication, and synergism observed in the kinetic studies of *Am*GSTF (PAPER I).

4.1.3 The glutathione (GSH) binding site (G-site)

The *Am*GSTF structure was solved with a bound glutathione sulfenic acid (GSOH; GS8 in PDB 3-letter identification code) and succinic acid (SIN) in each of the subunits. Their presence in the structure allowed the delineation of G- and H-site residues that constitute the active site cavity. GSOH is assumed to be a probable intermediate of the GST´s glutathione peroxidase catalytic activity and it was possibly formed during crystallization following the oxidation of the -SH group of GSH. The GSOH corresponds to an intermediate that is formed between the reaction of GSH and hydroperoxides [190]. In the G-site, the glutamate portion of GS8 interacts with the Glu^{67} and Ser^{68} as in other GST structures. The main chain amide and hydroxyl groups of Ser^{68} are located 2.89Å and 2.63Å from the γ-Glu moiety and are involved in hydrogen-bond interactions. The complete conservation of Glu^{67} and Ser^{68} signifies their critical role in GSH binding. Pro^{56} and Asn^{14} are also involved in weak van der Waals interactions with the glutamate group of GS8. Similarly, Arg^{19} and Arg^{69}, which are highly conserved, also interact with the γ-Glu moiety and possibly play a role in the correct orientation of GSH in the active site. The cysteinyl moiety of the bound GSOH forms a hydrogen bond with Ile^{55} and the -SOH group is in the H-bonding distance (3.18 Å) with the hydroxyl group of Tyr^{118}. On the other hand, the catalytic residue Ser^{12} is positioned close to the GSOH moiety in the α1-helix. Similarly, the glycyl moiety of the bound GSOH forms an H-bond with Lys^{42} and Gln^{54}. The glycyl portion also interacts via weak van der Waals forces with Gly^{53}, located in the α2-β3 region (Fig. 13a and b). Most of the residues involved in G-site binding are highly conserved, implying the importance of accurate orientation of glutathione in the enzyme cleft.

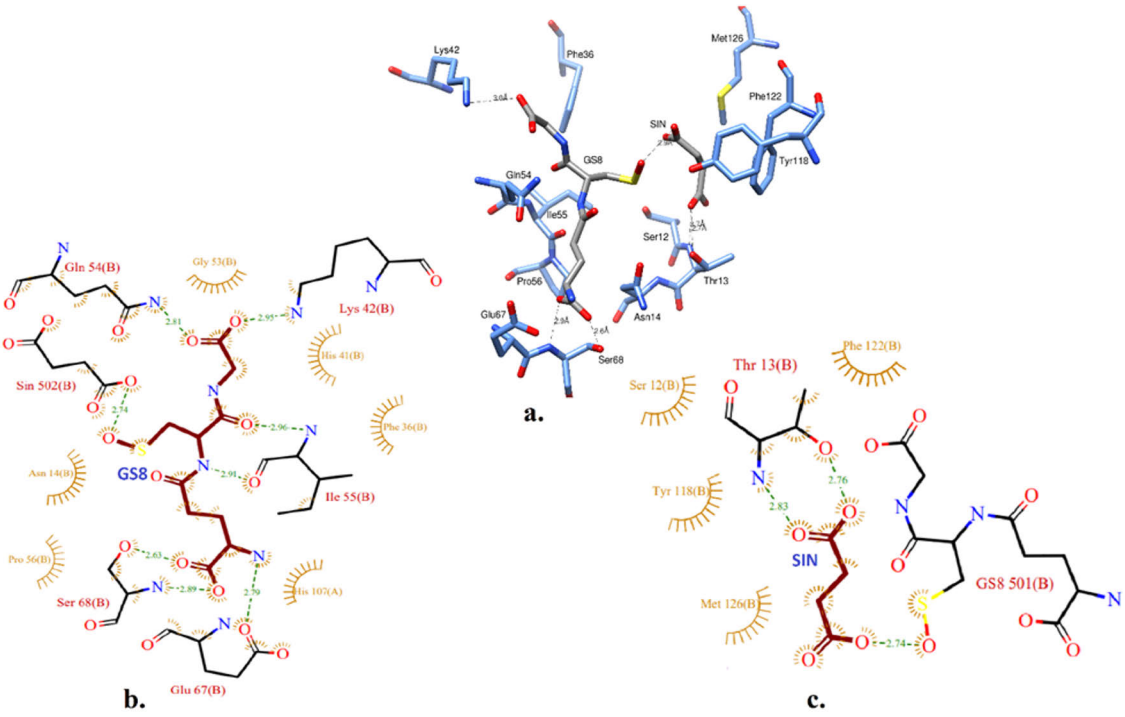


Figure 13. a. GS8 and SIN (colored in grey) in the *Am*GSTF active site. The active site residues are colored in light blue. **b.** Details of GS8 interacting residues in the G-site **c.** SIN interacting residues in the H-site.

4.1.4 The substrate-binding site (H-site)

The H-site of *Am*GSTF is located next to the G-site and formed by residues from three regions, namely, Met^{11}-Asn^{14} located at the end of the β1-α1 loop and the beginning of α1 helix in the N-terminal domain, residues Asp^{35}-Glu^{40} from the β2-α2 loop, and residues Tyr^{118}-Arg^{12} located at the C-terminal end of α4-helix. A molecule of succinic acid (SIN) from the crystallization buffer was found bound at the H-site. In fact, this is the first time that a negatively charged molecule is found in the H-site, suggesting that *Am*GSTF may participate in additional regulatory roles and function in primary and secondary metabolism. The H-site is a strict hydrophobic pocket created by residues Met^{126}, Phe^{122}, Tyr^{118}, Tyr^{178}, Tyr^{175}, and an STN motif formed by Ser^{12}, Thr^{13}, and Asn^{14} residues. The STN motif is part of a conserved STNV catalytic motif located in the α1 helix of the G-site. All these residues are orientated towards the center of the H-site cavity. These hydrophobic pocket residues are mostly not conserved, which probably modulates the H-site shape, size, and specificity depending upon the variability of the hydrophobic substrates. The carboxylate group of SIN at one end was found to participate in H-

bonding interactions with -OH and the main chain -NH group of Thr[13] located at 2.7 Å and 2.8 Å distance respectively (Fig.13a and c). However, at the other carboxylate end of SIN, the interaction is weak with the hydroxyl group of Tyr[118] (distance 3.3Å). The SIN molecule was also found in hydrogen bonding distance with the -SOH group of the bound GS8 (distance 2.7Å). The binding residues and localization of SIN in the H-site cleft are shown in Fig. 14.

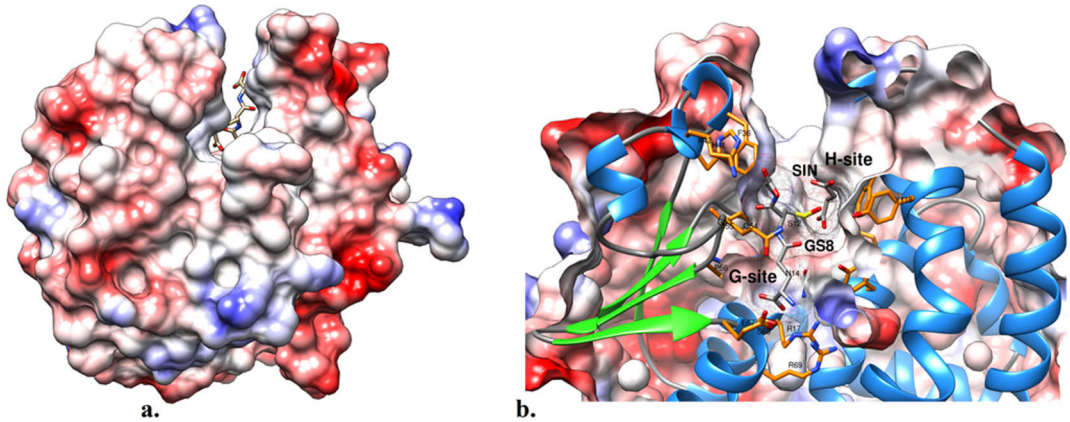


Figure 14. a. Molecular surface view of *Am*GSTF monomer. The surface is colored according to the electrostatic potential (blue=positive, red=negative). **b.** close-up view of the G- and H-sites. SIN and GS8 are depicted with mesh surface and neighboring active-site residues are labeled.

4.1.5 Comparison of *Am*GSTF with *Pt*GSTF1

The *Am*GSTF structure is superimposed and aligned with *Populus trichocarpa* phi class GST (*Pt*GSTF1) with an RMSD value of 1.62 Å (37% sequence identity for 201 aligned residues). The most significant differences were observed in three structural zones. First, the linker loop region between N- and C-terminal domains (α3-α4); Second, the α4 helix and third the β2-α2 region (Fig. 15a, b). Importantly, the upper part of the α4-helix is involved in the formation of the H-site cavity. In general, it adopts different conformations that may have an impact on the H-site properties of ligand binding. For example, the *Am*GSTF α4-helix is longer by three residues (Pro[124]-Met[126]) than its *Pt*GSTF1 counterpart while in *Pt*GSTF8 it moves away from the H-site. Furthermore, comparative studies of the α4-helix conformation with tau class GSTs [191] revealed that the *Am*GSTF α4-helix occupies the C-terminal H9 helix of tau class GSTs and provides a lid-like architecture over the H-site (Fig.16a). To complement the length and `lid-like´ design of the *Am*GSTF α4 helix, the α8-helix points away from the H-site, and thus clashes between the α4 and α8-helices are not possible (Fig. 16a).

Furthermore, as discussed earlier, the sharp kink displayed by the α4- helix at its middle results in a slight curvature of its second half over the active site. A moderately conserved Pro[111] is spotted in the proximity of the kink (Fig. 16b). However, it is substituted by an Ala and Asp in other GSTFs with similar kinks, thus Pro[111] might not solely be responsible for the kink. Nevertheless, GSTF structures with conserved Pro at this location displayed more prominent bending towards the active site. For example, in *Pt*GSTF8, which has an Asp instead of Pro at this location, the α4-helix has moved away from the H-site.

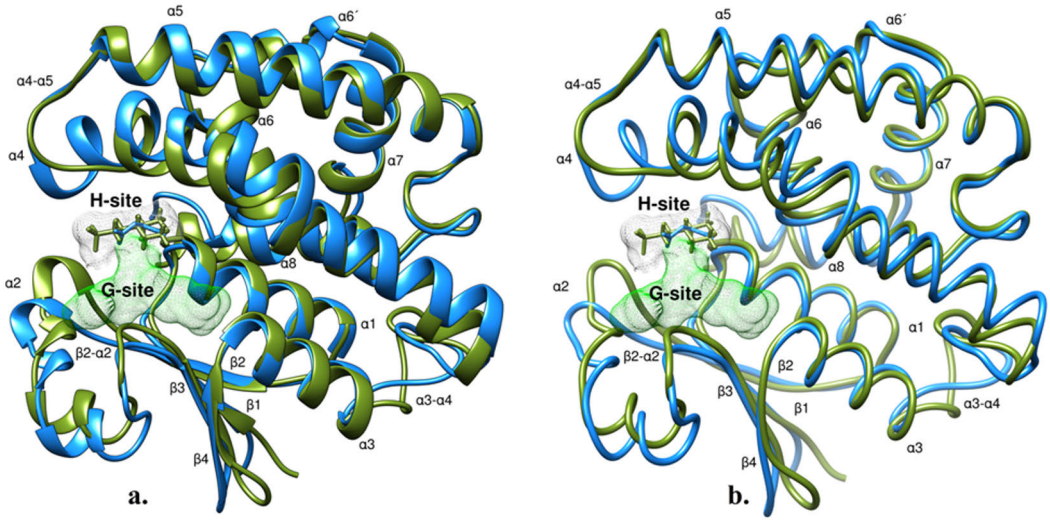
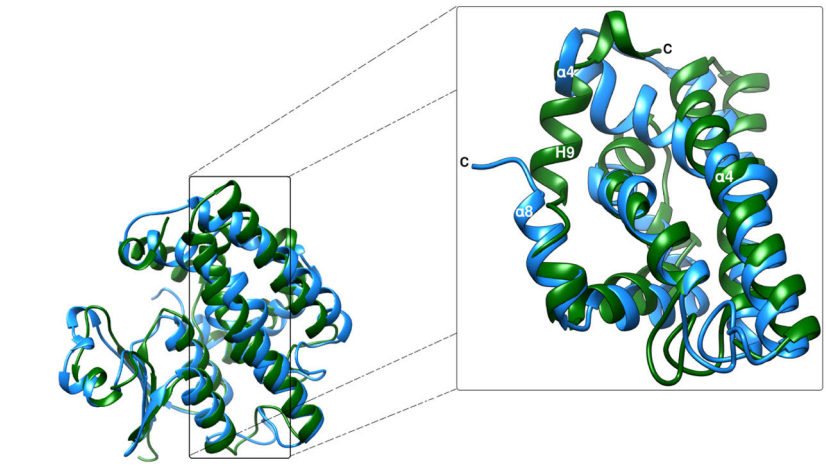


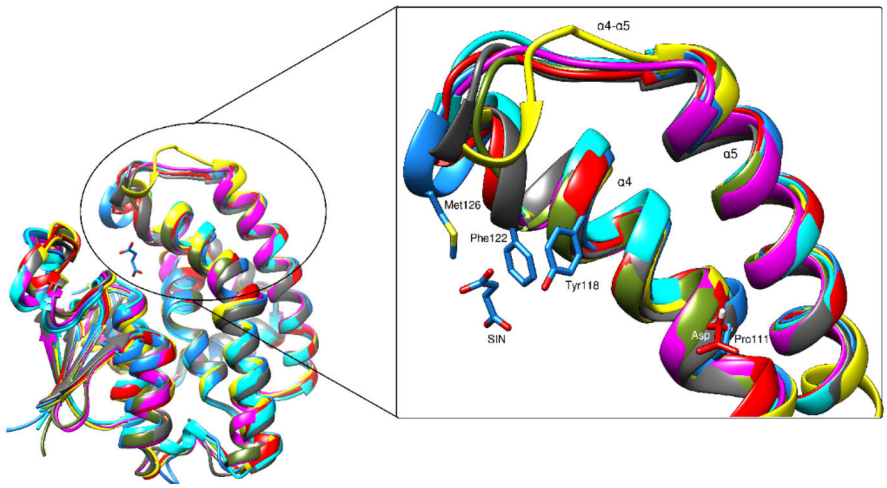
Figure 15. **a.** A ribbon representation of superimposed *Am*GSTF monomer (light blue) onto *Pt*GSTF1 (olive). The G- and H-sites are designated with mesh surfaces and labeled. The GS8 and GSH binding in the G-site is omitted for visual clarity of the active site. MES and SIN bound in the respective H-sites of *Pt*GSTF1 and *Am*GSTF are displayed in sticks and colored according to the parent monomer. **b.** A licorice representation of the aligned enzymes for good visual.

On the other hand, the SIN position in the H-site of *Am*GSTF is also found similar to 2-(N-morpholino)-ethanesulfonic acid (MES) binding in the *Pt*GSTF1 structure [192]. However, the SIN was refined with full occupancy in the *Am*GSTF structure unlike the reduced occupancy assigned to MES in the *Pt*GSTF1 structure. This suggests that SIN can occupy the active site at the same time with GSH. However, the contact between SIN and GS8 is largely via weak interactions with no covalent bonds found. MES participates in hydrogen bonding with *Pt*GSTF1 His[119], which is structurally equivalent to Tyr[118] residue in *Am*GSTF but is not highly conserved in phi class GSTs. In other phi class GSTs, the Tyr[118] in *Am*GSTF is found substituted by other residues namely Phe[118] (aromatic), His (charged residue), or Leu (non-polar residue), suggesting that it may play a vital role in ligand specificity, recognition,

and binding. In addition, Phe[122] involved in van der Waals interaction with SIN and MES molecules is replaced by other non-polar/hydrophobic residues in other GSTFs (Fig. 17), emphasizing the role of Phe[122] in substrate recognition and binding.



a.



b.

Figure 16. **a**. Superposition of *Am*GSTF (light blue) and tau class *Glycine max* GSTU4 (*Gm*GSTU4) (dark green; PDB id 4top). The magnified view of C-terminal helices of *Am*GSTF and *Gm*GSTU4 are shown at the right and labeled α8 and H9 respectively; the α4 helix is also labeled. **b**. Structure-based alignment of *Am*GSTF (light blue) with *Pt*GSTF1 (olive); *Pt*GSTF2 (magenta); *Zm*GST-I (cyan); *At*GSTF9 (yellow); *Pt*GSTF7 (gray); *Pt*GSTF8 (red). On the right, is the closeup view of an upper α4-helix region of the superimposed GSTF structures.

4.1.6 Comparison with phi class GST enzymes

The secondary structure matching showed that a 40-50% sequence identity and almost 70-90% conservation in secondary structural elements (SSEs) in *Am*GSTF with other phi class GSTs. The lowest RMSD value of 0.88Å was found with *Zea mays* GST-I (PDB id 1axd). The corresponding sequence identity and SSEs similarities were 65% and 80% respectively for the alignment. The RMSD of *Am*GSTF chain A for the chosen superimposed structures was found less than 2 Å. The structural motif ^{67}ESR, which is involved in stabilizing the γ-glutamyl moiety of GSH via hydrogen bonding and Columbic interactions, is located at the N-terminal end of α3-helix is conserved in all the structures. This again emphasizes the critically important orientation of GSH in the active site for enzyme function. The conservation of core residues is linked to the same set of catalytic properties on the same xenobiotic and thiol substrates observed in the substrate specificity studies (PAPER I). Also, another structural motif, ^{160}LGAD, a linker sequence between the α5 and α6 helices, is conserved in every structure and may play an important role in maintaining structural integrity and GST´s conserved enzyme fold (Fig. 17). Pro^{51}and Gly^{53} located at the end of α2-helix and Pro^{56} at the beginning of the β3 strand are strictly conserved in all structures and are responsible for the tight turn essential for the formation of a cleft like structure for GSH binding in the active site.

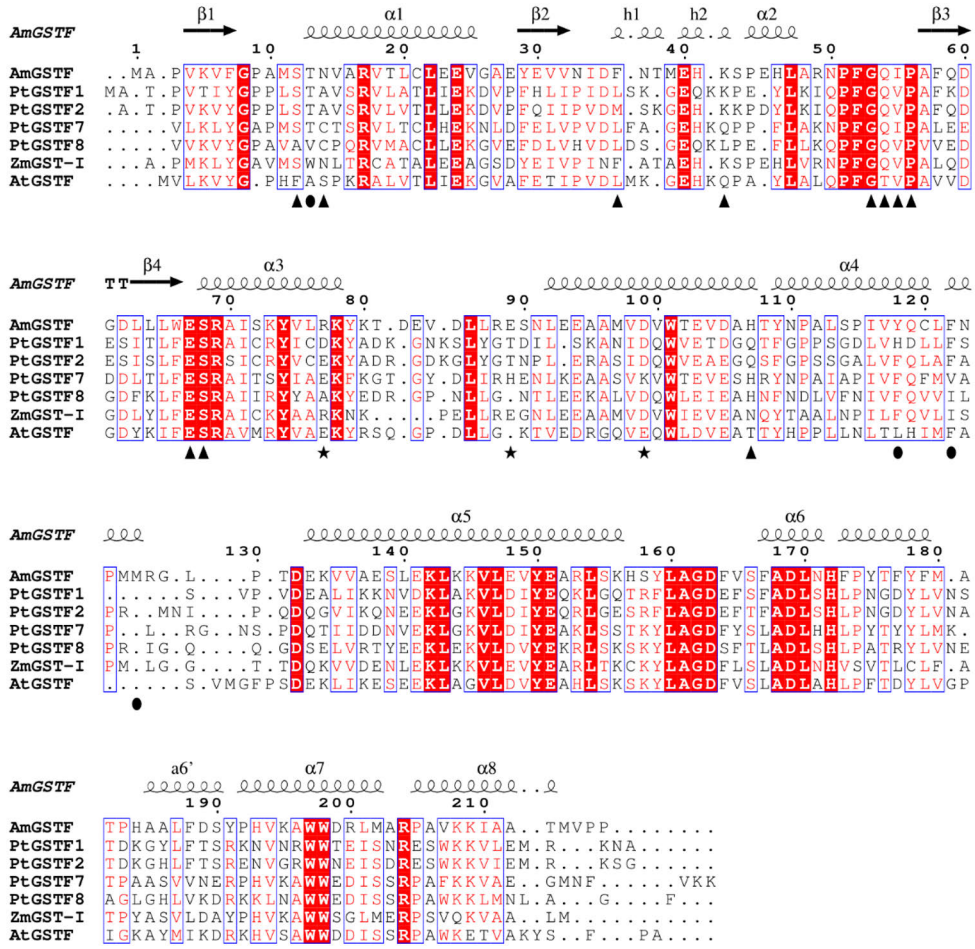


Figure 17. Structure-based sequence alignment of members of the GST phi family. The secondary structure elements of *Am*GSTF are displayed on top. The *Am*GSTF residue numbering is shown above the alignment and conserved areas are shaded. A column is framed if more than 70% of its residues are similar according to physio-chemical properties. Residues involved in GSH-binding (G-site) are shown with triangles. The residues involved in SIN binding (H-site) are in circles and those that make hydrogen bonds at the dimer interface (cut-off distance 3.6 Å) in stars. *Zea mays* GST-I, PDB id 1axd; *Populus trichocarpa* GSTF7, 5f06; *Arabidopsis thaliana* GSTF9, 6f01; *Pt*GSTF2, 5ey6; *Pt*GSTF8, 5f07; *Pt*GSTF1, 4ri6.

4.2 Structural studies of *Lolium rigidum* GSTF

4.2.1 Crystal structure of *Lr*GSTF

The crystal structure of *Lr*GSTF was determined with six molecules in the asymmetric unit, namely chains A, B, C, D, E, and F. Each of these six molecules of

*Lr*GSTF has 222, 215, 227, 215, 215, and 215 amino acids respectively. The variability in the residue numbers built into the structure is due to the flexibility at the C-terminal and the inadequate electron density. Of the six molecules, molecule C possesses the lowest flexibility, which resulted in the position of the entire C-terminal 6xHis-tag. The typical *Lr*GSTF dimers are formed by molecules A and E, F and C, and B and D. The *Lr*GSTF monomer consists of a conserved N-terminal thioredoxin-like domain (β1α1β2α2β3β4α3) and a variable all-helical C-terminal domain (α4α5α6 α6'α7α8). The G-site responsible for GSH binding and the H-site responsible for hydrophobic substrate binding are located in the N- and C-terminal, respectively (Fig 18a and b). Both the G-site and the H-site together comprise a large open active site cavity located in between the N- and C-terminal domains.

Apart from molecule D, all the active sites of the five molecules consist of a bound S-(p-nitrobenzyl)-glutathione (Nb-GSH) whereas molecule D has a bound GSH moiety instead. Similarly, a glycerol molecule is located in the G-site of all the molecules. A second glycerol molecule is spotted near the nitro group of Nb-GSH in molecule A but not in other molecules. Enzyme interface analysis shows that the largest interface exists between A–E, C–F, and B–D dimers at about 1407 $Å^2$. Apart from this, other interfaces are significantly smaller (A–D ~569.3 $Å^2$, A–C ~612.5 $Å^2$, and B–F ~469.8 $Å^2$). In addition, between molecules A–E, C–F, B–D, A–C, A–D, C–E, and B–F 9, 8, 9, 7, 2, 2, and 1 salt bridges are spotted while the number of hydrogen bonds between the subunits is 2, 1, 2, 7, 3, 1 and 3, respectively.

4.2.2 Structural comparison of *Lr*GSTF with its homologue *Am*GSTF

Comparative structural studies involved the superposition and alignment of the *Lr*GSTF monomer with the *Am*GSTF structure. The RMSD value of the alignment is 0.722 Å, the sequence identity being 89.4% with 214 residues aligned. The RMSD value suggests a close structural similarity between the enzymes (Fig. 18a and b). The alignment with *Am*GSTF showed that a total of 20 amino acids substitution are present in *Lr*GSTF. Most of the residue substitutions are at the variable C-terminal domain with 11 substitutions versus 9 substitutions at the N-terminal domain (Fig. 19). The analysis showed that the ^{12}STNV structural motif harboring the catalytic residues is well conserved in the α1-helix at the G-site. All the residues involved in substrate interaction at the H-site are identical as well. The most significant difference between the two structures is observed towards the end of the α4-helix, which bends slightly towards the H-site of the enzyme (Fig. 20). Also, a subtle difference is present in the connecting loop between the α3- and α4-helix. Most of the residues in this loop region are conserved except for Gly^{82} and Gly^{90} which are substituted by polar residues Asp^{82} and Ser^{90} in *Am*GSTF.

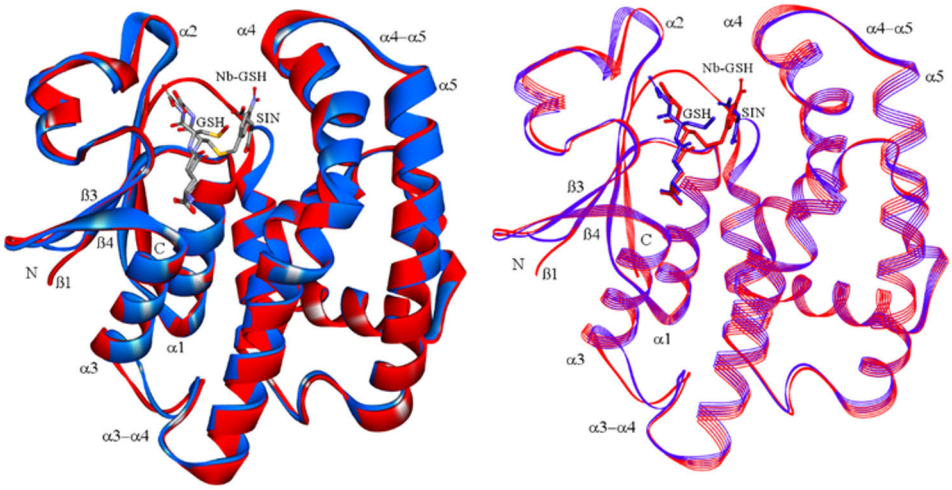


Figure 18. a. Superposition of *Lr*GSTF (red) with *Am*GSTF monomer (blue) in solid ribbon diagram. The active G- and H-sites are denoted with bound GSH/Nb-GSH and SIN respectively. **b.** Line ribbon diagram of (a) for visual clarity.

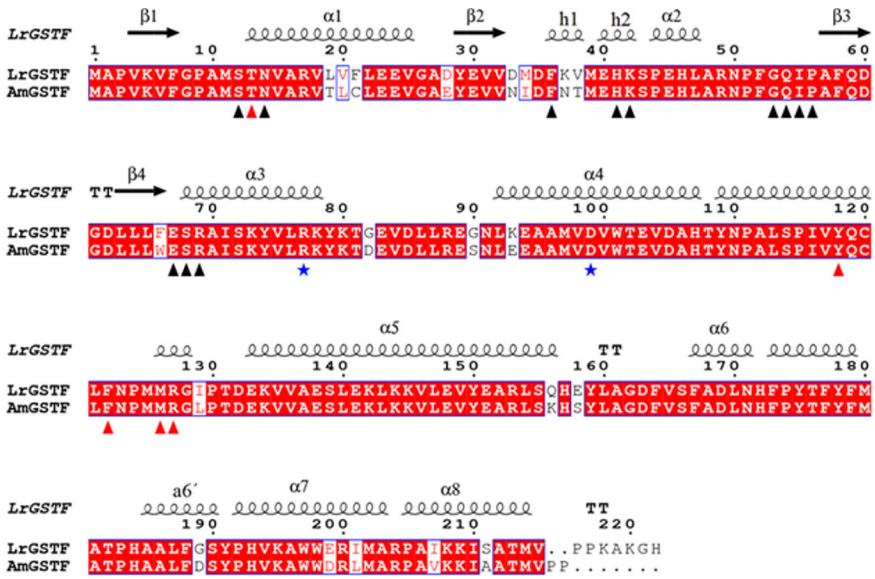


Figure 19. Structure-based sequence alignment of *Lr*GSTF and *Am*GSTF monomer. The *Lr*GSTF numbering is shown above the alignment. The conserved areas are shaded red. Residues involved in Nb-GSH binding are denoted with triangles. The red triangles represent the residues interacting with the N-nitro benzyl moiety of Nb-GSH in *Lr*GSTF. These residues also correspond with residues interacting with SIN in *Am*GSTF. The blue stars depict residues that participate in hydrogen bonds at the dimer interface.

The active site of the enzyme in both the *Lr*GSTF and *Am*GSTF structures is located in a cleft between the N- and C-terminal domains. The GSH moiety is bound at the

active site in the same conformation and similar position in both structures (Fig. 20). In *Lr*GSTF, the nitrobenzyl group in Nb-GSH occupies the H-site hydrophobic pocket where the hydrophobic electrophiles usually bind, thus acting as an enzyme inhibitor. The structural superposition between *Lr*GSTF and *Am*GSTF allowed the inspection of conformational changes upon substrate binding. The residue Arg^{127} present at the end of the α4-helix forms an H-bonding interaction with the nitrobenzyl group of Nb-GSH. In the case of the *Am*GSTF structure, the corresponding Arg residue points away from the enzyme cavity, almost like an open cap (Fig. 20). Similarly, Met^{126} is oriented away from the active site cavity upon the binding of bulky nitrobenzyl group in *Lr*GSTF which otherwise is pointed towards the active site cavity in *Am*GSTF. Also, Met^{126} is actively involved in van der Waals interactions with succinic acid (SIN) molecule in *Am*GSTF and the corresponding nitrobenzyl ring of Nb-GSH in *Lr*GSTF (Fig. 20). However, the orientation of Met^{126} and Arg^{127} residues were also found similar in *Lr*GSTF molecule D which only has GSH in the active site, thus these residue arrangements cannot be attributed solely to the inhibitor binding.

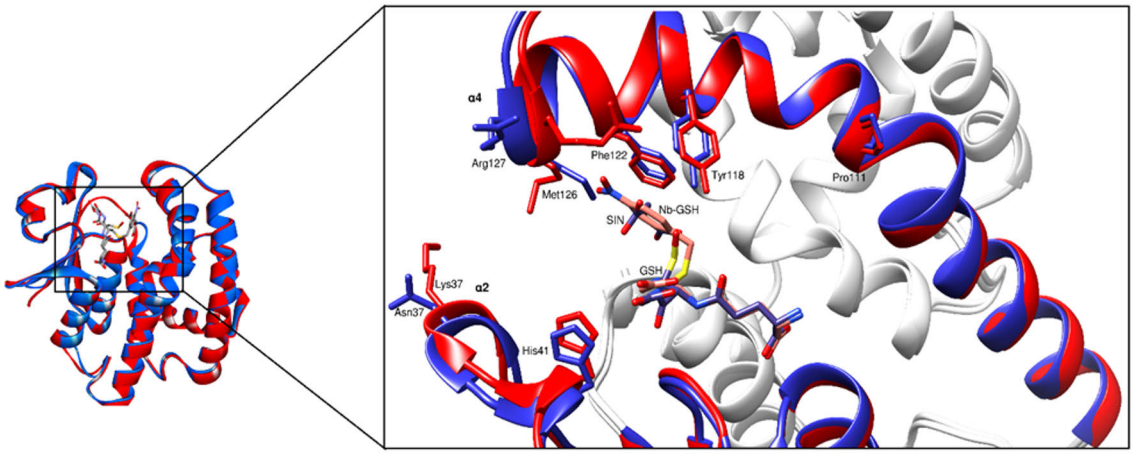


Figure 20. Close up view of the α4 and α2 helices comparison in *Lr*GSTF (red) and *Am*GSTF (blue) structures. The ligands bound in the active sites are colored according to the parent molecule. The H-site of *Lr*GSTF is occupied by the nitro-benzyl ring of Nb-GSH and is bordered by residues in the α4-helix, which is comparatively shorter in *Lr*GSTF than in *Am*GSTF. The conformational changes in the α4-helix in *Am*GSTF are shown to accommodate succinic acid (SIN) in the active site.

Furthermore, the nitrobenzyl ring of Nb-GSH in *Lr*GSTF is involved in the formation of two π-π interactions with Phe^{122} and Tyr^{118} residues of the enzyme. The orientation of Phe^{122} was found similar in both enzymes, probably owing to the bound SIN in *Am*GSTF. In molecule D of *Lr*GSTF, Phe^{122} points towards the H-site

but upon Nb-GSH binding, the side chain of Phe^{122} swings away from the active site cavity to provide space for the nitro-benzyl group of Nb-GSH. Tyr^{118} has a similar position in all molecules of *Lr*GSTF apart from molecule C, suggesting some flexibility. Thus, both the residues Phe^{122} and Tyr^{118} may play important roles in ligand specificity and binding (PAPER II). Besides, other residues, such as Lys^{42}, Gln^{54}, Ile^{55}, Glu^{67}, and Ser^{68}, also participate in H-bonding interactions with Nb-GSH. Residues Ser^{12}, Thr^{13}, Asn^{14}, Phe^{36}, His^{41}, Gly^{53}, and Arg^{69} are involved in van der Waals interactions with Nb-GSH (Fig. 21). These interacting residues are found strictly conserved in *Am*GSTF as well.

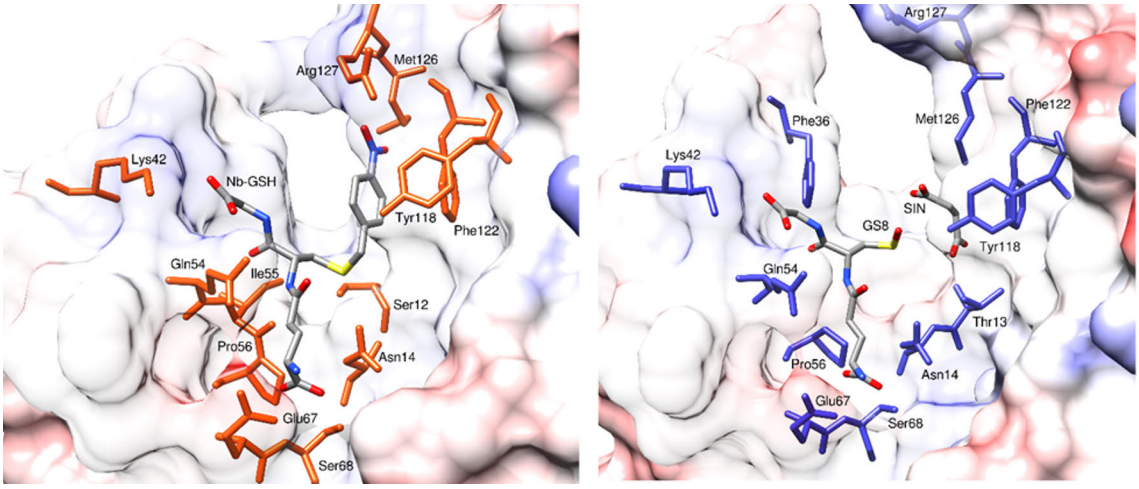


Figure 21. The active site residues comparison between *Lr*GSTF and *Am*GSTF structures. **a.** *Lr*GSTF active site cleft. The active site residues are shown (orange) and the position of bound ligand Nb-GSH is shown (grey). **b.** The *Am*GSTF active site cavity with interacting residues (blue). The bound ligands GS8 and SIN are shown in grey color.

Molecules C and D of *Lr*GSTF along with the bound Nb-GSH and GSH were superimposed with an RMSD value of 0.41 Å. Subsequently, *Am*GSTF chain A with bound GSH was also superimposed into the *Lr*GSTF molecules (Fig. 22a, b). The structural superposition showed that the binding of inhibitor Nb-GSH into the active site narrows the active site opening, thus limiting the accessibility of the substrate to the active site. The accessibility is limited by the slight movement (~ 2.0 Å) of the beginning of the α2 and the end of the α4 helix towards each other. On the contrary, molecule D with bound GSH possesses a comparatively wider gateway for substrate binding in the H-site (Fig. 22a). Besides, in *Am*GSTF with SIN bound at the H-site, the α2 and α4-helix movement narrows the active site opening but to a lesser extent than that of inhibitor bound molecule C of *Lr*GSTF (Fig. 22b). Based on structural

alignment results and RMSD values it can be concluded that the ligand binding is not accompanied by extensive conformational changes in the GSTF enzymes. However, conformational changes in individual residues appear necessary for efficient binding.

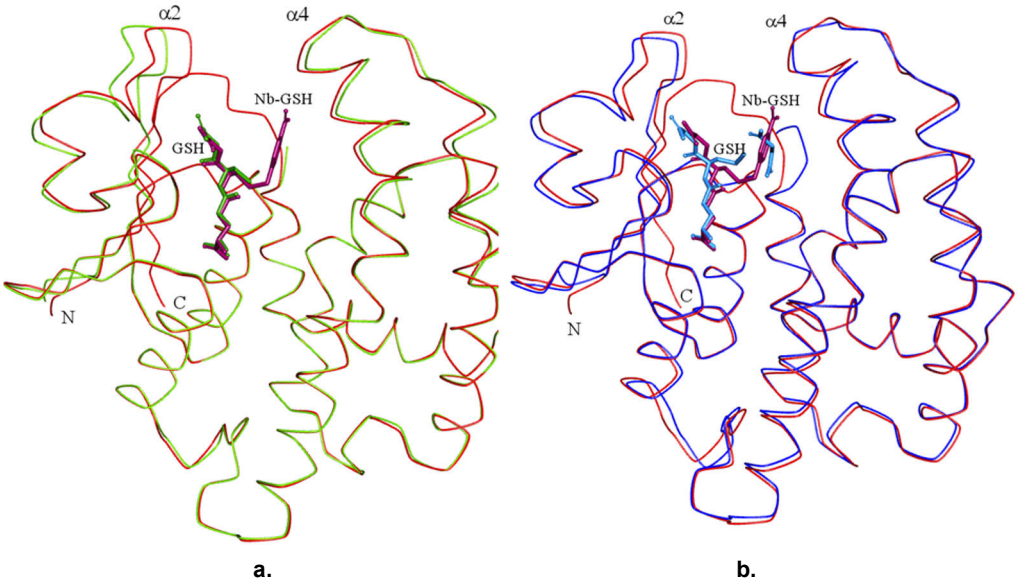
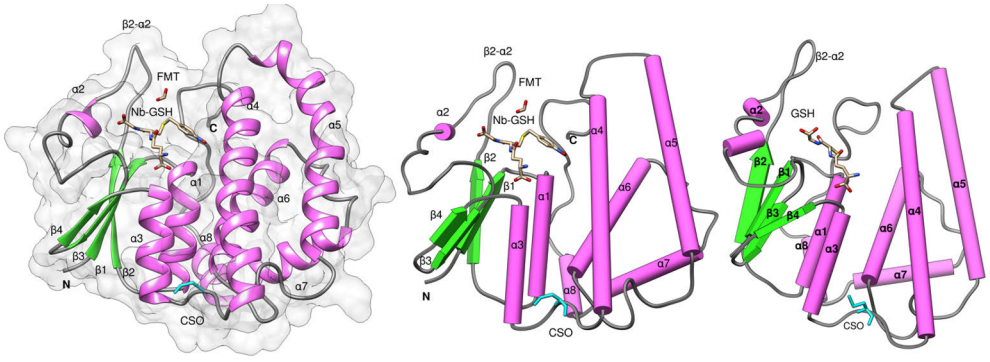


Figure 22. a. The comparison between *Lr*GSTF molecule C (red) and molecule D (green) is shown. **b.** *Lr*GSTF molecule C (red) and *Am*GSTF chain C (blue). The enzyme backbone is shown in the molecule trace diagram. Nb-GSH and GSH bound in the active sites are colored according to the parent molecule.

4.3 Structural studies of *Camelus Dromedarius* GSTs

4.3.1 Crystal structure of *Cd*GSTM1-1

The three-dimensional enzyme structure is similar to other GST classes discussed earlier. The enzyme crystallized with four molecules in the asymmetric unit, namely A, B, C, and F. The typical *Cd*GSTM1-1 dimer is formed by molecules A and C, B and F. Each of the *Cd*GSTM1-1 molecules comprises 217 amino acid residues. The crystal structure of *Cd*GSTM1-1 was determined with bound substrate GSH or a substrate analogue Nb-GSH in the active site (Fig. 23).



a. **b.** **c.**

Figure 23. a. Ribbon representation of chain A of *Cd*GSTM1 dimer; molecular surface for the monomer is shown in gray. The monomer is colored according to the secondary structural elements (α-helices in magenta, β-strands in green, and loops in dark gray). CSO and the active site ligands, Nb-GSH and FMT are shown as sticks and colored according to atom type. **b.** Cylindrical representation of Nb-GSH bound *Cd*GSTM1 monomer. **c.** GSH bound to the *Cd*GSTM1 monomer.

4.3.2 Interface analysis and comparisons

The interface comparison of *Cd*GSTM1-1 with other mu-class GSTs from *Homo sapiens* (PDB id 1xw6, 1ab6, 4gtu), *Mus musculus* (2dc5), *Gallus gallus* (1gsu), *Litopenaeus vannamei* (5an1) and *Rattus rattus* (6gst) revealed similar interface area of 1249.2 Å^2, 1383.6 Å^2, 1353.6 Å^2, 1358.8 Å^2, 1212.1 Å^2, 1157.5 Å^2, and 1328.5 Å^2 respectively. The interface area of *Cd*GSTM1-1 was found almost similar at 1379.5 Å^2. In the A–C dimer, 203 residues from subunit A and 201 from subunit C are present at the surface. Of the surface residues, 36 and 35 residues from subunits A and C (approximately 16% of total residues) are involved in interactions at the dimer interface. The solvation free energy gain upon the formation of the A–C interface is -85 kcal/M. Similarly, at the B–F dimer interface, 203 residues each from B and F monomers are present at the surface. Of the surface exposed residues, 39 and 34 residues from chains B and F (approximately 16 of the total residues) take part in the interface interactions. The solvation free energy gain from the B-F interface formation is -7.8 kcal/M. On the other hand, 22 salt bridges and 18 hydrogen bonds are spotted at the A–C interface while 20 salt bridges and 17 hydrogen bonds are marked in the B–F interface.

4.3.3 Structure comparison with human GSTM1 and other GSTs

The *Cd*GSTM1-1 sequence identity with other mu-class GSTs ranges from 50-80% but with a high level of secondary structure elements similarity (90-100%) resulting in root mean square deviations (RMSDs) between 0.44-0.85 Å (Table 7).

Table 7. *Cd*GSTM1-1 comparison statistics with other mu-class GSTs.

μ-GSTs	Organisms	PDB entries	RMSD values (Å)
hGSTM1	*Homo sapiens*	1xw6	0.44
hGSTM2	*Homo sapiens*	2ab6	0.53
hGSTM4	*Homo sapiens*	4gtu	0.73
Mu(GSTM7)	*Mus musculus*	2dc5	0.79
cGSTM1-1	*Gallus gallus*	1gsu	0.72
***Lv*GSTM**	*Litopenaeus vannamei*	5an1	0.85
***Rr*GSTM**	*Rattus rattus*	6gst	0.50

The *Cd*GSTM1-1 monomer consists of two domains. Domain I (residues 1-81) is located at the N-terminal and has an α/β topology while domain II (89-217) is located at the C-terminal and adopts an all α-helical topology. The N-terminal domain adopts a typical thioredoxin fold which is highly conserved through all GST classes. Domain I and domain II consist of GSH binding site (G-site) and the hydrophobic substrate binding site (H-site) respectively. The G-site topology appears to be highly conserved among the whole mu-class GSTs whereas the H-site is a more variable region. The G- and H-site of *Cd*GSTM1-1 together comprise the active site cleft. Each of the *Cd*GSTM1-1 monomers consists of an independent and separate active site. The ligands, Nb-GSH and formic acid (FMT) were found bound to the catalytic G-site and H-site, respectively (Fig. 23).

The structural-based sequence alignment revealed several conserved structural elements in *Cd*GSTM1-1 (Fig. 24). A strictly conserved structural motif ^{58}PNLP lining the active site cavity at the beginning of the β3 strand is present in all compared structures. While residues Pro^{58} and Pro^{61} are responsible to give a characteristic turn to the loop, Asn^{59} and Leu^{60} are involved in interactions with various polar groups of Nb-GSH. Also, Leu^{60} forms a conventional H-bond with Nb-GSH. Another conserved structural motif ^{72}QSN is present at the end of β4 and the beginning of the α3 region, which also lines the active site cavity of the enzyme. The carbonyl oxygen of Gln^{72} and the carboxyl group of a conserved Asp^{106} are likely to form an H-bond with the γ-Glu amino moiety of Nb-GSH. The Ser^{73} hydroxyl and amide groups are located at 2.6 Å and 2.8 Å, respectively, from the α-carboxyl group of Nb-GSH and are actively involved in H-bonding.

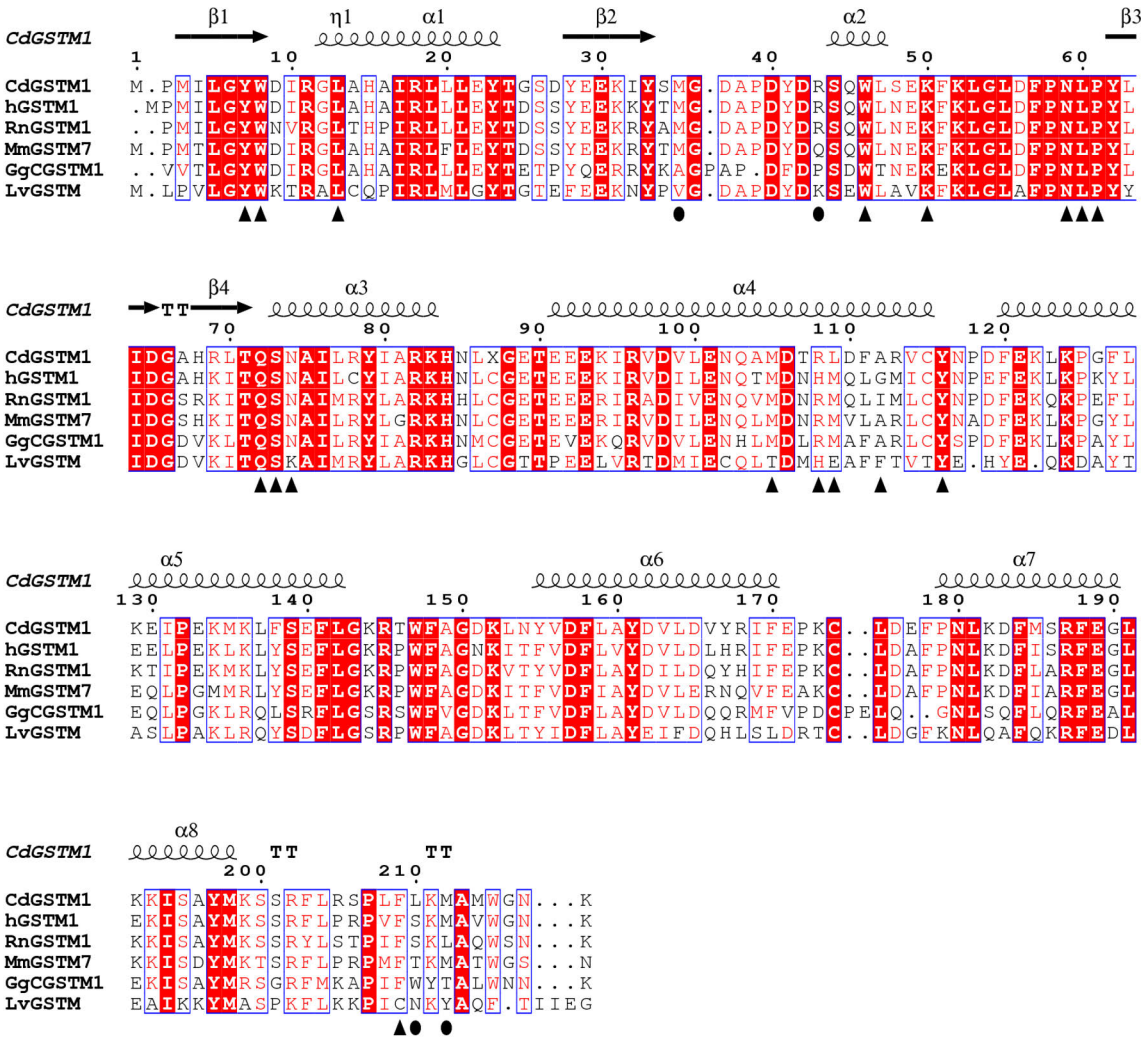


Figure 24. The structure-based sequence alignment of different members of mu-class GST families. The secondary structure elements and *Cd*GSTM1 numbering are shown above the alignment. The conserved areas are shaded, and a column is framed if more than 70% of physiological properties are similar. Residues involved in Nb-GSH binding (the G-site) are represented with triangles and residues involved in FMT binding (the H-site) are denoted in circles. *Homo sapiens* hGSTM1, PDB id 1xw6; *Rattus norvegicus* *Rn*GSTM1, PDB id 4gst; *Mus musculus* *Mm*GSTM7, PDB id 2dc5; *Gallus gallus* *Gg*CGSTM1, PDB id 1gsu; *Litopenaeus vannamei* *Lv*GSTM, PDB id 5an1.

Furthermore, a Pro^{39} residue located in the connecting loop between β2 and α2 region is highly conserved in all mu-class GSTs and is responsible to provide a characteristic extended loop conformation in the connector region. The β2-α2 region also lines the active site activity, Arg^{43} and Lys^{50} in the region are located at 3.09 Å and 3.52 Å respectively from the carboxylate glycine end of Nb-GSH and are

involved in the formation of a salt bridge and thus contribute to the stability of Nb-GSH in the active site. A conserved Tyr^{7} is in the active site β1 strand and acts as the catalytic residue for GSH activation. Similarly, Trp^{8} and Leu^{13} are conserved in all compared structures, and they contribute to Nb-GSH stabilization in the active site via H-bonding interaction with the cysteinyl moiety of Nb-GSH and pi-alkyl interactions with the nitrobenzyl ring, respectively. In addition, an active site residue, His^{107} (hGSTM1 numbering) in the α4-helix, which is present in hGSTM1-1 and has a marked influence in the catalysis of nucleophilic aromatic substitution reactions [193] is substituted by Arg^{108} in all compared rodent and avian mu-GSTs except *Litopenaeus vannamei* GST (*Lv*GSTM1). The guanidinium group of Arg^{108} is oriented away from the active site cavity and the nitrobenzyl ring of Nb-GSH. Thus, only weak van der Waals interactions are possible between the Arg^{108} residue and the bound substrate (Fig. 25 b).

The G-site for each of the compared structures are shown in Fig. 26. Although, residues that interact with GSH are strictly conserved throughout the compared structures, some differences are observed in the N- and C- terminal region of bound GSH analogs. Nevertheless, the overall distribution of charge and the van der Walls interaction of the G-site is strictly conserved throughout the compared structures, allowing for the formation of a cavity with strictly conserved polar/ hydrophobic features.

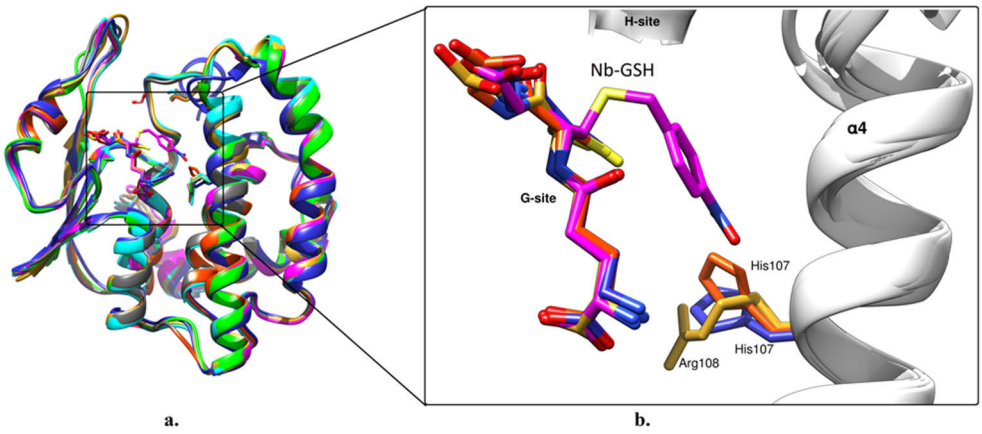


Figure 25. a. The superposition of Nb-GSH and GSH bound *Cd*GSTM1 (colored magenta and gold respectively) with hGSTM1 (orange red), *Gg*CGSTM1 (green), *Mm*GSTM7 (gray) and *Lv*GSTM1 (blue). **b.** Magnified view of hGSTM1 His^{107} (orange) and *Lv*GSTM1 His^{107} (blue) residues substituted with Arg^{108} (gold) in *Cd*GSTM1. Nb-GSH is displayed in magenta.

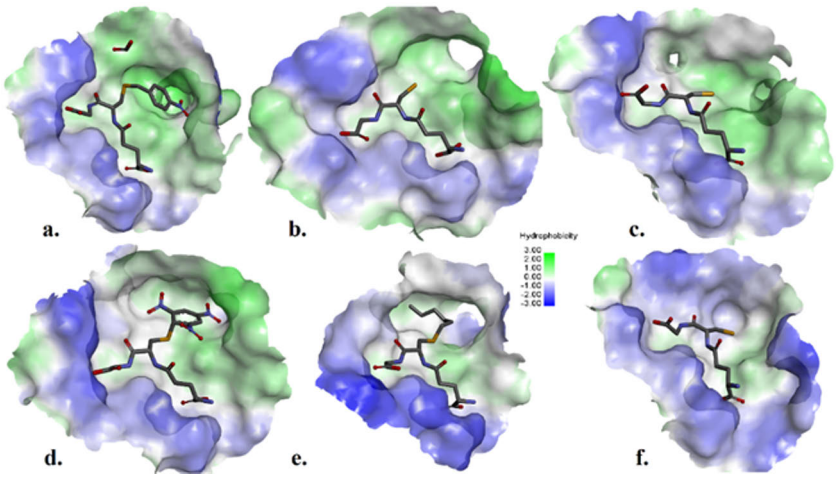


Figure 26. The active site cleft comparison in different mu-class GSTs with a bound ligand. **a.** Nb-GSH bound to *Cd*GSTM1. **b.** GSH bound to *Cd*GSTM1. **c.** GSH bound to hGSTM1 (PDB id 1xw6) [53] . **d.** GTD bound to *Rn*GSTM1-1(PDB id 4gst) [194]. **e.** GTX bound to *Gg*CGSTM1-1 (PDB id 1gsu) [195]. **f.** GSH bound to *Lv*GSTM (PDB id 5an1) [196]. The analysis indicates a similar binding mode and only a subtle difference in the active site architecture among the class.

On the other hand, the H-site of the enzyme is near to the G-site and comprises residues from the C-terminal region. The H-site in all the structures analyzed is lined mostly by hydrophobic residues such as Met^{35}, Arg^{43}, Tyr^{116}, Phe^{209}, Leu^{210}, and Met^{212}. Tyr^{116} is strictly conserved in all the mu-class GSTs compared and Phe^{209} is mostly conserved, substituted by a Cys residue in *Lv*GSTM at the same location. Leu^{210} and Met^{212} are two highly variable residues in the H-site, which probably are responsible for wide substrate recognition and binding at the H-site. For example, Leu^{210} is replaced by a Ser, Thr, and Trp residues at equivalent positions in hGSTM1, *Mm*GSTM7, and *Gg*GSTM1 respectively. Similarly, Met^{212} is replaced by a Tyr, Thr, or Leu residue at the equivalent positions in *Lv*GSTM, *Gg*GSTM1, and *Rn*GSTM1, respectively (Fig. 24). In addition, to its role in stabilizing the Nb-GSH moiety in the G-site, Arg^{43} also interacts via weak van der Waals forces with the formic acid (FMT) molecule in the H-site. However, Arg^{43} is not highly conserved in the aligned mu-class GSTs, it is substituted by a Gln, Pro, and Lys residues in *Mm*GSTM7, *Gg*GSTM1, and *Lv*GSTM respectively. The FMT molecule in the H-site is also found to interact weakly with the Nb-GSH molecule in the G-site with van der Waals forces.

Furthermore, the superposition of Nb-GSH and GSH bound *Cd*GSTM1 showed the active site cavity is almost a rigid part of the protein. The RMSD value of the alignment is 0.41 Å. Notably, a movement of the β2-α2 loop region was observed upon binding of Nb-GSH (~4 Å) which may suggest an induced-fit mechanism to

facilitate the binding of various substrates. The loop movement may favorably contribute to the overall structural stability of the enzyme–Nb-GSH complex. The induced fit mechanism operated by *Cd*GSTM1-1 probably explains the substrate concentration-dependent shift in T_m value in thermal stability studies (PAPER III). On the other hand, side chains of amino acid residues lining the cavity retain their orientations upon binding of either of the glutathione ligands except for Arg[43] and the overall enzyme structure essentially remains unchanged (Fig. 27).

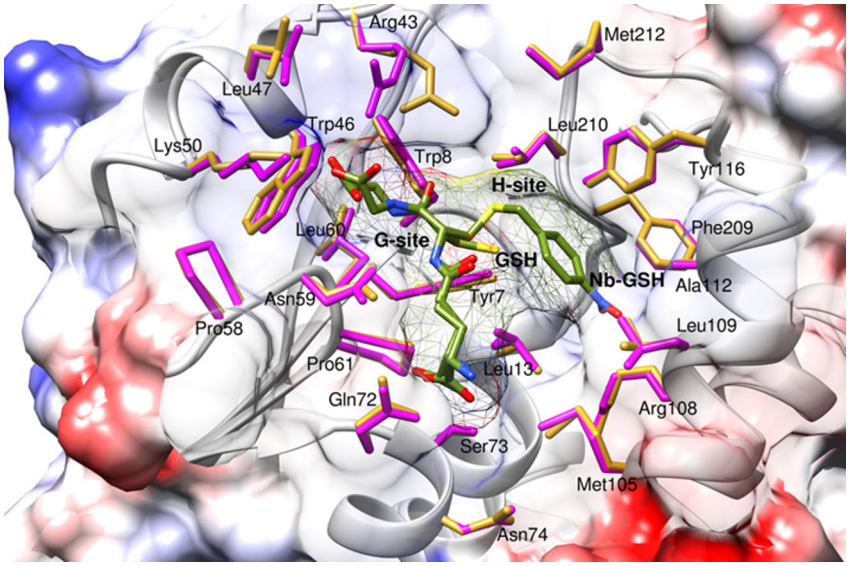


Figure 27. The active site cavity of *Cd*GSTM1. Superposition of active site residues of GSH-bound (gold) and Nb-GSH-bound (magenta) *Cd*GSTM1 structures. The aligned Nb-GSH and GSH moieties are shown as sticks (olive). The position of bound ligands is highlighted by a mesh surface. The G-and H-sites are denoted in the active site.

4.4 Structural studies of human glutathione transferase M1-1

4.4.1 Crystal structure of hGSTM1-1

The crystal structure of free hGSTM1 was solved at 1.59 Å resolution. Previously, a ligand-free crystal structure of hGSTM1 was only reported at 2.68 Å (PDB id 1gtu) [193]. The high-resolution structure reported here provides better atomic positions as well as detailed information on the induced-fit mechanism of ligand binding and catalysis (PAPER IV). The 1.59 Å structure comprises 1149 water molecules compared to only 64 in the low-resolution structure. The enzyme was crystallized in the $P2_1$ space group with four molecules (A, B, C, D) in the crystallographic

asymmetric unit. An earlier, low-resolution structure, 1gtu, was crystallized in the $P2_12_12_1$ space group but with the identical four molecules in the asymmetric unit [193]. Each of the four molecules consists of 218 amino acid residues whereas the low-resolution structure has 217 residues in each molecule.

4.4.2 Interface analysis and comparisons

The hGSTM1-1 forms a typical GSTs homodimer structure with D–A and C–B subunits. There are 36 and 34 residues from chains D and A, respectively, in the D–A interface region. The D–A interface area corresponds to 1341.9 $Å^2$. In the C–B contact region, 35 and 36 amino acids from chains C and B participate. The C–B interface area is 1313.1$Å^2$. Both dimers also have a significant number of H-bonds and salt bridges between them. The D–A interface has 13 H-bonds and 11 salt bridges while the C–B interface consists of 12 H-bonds and 13 salt bridges. On the other hand, in the low-resolution structure (1gtu), the dimer interface is like the high resolution one, however, the number of H-bonds and salt bridges is lower, possibly owing to the low resolution that affects the accurate positioning of the atoms. In addition, the interface areas between B–D and C–A subunits are much lower at 375 $Å^2$ and 350 $Å^2$ respectively with only two salt bridges in both the interfaces and 7 and 6 H-bonds between them. Also, the interaction between the dimers is mediated via helix α7 (residues 178-189).

4.4.3 Comparison of the human GSTM1-1 crystal structures

The structural alignment between the low and high-resolution hGSTM1-1 enzymes showed some subtle changes between the structures. Some differences are observed in the β2-α2 extended loop and the connector region between the α3-α4 helix (Fig. 28a). The RMSD value of the β2-α2 region is 1.36 Å which is higher than the RMSD of the overall alignment (0.65 Å). Besides, these two discrepancies in the loop regions, the Arg^{42} residue located at the beginning of the α2-helix adopts a different orientation in the high-resolution structure. Arg^{42} points closely towards the active site cavity than its low-resolution structure (Fig. 28b). This Arg^{42} is involved in electrostatic interactions with bound ligands like glutathione. Other residues such as Ser^{72} located at the top of the α3-helix, and His^{107} and Met^{104} located in the α4 helix and lining the active site cavity, also display noticeable conformational differences (Fig. 28c). Apart from these differences, other secondary structural elements and active site residues display identical conformation (Fig. 28). Other major differences between previously solved hGSTM1-1 structures and the current structure are found between the occupied and free active site cavities. Previous structures are mostly occupied by GSH (PDB id: 1xw6), Nb-GSH (PDB id: 1xwk) and GTD (1-(S-

glutathionyl)-2, 4, 6-trinitrocyclohexa-2, 5-diene) (PDB id: 2f3m) at 1.9 Å, 2.3 Å and 2.7 Å, respectively. The RMSD values upon a comparison of the high-resolution hGSTM1-1 with the ligand bound 1xw6, 1xwk, and 2f3m, are found similar (~0.36Å), suggesting a close resemblance.

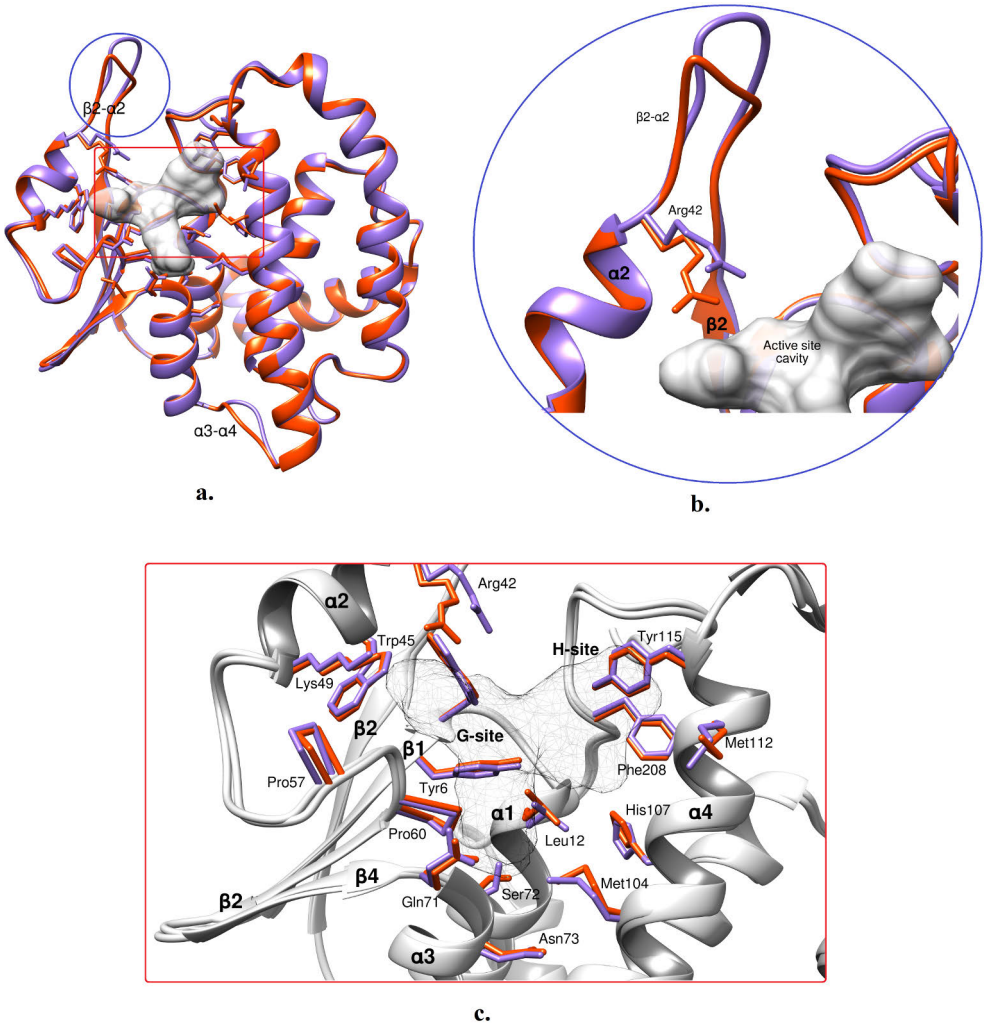


Figure 28. a. Structure superposition of high resolution (orange) and low resolution hGSTM1-1 (light purple) (PDB id: 1gtu). The area inside the red rectangle and the blue circle indicates the active site and β2-α2 loop regions, respectively**. b.** Close-up view of the loop region connecting β2-strand and α2-helix. The orientation of Arg^{42} is shown (in stick representation)**. c.** Superposition of active site residues in the high (1.59 Å) and low resolution (2.68 Å) hGSTM1-1 structures. The active site location is indicated with a mesh surface created for GTD.

5 Conclusions and Future Perspectives

Work in this thesis has led to the crystal structure determination of:

- two GSTF enzymes, *Am*GSTF (PDB id 6riv) and *Lr*GSTF (PDB id 6zb6) at 1.33 Å and 1.90 Å resolution, respectively.
- the mu-class glutathione transferases (GSTM1) from camel (*Camelus dromedarius*) with bound glutathione and nitrobenzyl glutathione (*Cd*GSTM1-GSH and *Cd*GSTM1- Nb-GSH) at 2.55 Å and 2.05 Å with PDB ids 7opz and 7opy, respectively.
- the human mu-class GST (hGSTM1) at high resolution (1.59 Å) with PDB id 7beu.

Together with molecular modeling tools, these solved structures have provided structural and functional insights into different GST classes.

The main conclusions of this study are as follows:

1. *Am*GSTF – GS8 complex

    - The crystal structure of *Am*GSTF revealed the presence of two distinct domains.
    - Ser^{12} residue in *Am*GSTF acts as the catalytic residue and was found close to the sulfur atom of GS8.
    - The H-site is lined by hydrophobic residues Met^{11}, Phe^{36}, Tyr^{118}, Phe^{122}, Met^{126}, Tyr^{178}, and Tyr^{175}, which are mostly not conserved and therefore play a role in modulating the shape, size, and specificity of the H-site.
    - The binding of the hydrophilic and negatively charged SIN molecule in the active site probably suggests the participation and function of the enzyme in primary or secondary metabolism.
    - The α4- helix provides a lid-like architecture over the H-site. The upper part of the α4-helix adopts different conformations in different GSTFs that may affect the ligand binding at the H-site.

2. *Lr*GSTF – Nb-GSH complex

   - The crystal structure of *Lr*GSTF in complex with the inhibitor S-(4-nitrobenzyl) glutathione was determined at 1.90 Å.
   - Close similarities with *Am*GSTF were revealed.
   - The nitrobenzyl group of Nb-GSH occupies the H-site of *Lr*GSTF and bends the end of the α4-helix.
   - The binding of Nb-GSH into the active site appears to narrow the active site opening, thus, displaying an induced-fit mechanism upon ligand binding.
   - Residues Tyr^{118} and Phe^{122} were found to play a vital role in ligand specificity.

3. *Cd*GSTM1-1 – GSH and *Cd*GSTM1-1 – Nb-GSH complex

   - The crystal structure of *Cd*GSTM1-1 in complex with GSH or Nb-GSH was resolved by X-ray crystallography at 2.55 Å and 2.05 Å resolution, respectively.
   - Tyr^{7} residue is conserved in all the compared mu-class GSTs, underlying its catalytic role for GSH activation.
   - Leu^{210} and Met^{212} are highly variable residues at the H-site and may contribute to wide substrate specificity and binding at the H-site.
   - The H-site is mostly formed by non-conserved residues which imply that the diversification in the evolution of these genes is primarily in the substrate-binding regions. This allows the adaptation of *Camelus dromedarius* to harsh climatic conditions.
   - Movement in β2-α2 loop (approx. 4Å) upon Nb-GSH binding suggests an induced-fit mechanism to facilitate the binding of various substrates.

4. hGSTM1-1

   - The crystal structure of ligand-free hGSTM1-1 was determined at 1.59 Å resolution. The high-resolution structure provided better information regarding the atomic positions, detailed structural features, and induced fit mechanism upon ligand binding.
   - The comparative studies of the present high-resolution structure with the previously solved low-resolution structure 1gtu allowed the better mapping of the interactions at the dimer interface.

- Differences in the β2-α2 and α3-α4 loop regions between the high and low-resolution structures were identified. Ser^{72}, Met^{104}, and His^{107} residues displayed subtle conformation differences. The most noticeable difference is identified in Arg^{42}.

The results reported here could pave the way to a better understanding of the cellular detoxification process to facilitate the design of better GST enzymes and more efficient inhibitors for use in biomedical and biotechnological applications. Future work on GSTs will also be expected to lead to better diagnostics tools (e.g., biosensors) for the detection of toxic compounds present in environmental samples and to monitoring the concentration of anticancer drugs in cancer chemotherapy for dosage optimization. GSTs are a highly promising candidate for protein engineering that can potentially be altered for novel functions, such as bioremediation. Hence, the study and understanding of GST enzymatic functions and their structures could guide future protein engineering experiments so that GSTs can be employed efficiently in biotechnological and biomedical applications.

Acknowledgements

The work was performed at the Protein Structure and Function Laboratory, Turku Bioscience Centre, University of Turku. Financial support from the Magnus Ehrnrooth Foundation and the University of Turku Foundation are gratefully acknowledged.

I would like to thank my supervisor Adj. Prof. Anastassios Papageorgiou for his extraordinary guidance and support during my research. I am truly grateful for his positive attitude, patience, and encouragement throughout these years. The suggestions, considerations, and critical remarks he provided were essential for my scientific growth. I also would like to thank my research director Prof. Jyrki Heino for his inspiration and direction. Docents Evangelia Chrysina and Rajaram Venkatesan are acknowledged for critically examining my dissertation and providing constructive comments and recommendations.

I also would like to thank all our collaborators from Agricultural University of Athens for all their work and vision. I am enormously grateful for all my previous and present workmates. I truly admire people from different culture and background that made my journey memorable. I want to thank Polychronis, Herbert and Imran for their witty scientific discussions during work hours. All my previous mentors in structural bioinformatics are acknowledged, Prof. Tiina Salminen, Dr. Martti Tolvanen and Prof. Mikko Metsä-Ketelä. In addition, my present and past faculty co-ordinators Nina Lehtimäki and Sari Järvi are thanked for their continuous support throughout these years.

My sincere gratitude goes to Turku Bioscience Centre for exceptional infrastructure support for research. Sirkku Grönroos and Paula Luoma are thanked for their administrative support. Technical support from Mikael Wasberg, Pasi Viljakainen, and Juha Strandén is acknowledged.

Finally, I would like to extend my adoration to my parents, siblings, and my wife Priyanka. My sons Nirvik and Prithak are thanked for being in my heart and for their gracious smiles.

Turku, August 2022
Nirmal Poudel

List of References

1. Booth J, Boyland E, Sims P. An enzyme from rat liver catalysing conjugations with glutathione. Biochem J. 1961;79:516–24.
2. Combes B, Stakelum GS. A liver enzyme that conjugates sulfobromophthalein sodium with glutathione. J Clin Invest. 1961;40:981–8.
3. Armstrong RN. Structure, Catalytic Mechanism, and Evolution of the Glutathione Transferases. Chem Res Toxicol. 1997;10:2–18.
4. Reinemer P, Dirr HW, Ladenstein R, Schäffer J, Gallay O, Huber R. The three-dimensional structure of class pi glutathione S-transferase in complex with glutathione sulfonate at 2.3 A resolution. EMBO J. 1991;10:1997–2005.
5. Hayes JD, McLellan LI. Glutathione and glutathione-dependent enzymes represent a co-ordinately regulated defence against oxidative stress. Free Radic Res. 1999;31:273–300.
6. Wilce MCJ, Parker MW. Structure and function of glutathione S-transferases. Biochim Biophys Acta BBA - Protein Struct Mol Enzymol. 1994;1205:1–18.
7. Dixon DP, Lapthorn A, Edwards R. Plant glutathione transferases. Genome Biol. 2002;3:REVIEWS3004.
8. Hayes JD, Flanagan JU, Jowsey IR. Glutathione transferases. Annu Rev Pharmacol Toxicol. 2005;45:51–88.
9. Tew KD, Townsend DM. Glutathione-S-Transferases As Determinants of Cell Survival and Death. Antioxid Redox Signal. 2012;17:1728–37.
10. Atkinson HJ, Babbitt PC. Glutathione Transferases Are Structural and Functional Outliers in the Thioredoxin Fold. Biochemistry. 2009;48:11108–16.
11. Di Pietro G, Magno LAV, Rios-Santos F. Glutathione S-transferases: an overview in cancer research. Expert Opin Drug Metab Toxicol. 2010;6:153–70.
12. Oakley A. Glutathione transferases: a structural perspective. Drug Metab Rev. 2011;43:138–51.
13. Mannervik B, Board PG, Hayes JD, Listowsky I, Pearson WR. Nomenclature for mammalian soluble glutathione transferases. Methods Enzymol. 2005;401:1–8.
14. Holm PJ, Bhakat P, Jegerschöld C, Gyobu N, Mitsuoka K, Fujiyoshi Y, et al. Structural basis for detoxification and oxidative stress protection in membranes. J Mol Biol. 2006;360:934–45.
15. Morgenstern R, Zhang J, Johansson K. Microsomal glutathione transferase 1: mechanism and functional roles. Drug Metab Rev. 2011;43:300–6.
16. Morel F, Aninat C. The glutathione transferase kappa family. Drug Metab Rev. 2011;43:281–91.
17. Harris JM, Meyer DJ, Coles B, Ketterer B. A novel glutathione transferase (13-13) isolated from the matrix of rat liver mitochondria having structural similarity to class theta enzymes. Biochem J. 1991;278 Pt 1:137–41.
18. Chronopoulou EG, Labrou NE. Glutathione transferases: emerging multidisciplinary tools in red and green biotechnology. Recent Pat Biotechnol. 2009;3:211–23.
19. Edwards R, Dixon DP, Walbot V. Plant glutathione S-transferases: enzymes with multiple functions in sickness and in health. Trends Plant Sci. 2000;5:193–8.

20. Thom R, Cummins I, Dixon DP, Edwards R, Cole DJ, Lapthorn AJ. Structure of a tau class glutathione S-transferase from wheat active in herbicide detoxification. Biochemistry. 2002;41:7008–20.
21. Sheehan D, Meade G, Foley VM, Dowd CA. Structure, function and evolution of glutathione transferases: implications for classification of non-mammalian members of an ancient enzyme superfamily. Biochem J. 2001;360 Pt 1:1–16.
22. Edwards R, Dixon DP. Plant Glutathione Transferases. In: Sies H, Packer L, editors. Methods in Enzymology. Academic Press; 2005. p. 169–86.
23. Dixon DP, Edwards R. Glutathione Transferases. Arab Book Am Soc Plant Biol. 2010;8:e0131.
24. Taylor VL, Cummins I, Brazier-Hicks M, Edwards R. Protective responses induced by herbicide safeners in wheat. Environ Exp Bot. 2013;88:93–9.
25. Pascal S, Scalla R. Purification and characterization of a safener-induced glutathione S-transferase from wheat (Triticum aestivum). Physiol Plant. 1999;106:17–27.
26. Cummins I, Cole DJ, Edwards R. Purification of Multiple Glutathione Transferases Involved in Herbicide Detoxification from Wheat (Triticum aestivumL.) Treated with the Safener Fenchlorazole-ethyl. Pestic Biochem Physiol. 1997;59:35–49.
27. Axarli I, Dhavala P, Papageorgiou AC, Labrou NE. Crystallographic and functional characterization of the fluorodifen-inducible glutathione transferase from Glycine max reveals an active site topography suited for diphenylether herbicides and a novel L-site. J Mol Biol. 2009;385:984–1002.
28. Cummins I, Dixon DP, Freitag-Pohl S, Skipsey M, Edwards R. Multiple roles for plant glutathione transferases in xenobiotic detoxification. Drug Metab Rev. 2011;43:266–80.
29. Dixon DP, McEwen AG, Lapthorn AJ, Edwards R. Forced evolution of a herbicide detoxifying glutathione transferase. J Biol Chem. 2003;278:23930–5.
30. Karavangeli M, Labrou NE, Clonis YD, Tsaftaris A. Development of transgenic tobacco plants overexpressing maize glutathione S-transferase I for chloroacetanilide herbicides phytoremediation. Biomol Eng. 2005;22:121–8.
31. Habig WH, Pabst MJ, Jakoby WB. Glutathione S-transferases. The first enzymatic step in mercapturic acid formation. J Biol Chem. 1974;249:7130–9.
32. Albarakati N, Khayyat D, Dallol A, Al-Maghrabi J, Nedjadi T. The prognostic impact of GSTM1/GSTP1 genetic variants in bladder Cancer. BMC Cancer. 2019;19:991.
33. Perperopoulou F, Ataya FS, Fouad D, Malik A, Saeed HM, Labrou NE. Biochemical Characterization of the Detoxifying Enzyme Glutathione Transferase P1-1 from the Camel Camelus Dromedarius. Cell Biochem Biophys. 2016;74:459–72.
34. Habig WH, Pabst MJ, Fleischner G, Gatmaitan Z, Arias IM, Jakoby WB. The identity of glutathione S-transferase B with ligandin, a major binding protein of liver. Proc Natl Acad Sci U S A. 1974;71:3879–82.
35. Coles BF, Kadlubar FF. Human Alpha Class Glutathione S-Transferases: Genetic Polymorphism, Expression, and Susceptibility to Disease. In: Sies H, Packer L, editors. Methods in Enzymology. Academic Press; 2005. p. 9–42.
36. Board PG, Menon D. Glutathione transferases, regulators of cellular metabolism and physiology. Biochim Biophys Acta. 2013;1830:3267–88.
37. Ross VL, Board PG, Webb GC. Chromosomal Mapping of the Human Mu Class Glutathione S-Transferases to 1p13. Genomics. 1993;18:87–91.
38. Board PG. Biochemical genetics of glutathione-S-transferase in man. Am J Hum Genet. 1981;33:36–43.
39. Campbell E, Takahashi Y, Abramovitz M, Peretz M, Listowsky I. A distinct human testis and brain mu-class glutathione S-transferase. Molecular cloning and characterization of a form present even in individuals lacking hepatic type mu isoenzymes. J Biol Chem. 1990;265:9188–93.

40. Hall AG, Autzen P, Cattan AR, Malcolm AJ, Cole M, Kernahan J, et al. Expression of μ Class Glutathione S-Transferase Correlates with Event-free Survival in Childhood Acute Lymphoblastic Leukemia. Cancer Res. 1994;54:5251–4.
41. Takanashi M, Morimoto A, Yagi T, Kuriyama K, Kano G, Imamura T, et al. Impact of glutathione S-transferase gene deletion on early relapse in childhood B-precursor acute lymphoblastic leukemia. Haematologica. 2003;88:1238–44.
42. Board PG, Webb GC, Coggan M. Isolation of a cDNA clone and localization of the human glutathione S-transferase 3 genes to chromosome bands 11q13 and 12q13-14. Ann Hum Genet. 1989;53:205–13.
43. Townsend DM, Tew KD. The role of glutathione-S-transferase in anti-cancer drug resistance. Oncogene. 2003;22:7369–75.
44. Allan JM, Wild CP, Rollinson S, Willett EV, Moorman AV, Dovey GJ, et al. Polymorphism in glutathione S-transferase P1 is associated with susceptibility to chemotherapy-induced leukemia. Proc Natl Acad Sci. 2001;98:11592–7.
45. Landi S. Mammalian class theta GST and differential susceptibility to carcinogens: a review. Mutat Res Mutat Res. 2000;463:247–83.
46. Aguilera I, Alvarez-Marquez A, Gentil MA, Fernandez-Alonso J, Fijo J, Saez C, et al. Anti-glutathione S-transferase T1 antibody-mediated rejection in C4d-positive renal allograft recipients. Nephrol Dial Transplant. 2008;23:2393–8.
47. Fernández-Cañón JM, Hejna J, Reifsteck C, Olson S, Grompe M. Gene Structure, Chromosomal Location, and Expression Pattern of Maleylacetoacetate Isomerase. Genomics. 1999;58:263–9.
48. Board PG. The omega-class glutathione transferases: structure, function, and genetics. Drug Metab Rev. 2011;43:226–35.
49. Kölsch H, Linnebank M, Lütjohann D, Jessen F, Wüllner U, Harbrecht U, et al. Polymorphisms in glutathione S-transferase omega-1 and AD, vascular dementia, and stroke. Neurology. 2004;63:2255–60.
50. Sinning I, Kleywegt GJ, Cowan SW, Reinemer P, Dirr HW, Huber R, et al. Structure Determination and Refinement of Human Alpha Class Glutathione Transferase A1-1, and a Comparison with the Mu and Pi Class Enzymes. J Mol Biol. 1993;232:192–212.
51. Tars K, Olin B, Mannervik B. Structural Basis for Featuring of Steroid Isomerase Activity in Alpha Class Glutathione Transferases. J Mol Biol. 2010;397:332–40.
52. Bruns CM, Hubatsch I, Ridderström M, Mannervik B, Tainer JA. Human glutathione transferase A4-4 crystal structures and mutagenesis reveal the basis of high catalytic efficiency with toxic lipid peroxidation products 11Edited by D. Rees. J Mol Biol. 1999;288:427–39.
53. Patskovsky Y, Patskovska L, Almo SC, Listowsky I. Transition State Model and Mechanism of Nucleophilic Aromatic Substitution Reactions Catalyzed by Human Glutathione S-Transferase M1a-1a. Biochemistry. 2006;45:3852–62.
54. Raghunathan S, Chandross RJ, Kretsinger RH, Allison TJ, Penington CJ, Rule GS. Crystal Structure of Human Class mu Glutathione Transferase GSTM2-2: Effects of Lattice Packing on Conformational Heterogeneity. J Mol Biol. 1994;238:815–32.
55. Patskovska LN, Fedorov AA, Patskovsky YV, Almo SC, Listowsky I. Expression, crystallization and preliminary X-ray analysis of ligand-free human glutathione S-transferase M2–2. Acta Crystallogr D Biol Crystallogr. 1998;54:458–60.
56. Oakley AJ, Lo Bello M, Nuccetelli M, Mazzetti AP, Parker MW. The ligandin (non-substrate) binding site of human Pi class glutathione transferase is located in the electrophile binding site (H-site). J Mol Biol. 1999;291:913–26.
57. Tars K, Larsson A-K, Shokeer A, Olin B, Mannervik B, Kleywegt GJ. Structural Basis of the Suppressed Catalytic Activity of Wild-type Human Glutathione Transferase T1-1 Compared to its W234R Mutant. J Mol Biol. 2006;355:96–105.

58. Rossjohn J, McKinstry WJ, Oakley AJ, Verger D, Flanagan J, Chelvanayagam G, et al. Human theta class glutathione transferase: the crystal structure reveals a sulfate-binding pocket within a buried active site. Structure. 1998;6:309–22.
59. Aritake K, Kado Y, Inoue T, Miyano M, Urade Y. Structural and Functional Characterization of HQL-79, an Orally Selective Inhibitor of Human Hematopoietic Prostaglandin D Synthase*. J Biol Chem. 2006;281:15277–86.
60. Polekhina G, Board PG, Blackburn AC, Parker MW. Crystal Structure of Maleylacetoacetate Isomerase/Glutathione Transferase Zeta Reveals the Molecular Basis for Its Remarkable Catalytic Promiscuity,. Biochemistry. 2001;40:1567–76.
61. Board PG, Coggan M, Chelvanayagam G, Easteal S, Jermiin LS, Schulte GK, et al. Identification, Characterization, and Crystal Structure of the Omega Class Glutathione Transferases*. J Biol Chem. 2000;275:24798–806.
62. Zhou H, Brock J, Casarotto MG, Oakley AJ, Board PG. Novel Folding and Stability Defects Cause a Deficiency of Human Glutathione Transferase Omega 1*. J Biol Chem. 2011;286:4271–9.
63. Dixon DP, Cummins I, Cole DJ, Edwards R. Glutathione-mediated detoxification systems in plants. Curr Opin Plant Biol. 1998;1:258–66.
64. Marrs KA. THE FUNCTIONS AND REGULATION OF GLUTATHIONE S-TRANSFERASES IN PLANTS. Annu Rev Plant Physiol Plant Mol Biol. 1996;47:127–58.
65. Schmitt R, Sandermann H. Specific Localization of β-D-Glucoside Conjugates of 2,4-Dichlorophenoxyacetic Acid in Soybean Vacuoles. Z Für Naturforschung C. 1982;37:772–7.
66. Sandermann H. Plant metabolism of xenobiotics. Trends Biochem Sci. 1992;17:82–4.
67. Yun M-S, Yogo Y, Miura R, Yamasue Y, Fischer AJ. Cytochrome P-450 monooxygenase activity in herbicide-resistant and -susceptible late watergrass (Echinochloa phyllopogon). Pestic Biochem Physiol. 2005;83:107–14.
68. Bolwell GP, Bozak K, Zimmerlin A. Plant cytochrome P450. Phytochemistry. 1994;37:1491–506.
69. Schuler MA, Werck-Reichhart D. Functional genomics of P450s. Annu Rev Plant Biol. 2003;54:629–67.
70. Inui H, Ohkawa H. Herbicide resistance in transgenic plants with mammalian P450 monooxygenase genes. Pest Manag Sci. 2005;61:286–91.
71. Morant M, Bak S, Møller BL, Werck-Reichhart D. Plant cytochromes P450: tools for pharmacology, plant protection and phytoremediation. Curr Opin Biotechnol. 2003;14:151–62.
72. Letouzé A, GASQUEZ J. Enhanced activity of several herbicide-degrading enzymes: a suggested mechanism responsible for multiple resistance in blackgrass (Alopecurus myosuroides Huds.). Agronomie. 2003;23:601–8.
73. Coleman J, Blake-Kalff M, Davies E. Detoxification of xenobiotics by plants: chemical modification and vacuolar compartmentation. Trends Plant Sci. 1997;2:144–51.
74. Dixon DP, Skipsey M, Edwards R. Roles for glutathione transferases in plant secondary metabolism. Phytochemistry. 2010;71:338–50.
75. Klein M, Burla B, Martinoia E. The multidrug resistance-associated protein (MRP/ABCC) subfamily of ATP-binding cassette transporters in plants. FEBS Lett. 2006;580:1112–22.
76. Martinoia E, Grill E, Tommasini R, Kreuz K, Amrhein N. ATP-dependent glutathione S-conjugate “export” pump in the vacuolar membrane of plants. Nature. 1993;364:247–9.
77. Windsor B, Roux SJ, Lloyd A. Multiherbicide tolerance conferred by AtPgp1 and apyrase overexpression in Arabidopsis thaliana. Nat Biotechnol. 2003;21:428–33.
78. Mentewab A, Stewart CN. Overexpression of an Arabidopsis thaliana ABC transporter confers kanamycin resistance to transgenic plants. Nat Biotechnol. 2005;23:1177–80.
79. DeRidder BP, Goldsbrough PB. Organ-specific expression of glutathione S-transferases and the efficacy of herbicide safeners in Arabidopsis. Plant Physiol. 2006;140:167–75.

80. Powles SB, Yu Q. Evolution in action: plants resistant to herbicides. Annu Rev Plant Biol. 2010;61:317–47.
81. Ryan GF. Resistance of Common Groundsel to Simazine and Atrazine. Weed Sci. 1970;18:614–6.
82. Gressel J. Evolving understanding of the evolution of herbicide resistance. Pest Manag Sci. 2009;65:1164–73.
83. Michel H, Deisenhofer J. Relevance of the photosynthetic reaction center from purple bacteria to the structure of photosystem II. Biochemistry. 1988;27:1–7.
84. Heap, I. The International Herbicide-Resistant Weed Database. Online. www.weedscience.org. 2021. http://www.weedscience.org/Home.aspx. Accessed 31 Aug 2021.
85. Gardner SN, Gressel J, Mangel M. A revolving dose strategy to delay the evolution of both quantitative vs major monogene resistances to pesticides and drugs. Int J Pest Manag. 1998;44:161–80.
86. Gressel J. Molecular biology of weed control. Transgenic Res. 2000;9:355–82.
87. Arntzen CJ, Pfister K, Steinback KE. The mechanism of chloroplast triazine resistance: alterations in the site of herbicide action [Weed biotypes, photosynthesis]. 1982.
88. Gronwald JW. Resistance to PS II Inhibitor Herbicides. In: De Prado R, Jorrín J, García-Torres L, editors. Weed and Crop Resistance to Herbicides. Dordrecht: Springer Netherlands; 1997. p. 53–9.
89. Moss SR, Cussans GW. Variability in the susceptibility of Alopecurus myosuroides (black-grass) to chlortoluron and isoproturon. Asp Appl Biol. 1985.
90. Heap I, Knight R. The occurrence of herbicide cross-resistance in a population of annual ryegrass, Lolium rigidum, resistant to diclofop-methyl. Aust J Agric Res. 1986;37:149–56.
91. Hall LM, Moss SR, Powles SB. Mechanisms of Resistance to Aryloxyphenoxypropionate Herbicides in Two Resistant Biotypes ofAlopecurus myosuroides(blackgrass): Herbicide Metabolism as a Cross-Resistance Mechanism. Pestic Biochem Physiol. 1997;57:87–98.
92. Cummins I, Wortley DJ, Sabbadin F, He Z, Coxon CR, Straker HE, et al. Key role for a glutathione transferase in multiple-herbicide resistance in grass weeds. Proc Natl Acad Sci U S A. 2013;110:5812–7.
93. Yuan JS, Tranel PJ, Stewart CN. Non-target-site herbicide resistance: a family business. Trends Plant Sci. 2007;12:6–13.
94. Rieger MA, Lamond M, Preston C, Powles SB, Roush RT. Pollen-Mediated Movement of Herbicide Resistance Between Commercial Canola Fields. Science. 2002;296:2386–8.
95. Owen MDK, Zelaya IA. Herbicide-resistant crops and weed resistance to herbicides. Pest Manag Sci. 2005;61:301–11.
96. Burnet MWM, Loveys BR, Holtum JAM, Powles SB. Increased Detoxification Is a Mechanism of Simazine Resistance in Lolium rigidum. Pestic Biochem Physiol. 1993;46:207–18.
97. Hall LM, Moss SR, Powles SB. Mechanism of Resistance to Chlorotoluron in Two Biotypes of the Grass Weed Alopecurus myosuroides. Pestic Biochem Physiol. 1995;53:180–92.
98. Christopher JT, Powles SB, Liljegren DR, Holtum JAM. Cross-Resistance to Herbicides in Annual Ryegrass (Lolium rigidum) 1: II. Chlorsulfuron Resistance Involves a Wheat-Like Detoxification System. Plant Physiol. 1991;95:1036–43.
99. Christopher JT, Preston C, Powles SB. Malathion Antagonizes Metabolism-Based Chlorsulfuron Resistance in Lolium rigidum. Pestic Biochem Physiol. 1994;49:172–82.
100. Cummins I, Cole DJ, Edwards R. A role for glutathione transferases functioning as glutathione peroxidases in resistance to multiple herbicides in black-grass. Plant J Cell Mol Biol. 1999;18:285–92.
101. Cocker KM, Northcroft DS, Coleman JOD, Moss SR. Resistance to ACCase-inhibiting herbicides and isoproturon in UK populations of Lolium multiflorum: mechanisms of resistance and implications for control. Pest Manag Sci. 2001;57:587–97.

102. Anderson MP, Gronwald JW. Atrazine Resistance in a Velvetleaf (Abutilon theophrasti) Biotype Due to Enhanced Glutathione S-Transferase Activity 1. Plant Physiol. 1991;96:104–9.
103. Gray JA, Balke NE, Stoltenberg DE. Increased Glutathione Conjugation of Atrazine Confers Resistance in a Wisconsin Velvetleaf (Abutilon theophrasti) Biotype. Pestic Biochem Physiol. 1996;55:157–71.
104. Plaisance KL, Gronwald JW. Enhanced Catalytic Constant for GlutathioneS-Transferase (Atrazine) Activity in an Atrazine-ResistantAbutilon theophrastiBiotype. Pestic Biochem Physiol. 1999;63:34–49.
105. Gardin JAC, Gouzy J, Carrère S, Délye C. ALOMYbase, a resource to investigate non-target-site-based resistance to herbicides inhibiting acetolactate-synthase (ALS) in the major grass weed Alopecurus myosuroides (black-grass). BMC Genomics. 2015;16:590.
106. Gaines TA, Lorentz L, Figge A, Herrmann J, Maiwald F, Ott M-C, et al. RNA-Seq transcriptome analysis to identify genes involved in metabolism-based diclofop resistance in Lolium rigidum. Plant J Cell Mol Biol. 2014;78:865–76.
107. Cummins I, Bryant DN, Edwards R. Safener responsiveness and multiple herbicide resistance in the weed black-grass (Alopecurus myosuroides). Plant Biotechnol J. 2009;7:807–20.
108. Chronopoulou EG, Ataya F, Labrou NE. A Microplate-based Platform with Immobilized Human Glutathione Transferase A1-1 for High-throughput Screening of Plant-origin Inhibitors. Curr Pharm Biotechnol. 2018;19:925–31.
109. Chronopoulou EG, Papageorgiou AC, Ataya F, Nianiou-Obeidat I, Madesis P, Labrou NE. Expanding the Plant GSTome Through Directed Evolution: DNA Shuffling for the Generation of New Synthetic Enzymes With Engineered Catalytic and Binding Properties. Front Plant Sci. 2018;9:1737.
110. Dalmizrak O, Teralı K, Asuquo EB, Ogus IH, Ozer N. The Relevance of Glutathione Reductase Inhibition by Fluoxetine to Human Health and Disease: Insights Derived from a Combined Kinetic and Docking Study. Protein J. 2019;38:515–24.
111. Green JM. Current state of herbicides in herbicide-resistant crops. Pest Manag Sci. 2014;70:1351–7.
112. Heap I. Global perspective of herbicide-resistant weeds. Pest Manag Sci. 2014;70:1306–15.
113. Kataria R, Khatkar A. Molecular Docking of Natural Phenolic Compounds for the Screening of Urease Inhibitors. Curr Pharm Biotechnol. 2019;20:410–21.
114. Walsh MJ, Powles SB. Management of herbicide resistance in wheat cropping systems: learning from the Australian experience. Pest Manag Sci. 2014;70:1324–8.
115. Shehu D, Alias Z. Functional Role of Tyr12 in the Catalytic Activity of Novel Zeta-like Glutathione S-transferase from Acidovorax sp. KKS102. Protein J. 2018;37:261–9.
116. Karatolos N, Williamson MS, Denholm I, Gorman K, Ffrench-Constant RH, Bass C. Over-expression of a cytochrome P450 is associated with resistance to pyriproxyfen in the greenhouse whitefly Trialeurodes vaporariorum. PloS One. 2012;7:e31077.
117. Frear DS, Swanson HR. Biosynthesis of S-(4-ethylamino-6-isopropylamino- 2-s-triazino) glutathione: Partial purification and properties of a glutathione S-transferase from corn. Phytochemistry. 1970;9:2123–32.
118. Mozer TJ, Tiemeier DC, Jaworski EG. Purification and characterization of corn glutathione S-transferase. Biochemistry. 1983;22:1068–72.
119. Edwards R, Cole DJ. Glutathione Transferases in Wheat (Triticum) Species with Activity toward Fenoxaprop-Ethyl and Other Herbicides. Pestic Biochem Physiol. 1996;54:96–104.
120. Neuefeind T, Huber R, Dasenbrock H, Prade L, Bieseler B. Crystal structrure of herbicide-detoxifying maize glutathione S-transferase-I in complex with lactoylglutathione: evidence for an induced-fit mechanism 11Edited by D. Rees. J Mol Biol. 1997;274:446–53.
121. Neuefeind T, Huber R, Reinemer P, Knäblein J, Prade L, Mann K, et al. Cloning, sequencing, crystallization and X-ray structure of glutathione S-transferase-III from Zea mays var. mutin: a

leading enzyme in detoxification of maize herbicides11Edited by D. Rees. J Mol Biol. 1997;274:577–87.
122. Reinemer P, Prade L, Hof P, Neuefeind T, Huber R, Zettl R, et al. Three-dimensional structure of glutathione S-transferase from Arabidopsis thaliana at 2.2 Å resolution: Structural characterization of herbicide-conjugating plant glutathione S-transferases and a novel active site architecture. J Mol Biol. 1996;255:289–309.
123. Axarli I, Dhavala P, Papageorgiou AC, Labrou NE. Crystal structure of Glycine max glutathione transferase in complex with glutathione: investigation of the mechanism operating by the Tau class glutathione transferases. Biochem J. 2009;422:247–56.
124. Armstrong RN. Glutathione S-transferases: structure and mechanism of an archetypical detoxication enzyme. Adv Enzymol Relat Areas Mol Biol. 1994;69:1–44.
125. Reinemer P, Dirr HW, Ladenstein R, Huber R, Lo Bello M, Federici G, et al. Three-dimensional structure of class pi glutathione S-transferase from human placenta in complex with S-hexylglutathione at 2.8 A resolution. J Mol Biol. 1992;227:214–26.
126. Prade L, Huber R, Bieseler B. Structures of herbicides in complex with their detoxifying enzyme glutathione S-transferase - explanations for the selectivity of the enzyme in plants. Struct Lond Engl 1993. 1998;6:1445–52.
127. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, et al. UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem. 2004;25:1605–12.
128. Lallement P-A, Brouwer B, Keech O, Hecker A, Rouhier N. The still mysterious roles of cysteine-containing glutathione transferases in plants. Front Pharmacol. 2014;0.
129. Dixon DP, Cole DJ, Edwards R. Dimerisation of maize glutathione transferases in recombinant bacteria. Plant Mol Biol. 1999;40:997–1008.
130. Mannervik B. Five decades with glutathione and the GSTome. J Biol Chem. 2012;287:6072–83.
131. Armstrong RN. Glutathione S-transferases: reaction mechanism, structure, and function. Chem Res Toxicol. 1991;4:131–40.
132. Labrou NE, Mello LV, Clonis YD. Functional and structural roles of the glutathione-binding residues in maize (Zea mays) glutathione S-transferase I. Biochem J. 2001;358 Pt 1:101–10.
133. Liu S, Zhang P, Ji X, Johnson WW, Gilliland GL, Armstrong RN. Contribution of tyrosine 6 to the catalytic mechanism of isoenzyme 3-3 of glutathione S-transferase. J Biol Chem. 1992;267:4296–9.
134. Mashiyama ST, Malabanan MM, Akiva E, Bhosle R, Branch MC, Hillerich B, et al. Large-Scale Determination of Sequence, Structure, and Function Relationships in Cytosolic Glutathione Transferases across the Biosphere. PLOS Biol. 2014;12:e1001843.
135. Dixon DP, Sellars JD, Edwards R. The Arabidopsis phi class glutathione transferase AtGSTF2: binding and regulation by biologically active heterocyclic ligands. Biochem J. 2011;438:63–70.
136. Dixon DP, Edwards R. Selective Binding of Glutathione Conjugates of Fatty Acid Derivatives by Plant Glutathione Transferases. J Biol Chem. 2009;284:21249–56.
137. Dixon DP, Lapthorn A, Madesis P, Mudd EA, Day A, Edwards R. Binding and Glutathione Conjugation of Porphyrinogens by Plant Glutathione Transferases*. J Biol Chem. 2008;283:20268–76.
138. Axarli IA, Rigden DJ, Labrou NE. Characterization of the ligandin site of maize glutathione S-transferase I. Biochem J. 2004;382 Pt 3:885–93.
139. Sau A, Pellizzari Tregno F, Valentino F, Federici G, Caccuri AM. Glutathione transferases and development of new principles to overcome drug resistance. Arch Biochem Biophys. 2010;500:116–22.
140. Pljesa-Ercegovac M, Savic-Radojevic A, Matic M, Coric V, Djukic T, Radic T, et al. Glutathione Transferases: Potential Targets to Overcome Chemoresistance in Solid Tumors. Int J Mol Sci. 2018;19:E3785.

141. Adler V, Yin Z, Fuchs SY, Benezra M, Rosario L, Tew KD, et al. Regulation of JNK signaling by GSTp. EMBO J. 1999;18:1321–34.
142. Laborde E. Glutathione transferases as mediators of signaling pathways involved in cell proliferation and cell death. Cell Death Differ. 2010;17:1373–80.
143. Tew KD. Glutathione-Associated Enzymes In Anticancer Drug Resistance. Cancer Res. 2016;76:7–9.
144. Allocati N, Masulli M, Di Ilio C, Federici L. Glutathione transferases: substrates, inihibitors and pro-drugs in cancer and neurodegenerative diseases. Oncogenesis. 2018;7:1–15.
145. Ruzza P, Calderan A. Glutathione Transferase (GST)-Activated Prodrugs. Pharmaceutics. 2013;5:220–31.
146. Ruzza P, Rosato A, Rossi CR, Floreani M, Quintieri L. Glutathione Transferases as Targets for Cancer Therapy. Anti-Cancer Agents Med Chem- Anti-Cancer Agents. 2009;9:763–77.
147. Xu W, Liu LZ, Loizidou M, Ahmed M, Charles IG. The role of nitric oxide in cancer. Cell Res. 2002;12:311–20.
148. Kapoli P, Axarli IA, Platis D, Fragoulaki M, Paine M, Hemingway J, et al. Engineering sensitive glutathione transferase for the detection of xenobiotics. Biosens Bioelectron. 2008;24:498–503.
149. Andreou VG, Clonis YD. Novel fiber-optic biosensor based on immobilized glutathione S-transferase and sol–gel entrapped bromcresol green for the determination of atrazine. Anal Chim Acta. 2002;460:151–61.
150. Chronopoulou EG, Vlachakis D, Papageorgiou AC, Ataya FS, Labrou NE. Structure-based design and application of an engineered glutathione transferase for the development of an optical biosensor for pesticides determination. Biochim Biophys Acta BBA - Gen Subj. 2019;1863:565–76.
151. Chronopoulou EG, Papageorgiou AC, Markoglou A, Labrou NE. Inhibition of human glutathione transferases by pesticides: Development of a simple analytical assay for the quantification of pesticides in water. J Mol Catal B Enzym. 2012;81:43–51.
152. Bernal C, Rodríguez K, Martínez R. Integrating enzyme immobilization and protein engineering: An alternative path for the development of novel and improved industrial biocatalysts. Biotechnol Adv. 2018;36:1470–80.
153. Oliveira TIS, Oliveira M, Viswanathan S, Fátima Barroso M, Barreiros L, Nunes OC, et al. Molinate quantification in environmental water by a glutathione-S-transferase based biosensor. Talanta. 2013;106:249–54.
154. Morou E, Ismail HM, Dowd AJ, Hemingway J, Labrou N, Paine M, et al. A dehydrochlorinase-based pH change assay for determination of DDT in sprayed surfaces. Anal Biochem. 2008;378:60–4.
155. Heydens WF, Wilson AG, Kier LD, Lau H, Thake DC, Martens MA. An evaluation of the carcinogenic potential of the herbicide alachlor4 to man. Hum Exp Toxicol. 1999;18:363–91.
156. Perperopoulou F, Fragoulaki M, Papageorgiou AC, Labrou NE. Directed Evolution of a Glutathione Transferase for the Development of a Biosensor for Alachlor Determination. Symmetry. 2021;13:461.
157. Materon EM, Jimmy Huang P-J, Wong A, Pupim Ferreira AA, Sotomayor MDPT, Liu J. Glutathione-s-transferase modified electrodes for detecting anticancer drugs. Biosens Bioelectron. 2014;58:232–6.
158. Benekos K, Kissoudis C, Nianiou-Obeidat I, Labrou N, Madesis P, Kalamaki M, et al. Overexpression of a specific soybean GmGSTU4 isoenzyme improves diphenyl ether and chloroacetanilide herbicide tolerance of transgenic tobacco plants. J Biotechnol. 2010;150:195–201.
159. Hu T. A glutathione s-transferase confers herbicide tolerance in rice. Crop Breed Appl Biotechnol. 2014;14:76–81.

160. Sharma R, Sahoo A, Devendran R, Jain M. Over-Expression of a Rice Tau Class Glutathione S-Transferase Gene Improves Tolerance to Salinity and Oxidative Stresses in Arabidopsis. PLOS ONE. 2014;9:e92900.
161. Milligan AS, Daly A, Parry MAJ, Lazzeri PA, Jepson I. The expression of a maize glutathione S-transferase gene in transgenic wheat confers herbicide tolerance, both in planta and in vitro. Mol Breed. 2001;7:301–15.
162. Skipsey M, Cummins I, Andrews CJ, Jepson I, Edwards R. Manipulation of plant tolerance to herbicides through co-ordinated metabolic engineering of a detoxifying glutathione transferase and thiol cosubstrate. Plant Biotechnol J. 2005;3:409–20.
163. Takesawa T, Ito M, Kanzaki H, Kameya N, Nakamura I. Over-expression of ζ glutathione S-transferase in transgenic rice enhances germination and growth at low temperature. Mol Breed. 2002;9:93–101.
164. Moons A. Regulatory and functional interactions of plant growth regulators and plant glutathione S-transferases (GSTs). Vitam Horm. 2005;72:155–202.
165. Georgakis ND, Karagiannopoulos DA, Thireou TN, Eliopoulos EE, Labrou NE, Tsoungas PG, et al. Concluding the trilogy: The interaction of 2,2′-dihydroxy-benzophenones and their carbonyl N-analogues with human glutathione transferase M1-1 face to face with the P1-1 and A1-1 isoenzymes involved in MDR. Chem Biol Drug Des. 2017;90:900–8.
166. Pouliou F, Perperopoulou F, Labrou NE. Comparative Analysis of Two Stress-Inducible tau Class Glutathione Transferases from Glycine max Revealed Significant Catalytic and Structural Diversification. Protein Pept Lett. 2017;24:922–35.
167. Bradford MM. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem. 1976;72:248–54.
168. Newman J, Egan D, Walter TS, Meged R, Berry I, Ben Jelloul M, et al. Towards rationalization of crystallization screening for small- to medium-sized academic laboratories: the PACT/JCSG+ strategy. Acta Crystallogr D Biol Crystallogr. 2005;61 Pt 10:1426–31.
169. Grimm C, Chari A, Reuter K, Fischer U. A crystallization screen based on alternative polymeric precipitants. Acta Crystallogr D Biol Crystallogr. 2010;66 Pt 6:685–97.
170. Incardona M-F, Bourenkov GP, Levik K, Pieritz RA, Popov AN, Svensson O. EDNA: a framework for plugin-based applications applied to X-ray experiment online data analysis. J Synchrotron Radiat. 2009;16:872–9.
171. Evans PR, Murshudov GN. How good are my data and what is the resolution? Acta Crystallogr D Biol Crystallogr. 2013;69:1204–14.
172. Matthews BW. Solvent content of protein crystals. J Mol Biol. 1968;33:491–7.
173. Kabsch W. Integration, scaling, space-group assignment and post-refinement. Acta Crystallogr D Biol Crystallogr. 2010;66 Pt 2:133–44.
174. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, et al. Overview of the CCP4 suite and current developments. Acta Crystallogr D Biol Crystallogr. 2011;67:235–42.
175. Keegan RM, Winn MD. MrBUMP: an automated pipeline for molecular replacement. Acta Crystallogr D Biol Crystallogr. 2008;64 Pt 1:119–24.
176. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. Phaser crystallographic software. J Appl Crystallogr. 2007;40 Pt 4:658–74.
177. Cowtan K. The Buccaneer software for automated model building. 1. Tracing protein chains. Acta Crystallogr D Biol Crystallogr. 2006;62 Pt 9:1002–11.
178. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr. 2004;60 Pt 12 Pt 1:2126–32.
179. Murshudov GN, Skubak P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, et al. REFMAC5 for the refinement of macromolecular crystal structures. Acta Crystallogr D Biol Crystallogr. 2011;67 Pt 4:355–67.
180. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. Phaser crystallographic software. J Appl Crystallogr. 2007;40 Pt 4:658–74.

181. Adams PD, Afonine PV, Bunkóczi G, Chen VB, Davis IW, Echols N, et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr. 2010;66 Pt 2:213–21.
182. Bunkóczi G, Read RJ. Improvement of molecular-replacement models with Sculptor. Acta Crystallogr D Biol Crystallogr. 2011;67 Pt 4:303–12.
183. Chen VB, Arendall WB 3rd, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, et al. MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr D Biol Crystallogr. 2010;66 Pt 1:12–21.
184. Krissinel E, Henrick K. Inference of macromolecular assemblies from crystalline state. J Mol Biol. 2007;372:774–97.
185. Krissinel E, Henrick K. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. Acta Crystallogr D Biol Crystallogr. 2004;60:2256–68.
186. Laskowski RA. PDBsum new things. Nucleic Acids Res. 2009;37 Database issue:D355–9.
187. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 1994;22:4673–80.
188. Gouet P, Courcelle E, Stuart DI, Métoz F. ESPript: analysis of multiple sequence alignments in PostScript. Bioinforma Oxf Engl. 1999;15:305–8.
189. Delano W. The PyMOL Molecular Graphics System. 2002. http://www.pymol.org. Accessed 24 Apr 2014.
190. Skopelitou K, Muleta AW, Papageorgiou AC, Chronopoulou E, Labrou NE. Catalytic features and crystal structure of a tau class glutathione transferase from Glycine max specifically upregulated in response to soybean mosaic virus infections. Biochim Biophys Acta BBA - Proteins Proteomics. 2015;1854:166–77.
191. Axarli I, Muleta AW, Vlachakis D, Kossida S, Kotzia G, Maltezos A, et al. Directed evolution of Tau class glutathione transferases reveals a site that regulates catalytic efficiency and masks co-operativity. Biochem J. 2016;473:559–70.
192. Pégeot H, Koh CS, Petre B, Mathiot S, Duplessis S, Hecker A, et al. The poplar Phi class glutathione transferase: expression, activity and structure of GSTF1. Front Plant Sci. 2014;5:712.
193. Patskovsky YV, Patskovska LN, Listowsky I. Functions of His107 in the catalytic mechanism of human glutathione S-transferase hGSTM1a-1a. Biochemistry. 1999;38:1193–202.
194. Ji X, Armstrong RN, Gilliland GL. Snapshots along the reaction coordinate of an SNAr reaction catalyzed by glutathione transferase. Biochemistry. 1993;32:12949–54.
195. Sun YJ, Kuan IC, Tam MF, Hsiao CD. The three-dimensional structure of an avian class-mu glutathione S-transferase, cGSTM1-1 at 1.94 A resolution. J Mol Biol. 1998;278:239–52.
196. Juárez-Martínez AB, Sotelo-Mundo RR, Rudiño-Piñera E. Crystal structure of a class-mu glutathione S-transferase from whiteleg shrimp Litopenaeus vannamei: structural changes in the xenobiotic binding H-site may alter the spectra of molecules bound. J Biochem Mol Toxicol. 2017;31.

TURUN
YLIOPISTO
UNIVERSITY
OF TURKU

Painosalama Oy, Turku, Finland 2022

ISBN 978-951-29-8963-8 (Print)
ISBN 978-951-29-8964-5 (PDF/Online)
ISSN 0082-7002 (Print)
ISSN 2343-3175 (Online)