Turun yliopisto
University of Turku

GLUTATHIONE TRANSFERASES AS DETOXIFICATION AGENTS:

Structural and Functional Studies

Abdi Worku Muleta

TURUN YLIOPISTON JULKAISUJA – ANNALES UNIVERSITATIS TURKUENSIS
Sarja - ser. D osa - tom. 1240 | Medica - Odontologica | Turku 2016

University of Turku

Faculty of Mathematics and Natural Sciences
Department of Biochemistry and Food Chemistry
Doctoral Degree Program in Molecular Life Sciences

Supervised by

Adj. Professor Anastassios C. Papageorgiou, PhD
Turku Centre for Biotechnology
University of Turku and Åbo Akademi University
University of Turku, Turku, Finland

Reviewed by

David Reverter, PhD
Institut de Biotecnologia i de Biomedicina
Universitat Autonoma de Barcelona, Barcelona, Spain

Adj. Professor Tiina Salminen, PhD
Structural Bioinformatics Laboratory, Biochemistry
Faculty of Science and Engineering
Åbo Akademi University, Turku, Finland

Opponent

Professor Savvas Savvides, PhD
Laboratory for Protein Biochemistry and Biomolecular Engineering
Department of Biochemistry and Microbiology
VIB Inflammation Research Center (IRC)
Ghent University, Ghent, Belgium

Cover photo: Abdi Worku Muleta

The originality of this thesis has been checked in accordance with the University of Turku quality assurance system using the Turnitin OriginalityCheck service.

ISBN 978-951-29-6538-0 (PRINT)
ISBN 978-951-29-6538-7 (PDF)
ISSN 0355-9483 (Print)
ISSN 2343-3213 (Online)
Painosalama Oy - Turku, Finland 2016

To my beloved family

LIST OF ORIGINAL PUBLICATIONS

This thesis is based on the following articles that are referred to by their Roman numerals in the text.

I. Skopelitou K, **Muleta AW**, Pavli O, Skaracis GN, Flemetakis E, Papageorgiou AC, Labrou NE. (2012) Overlapping protective roles for glutathione transferase gene family members in chemical and oxidative stress response in *Agrobacterium tumefaciens*. *Funct Integr Genomics.* 12(1):157-72

II. Skopelitou K*, **Muleta AW***, Papageorgiou AC, Chronopoulou E, Labrou NE. (2015) 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.*1854(2):166-77

III. Axarli I*, **Muleta AW***, Vlachakis D, Kossida S, Kotzia G, Maltezos A, Dhavala P, Papageorgiou AC, Labrou NE. (2016) Directed evolution of tau class glutathione transferases reveals a site that regulates catalytic efficiency and masks co-operativity. *Biochem J*. 473(5): 559-570

IV. Skopelitou K, **Muleta AW**, Papageorgiou AC, Chronopoulou EG, Pavli O, Flemetakis E, Skaracis GN, and Labrou NE. (2016) Characterization and functional analysis of a recombinant tau class glutathione transferase *Gm*GSTU2-2 from *Glycine max. Int J Biol Macrom*. (Epub ahead of print)

* Joint first authors.

ORIGINAL PUBLICATIONS NOT INCLUDED IN THE THESIS

I. Axarli I, **Muleta AW**, Chronopoulou E, Papageorgiou AC, Labrou NE. (2016) Directed evolution of glutathione transferases towards a selective glutathione-binding site and improved oxidative stability. *Biochim Biophys Acta.* (Submitted)

ABBREVIATIONS

2,4-D	2,4-dichlorophenoxyacetic acid
4NM	(4-nitrophenyl)methanethiol
ABC	ATP binding cassette
*Aa*GSTE2-2	Epsilon class GST2 from *Aedes aegypti*
*Atu*GSTs	*Agrobacterium tumefaciens* C58 GSTs
CDNB	1-chloro-2,4-dinitrobenzene
CHARMM	Chemistry at HARvard Macromolecular Mechanics
DHAR	Dehydroascorbate
DOPE	Discrete Optimized Protein Energy
DSSP	Define Secondary Structure of Proteins
EBI	European Bioinformatics Institute
FLAP	5-lipoxygenase activating protein
*Gm*GSTU	Tau class GST from *Glycine max*
*Gm*GSTU2-2	Tau class GSTU2 from *Glycine max*
*Gm*GSTU10-10	Tau class GSTU10 from *Glycine max*
*Gm*GSTU4-4	Tau class GSTU4 from *Glycine max*
GSH	Glutathione
GSOH	Glutathione sulfenic acid
GST	Glutathione transferase
GSTU	Tau class glutathione transferase
LTC_4	Leukotriene C_4 synthase
MAPEG	Membrane-Associated Proteins in Eicosanoid and Glutathione metabolism
MARCCD	MAR Charge Coupled Device
MD	Molecular Dynamics simulations
MR	Molecular Replacement
MRP	Multidrug Resistance Protein
NADPH	Nicotinamide adenine dinucleotide phosphate
Nb-GSH	*S*-(*p*-nitrobenzyl)-glutathione
NCBI	National Centre for Biotechnology Information
PDB	Protein Data Bank
PGES	Prostaglandin E synthase
PEG	Polyethylene glycol
RMSD	Root mean square deviation
Sh14	Chimeric clone 14 of homologous *Gm*GSTUs
SSM	Secondary Structure Matching
UDP	Uridine diphosphate

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
Sec	Selenocysteine	U
Ser	Serine	S
Thr	Threonine	T
Trp	Tryptophan	W
Tyr	Tyrosine	Y
Val	Valine	V

ABSTRACT

Glutathione transferases (GSTs) are a diverse family of enzymes that catalyze the glutathione-dependent detoxification of toxic compounds. GSTs are responsible for the conjugation of the tripeptide glutathione (GSH) to a wide range of electrophilic substrates. These include industrial pollutants, drugs, genotoxic carcinogen metabolites, antibiotics, insecticides and herbicides. In light of applications in biomedicine and biotechnology as cellular detoxification agents, detailed structural and functional studies of GSTs are required.

Plant tau class GSTs play crucial catalytic and non-catalytic roles in cellular xenobiotic detoxification process in agronomically important crops. The abundant existence of GSTs in *Glycine max* and their ability to provide resistance to abiotic and biotic stresses such as herbicide tolerance is of great interest in agriculture because they provide effective and suitable tools for selective weed control. Structural and catalytic studies on tau class GST isoenzymes from *Glycine max* (*Gm*GSTU10-10, *Gm*GSTU chimeric clone 14 (Sh14), and *Gm*GSTU2-2) were performed. Crystal structures of *Gm*GSTU10-10 in complex with glutathione sulfenic acid (GSOH) and Sh14 in complex with *S*-(*p*-nitrobenzyl)-glutathione (*Nb*-GSH) were determined by molecular replacement at 1.6 Å and 1.75 Å, respectively. Major structural variations that affect substrate recognition and catalytic mechanism were revealed in the upper part of helix H4 and helix H9 of *Gm*GSTU10-10. Structural analysis of Sh14 showed that the $Trp^{114}Cys$ point mutation is responsible for the enhanced catalytic activity of the enzyme. Furthermore, two salt bridges that trigger an allosteric effect between the H-sites were identified at the dimer interface between Glu^{66} and Lys^{104}.

The 3D structure of *Gm*GSTU2-2 was predicted using homology modeling. Structural and phylogenetic analysis suggested *Gm*GSTU2-2 shares residues that are crucial for the catalytic activity of other tau class GSTs–Phe^{10}, Trp^{11}, Ser^{13}, Arg^{20}, Tyr^{30}, Leu^{37}, Lys^{40}, Lys^{53}, Ile^{54}, Glu^{66} and Ser^{67}. This indicates that the catalytic and ligand binding site in *Gm*GSTU2-2 are well-conserved. Nevertheless, at the ligandin binding site a significant variation was observed. Tyr^{32} is replaced by Ser^{32} in *Gm*GSTU2-2 and this

may affect the ligand recognition and binding properties of *Gm*GSTU2-2. Moreover, docking studies revealed important amino acid residues in the hydrophobic binding site that can affect the substrate specificity of the enzyme. Phe^{10}, Pro^{12}, Phe^{15}, Leu^{37}, Phe^{107}, Trp^{114}, Trp^{163}, Phe^{208}, Ile^{212}, and Phe^{216} could form the hydrophobic ligand binding site and bind fluorodifen. Additionally, side chains of Arg^{111} and Lys^{215} could stabilize the binding through hydrogen bonds with the $–NO_2$ groups of fluorodifen.

GST gene family from the pathogenic soil bacterium *Agrobacterium tumefaciens* C58 was characterized and eight GST-like proteins in *A. tumefaciens* (*Atu*GSTs) were identified. Phylogenetic analysis revealed that four members of *Atu*GSTs belong to a previously recognized bacterial beta GST class and one member to theta class. Nevertheless, three *Atu*GSTs do not belong to any previously known GST classes. The 3D structures of *Atu*GSTs were predicted using homology modeling. Comparative structural and sequence analysis of the *Atu*GSTs showed local sequence and structural characteristics between different GST isoenzymes and classes. Interactions at the G-site are conserved, however, significant variations were seen at the active site and the H5b helix at the C-terminal domain. H5b contributes to the formation of the hydrophobic ligand binding site and is responsible for recognition of the electrophilic moiety of the xenobiotic. It is noted that the position of H5b varies among models, thus providing different specificities. Moreover, *Atu*GSTs appear to form functional dimers through diverse modes. *Atu*GST1, *Atu*GST3, *Atu*GST4 and *Atu*GST8 use hydrophobic ‘lock–and–key’-like motifs whereas the dimer interface of *Atu*GST2, *Atu*GST5, *Atu*GST6 and *Atu*GST7 is dominated by polar interactions. These results suggested that *Atu*GSTs could be involved in a broad range of biological functions including stress tolerance and detoxification of toxic compounds.

TIIVISTELMÄ

Glutationitransferaasit (GST:t) ovat monimuotoinen perhe entsyymejä, jotka katalysoivat myrkyllisten yhdisteiden glutationi-riippuvaista detoksikaatiota. GST:t ovat vastuussa tripeptidiglutationin (GSH) konjugaatiosta monenlaisiin elektrofiilisiin substraatteihin. Näihin sisältyvät teollisuuden saasteita, lääkkeitä, genotoksisia karsinogeenien metaboliitteja, antibiootteja, hyönteismyrkkyjä ja kasvimyrkkyjä. Ajatellen biolääketieteen ja biotekniikan sovelluksia, kuten solujen detoksifikaatio-aineita, tarvitaan GST:ien rakenteiden ja toimintojen tutkimuksia.

Kasvien tau-luokan GST:illa on erittäin tärkeä merkitys maataloudelle tärkeiden viljelykasvien katalyyttisissä ja ei- katalyyttisissä solujen xenobioottisissa detoksifikaatio-prosesseissa. GST:ien runsas esiintyminen ”*Glycine max”*:issa ja niiden kyky tarjota vastustuskykyä abioottiselle ja bioottiselle stressille, kuten kasvimyrkkyjen sietämiselle, on erittäin kiinnostavaa maataloudessa. GST:t siis tarjoavat tehokkaita ja sopivia työkaluja valikoivaan rikkakasvien torjuntaan. ”*Glycine max”*:in (*Gm*GSTU10-10, *Gm*GSTU kimeerinen klooni 14 (Sh14), ja *Gm*GSTU2-2) tau-luokan GST-isoentsyymien rakenteellisia ja katalyyttisiä tutkimuksia suoritettiin. *Gm*GSTU10-10:n kiderakenteet, yhdistyneenä glutationi-sulfeenihappoon (GSOH) ja Sh14:n kiderakenteet, yhdistyneenä *S*-(*p*-nitrobentsyyli)-glutationin (*Nb*-GSH) kanssa, määritettiin molekyylien korvauksella 1,6 Å ja 1,75 Å:ssa, vastaavasti. Suuria rakenteellisia muutoksia, jotka vaikuttavat substraattien tunnistamiseen ja katalyyttiseen mekanismiin, ilmeni *Gm*GSTU10-10:n kaksoiskierteen H4:n ja H9:n yläosassa. Sh14:n rakenteellinen analyysi osoitti, että $Trp^{114}Cys$ pistemutaatio on vastuussa entsyymin katalyyttisen aktiivisuuden parantamisesta. Lisäksi havaittiin, että kaksi suolasiltaa dimeerin rajapinnassa Glu^{66}:n ja Lys^{104}:n välissä aiheuttavat allosteerisen efektin H-kohtien välillä.

*Gm*GSTU2-2:n 3D-rakenne laskettiin käyttämällä homologiamallinnusta. Rakenteelliset ja fylogeneettiset analyysit viittasivat, että *Gm*GSTU2-2:lla on yhteisiä osia, jotka ovat ratkaisevia muiden tau-luokan GST–Phe^{10}, Trp^{11}, Ser^{13}, Arg^{20}, Tyr^{30}, Leu^{37}, Lys^{40}, Lys^{53}, Ile^{54}, Glu^{66} ja Ser^{67} katalyyttiselle aktiivisuudelle. . Tämä viittaa hyvin säilyneisiin katalyyttisiin ja ligandien sitoutumiskohtiin *Gm*GSTU2-2:ssa. Kuitenkin, ligandin sitoutumiskohdissa havaittiin merkittäviä eroja. Tyr^{32} korvataan

Ser^{32}:llä *Gm*GSTU2-2:ssa ja tämä voi vaikuttaa ligandin tunnistukseen ja *Gm*GSTU2-2:n sitoutumisominaisuuksiin.. Lisäksi telakointi-tutkimukset paljastivat tärkeitä aminohappotähteitä hydrofobisessa sitoutumiskohdassa, joka voivat vaikuttaa entsyymin substraattispesifisyyteen. Phe^{10}, Pro^{12}, Phe^{15}, Leu^{37}, Phe^{107}, Trp^{114}, Trp^{163}, Phe^{208}, Ile^{212} ja Phe^{216} voisivat muodostaa hydrofobisen ligandin sitoutumiskohdan ja sitoa fluorodifeenia. Lisäksi Arg^{111}:n ja Lys^{215}:n sivuketjut voisivat stabiloida sitoutumista vetysidosten avulla fluorodifeenin $–NO_2$ ryhmien kanssa.

GST-geeniperhe patogeenisestä maaperäbakteerista, *Agrobacterium tumefaciens* C58:sta karakterisoitiin ja kahdeksan GST-kaltaista proteiinia tunnistettiin *A. tumefaciens*:ssa (*Atu*GSTs). Fylogeneettinen analyysi paljasti, että neljä *Atu*GST:ista kuuluvat aikaisemmin tunnistettuun bakteerien beta-GST-luokkaan ja yksi theta-luokkaan. Kuitenkin, kolme *Atu*GST:ia eivät kuulu mihinkään aikaisemmin tunnettuun GST-luokkiin. *Atu*GST:ien 3D-rakenteet mallinnettiin käyttämällä homologiamallinnusta. *Atu*GST:ien vertaileva rakenne- ja sekvenssianalyysi s osoittivat paikallista sekvenssiominaisuuksia ja rakenteellisia ominaisuuksia eri GST-isoentsyymien ja luokkien välillä. Vuorovaikutukset G-kohdassa ovat säilyneet, mutta kuitenkin huomattavia eroja nähtiin aktiivisessa kohdassa ja H5b:n kaksoiskierteessä C-pään domeenissa. H5b edistää hydrofobisen ligandin sitoutumiskohdan muodostumista ja vastaa vierasaineen elektrofiilisen osan tunnistamisesta. On huomattu, että H5b:n paikka vaihtelee mallien välillä, jolloin muodostuu erilaisia ominaisuuksia. Lisäksi *Atu*GST:t näyttävät muodostavan toiminnallisia dimeerejä eri tavoilla. *Atu*GST1, *Atu*GST3, *Atu*GST4 ja *Atu*GST8 käyttävät hydrofobisia "lukko- ja avain" -kaltaisia rakennusosia, kun taas *Atu*GST2:n, *Atu*GST5:n, *Atu*GST6:n ja *Atu*GST7:n dimeeri-yhtymäkohdissa hallitsevat polaariset vuorovaikutukset. Nämä tulokset viittaavat siihen, että *Atu*GST:t voisivat olla mukana laajassa valikoimassa biologisia toimintoja, kuten stressinsietokyky ja myrkyllisten yhdisteiden detoksikointi.

TABLE OF CONTENTS

1. LITERATURE REVIEW

1.1 Introduction

Detoxification is a crucial metabolic process that provides protection to all living organisms against harmful agents. Glutathione transferases (GSTs; E.C. 2.5.1.18) are one of the most important families of enzymes in nature playing a significant role in cellular detoxification processes. The classic role of GSTs is to catalyze the conjugation of non-polar compounds that contain an electrophilic center to reduced glutathione (GSH); consequently, the resultant compounds are rendered to be more soluble and easy for cellular excretion. GSTs contribute to the metabolism of pesticides, herbicide, pharmaceuticals, and other foreign compounds (xenobiotics), including products of oxidative stress, such as peroxides [1-5].

Advancements in science in the last decades, such as the availability of large-scale genome data, advanced bioinformatics tools, expressed sequence tag (EST) databases, and the three-dimensional (3D) structures determination of enzymes by X-ray crystallography and *in silico* techniques have immensely broadened our knowledge of structure–function relationships in this important enzyme superfamily [6-10]. Currently, there are more than 37,000 entries at the PubMed MEDLINE database (www.ncbi.nlm.nih.gov/pubmed) for “Glutathione transferases” (and around 124,000 entries for “Glutathione”), indicating the importance of GSTs. This focus of interest on GSTs is owed to their biochemical, physiological, and toxicological features to agricultural, medicinal and clinical aspects, including herbicide bioremediation, carcinogenesis, and drug resistance [11].

1.1.1 Enzymatic detoxification systems

Living organisms are inevitably exposed to a myriad of xenobiotics during their lifetime, including a variety of toxic industrial chemicals, food components, pesticides, insecticides, herbicides and drugs. Most of the xenobiotic compounds show no or little relationship to previously metabolized compounds; nevertheless, living organisms have managed to detoxify themselves from harmful environmental exposures through highly evolved complex enzymatic detoxification systems. These

enzymatic systems have proven to be sufficient in protecting the organism from the potential damages that can be caused by toxic xenobiotics [12,13].

1.1.2 Discovery of detoxification reactions

The first detoxification reaction hypothesis: "*Xenobiotics consumed by animals are transformed to water-soluble substances and excreted through the urine.*" was officially tested in 1842 when the German chemist Wilhem Keller consumed a dose of *benzoic acid* and subsequently excreted *hippuric acid* in his urine [14]. This experiment shed light on how the body managed to remove various xenobiotics. Since then a variety of metabolites and a wide range of conjugation reactions and compounds including *taurine*, *sulfate*, *glycine*, *ornithine*, and *glutathione* have been identified [15].

1.1.3 Phases of enzymatic detoxification

The enzymatic transformation of toxic lipophilic (oil-like) xenobiotics to non-toxic or less toxic polar (water-soluble) compounds proceeds in three phases (I–III) [16-18]. In Phase I, functionalization, the parent compound is transformed to a more polar compound by the introduction of a reactive and polar group to form functional groups such as –SH, –OH, $–NH_2$, ±OH, and ±COOH. The main reactions that occur during this phase include deamination, O-, S- and N-dealkylation, S- and N-oxidation, epoxidation, peroxidation, aliphatic and aromatic hydroxylation. The cytochrome P450 (CYPs) enzymes are major players at this phase of xenobiotics biotransformation [19-21]. The compounds that have undergone Phase I are more polar than their parent compound, thus they can go to Phase II – conjugation. Phase II enzymes are responsible for forming a highly hydrophilic compound that is readily eliminated from the cell. Reactions at this phase are performed mainly by transferases including methyltransferases, sulfotransferases, glutathione transferases, and UDP-glucuronosyltransferases [22-24]. The main task of these enzymes is to perform conjugation reactions such as methylation, sulfation, amino acid and glutathione conjugation, etc. The final phase of enzymatic detoxification is export – Phase III. This phase uses specific transporters to excrete the newly formed Phase II products out of the cell [25].

In an attempt to understand detoxification systems and how living organisms manage to degrade such a wide range of compounds, the focus of research efforts has shifted in the past years into investigating protein structures and regulation of various enzymes involved in detoxification. Today, it is a well-known fact that living organisms use a battery of highly specific enzymes to manage this challenge. And GSTs are one of the major enzymes involved in enzymatic detoxification systems [26].

1.1.4 Xenobiotic detoxification in plants

Plants metabolize xenobiotics into reduced phytotoxic products using a variety of enzymes [27]. They achieve this in three phases (Fig. 1):

(I) Functionalization: Initial reactions–oxidation, reduction or hydrolysis, that serve to provide a functional group that is suitable for subsequent conjugation;

(II) Conjugation: Primary conjugation to an endogenous moiety such as GSH, sugars, or an amino acid;

(III) Export: Secondary conjugation, degradation, and compartmentation.

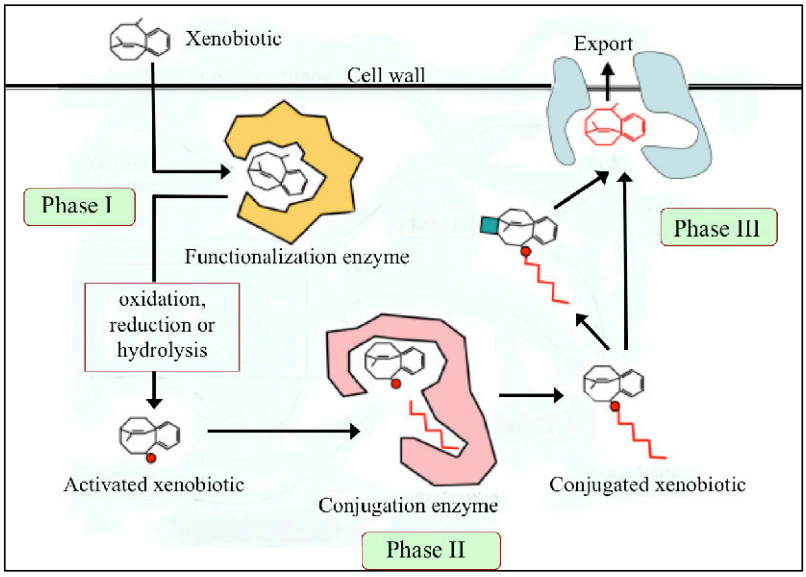

Fig. 1 The three phases of the enzymatic detoxification system in a hypothetical pathway representing the metabolism of a xenobiotic in plant tissues. Phase I, functionalization–activation of the xenobiotic by oxidation, reduction or hydrolysis; phase II, conjugation with an endogenous moiety; phase III, export of the conjugate into the cell wall or within the vacuole.

The role of P450 enzymes in xenobiotics oxidation

Plant Cytochrome P450 monooxygenases play a vital role in oxidation of xenobiotics. The most common Cytochrome P450-mediated reactions with xenobiotics include alkyl hydroxylation, aryl hydroxylation, and heteroatom release (N-demethylation). For catalytic activity NADPH, NADPH-Cytochrome P450 reductase and molecular oxygen are required [17,18]. The overall reaction is:

$$XH + O_2 + NADPH \longrightarrow XOH + H_2O + NADP^+$$

(*where XH is the substrate (xenobiotic) and XOH is the oxidized product*)

Sugar conjugation

Common sugars such as glucose are able to conjugate to a xenobiotic (X) consisting of phenolic (HO-X), carboxylic acid (COOH-X), or *N*-arylamile (H_2N-X) functional groups or to xenobiotics that passed Phase I. Enzymes such as UDP-glucosyltransferase, which requires UDP-Glucose as sugar substrate, catalyze these reactions [28]:

$$\text{UDP-Glucose} + \text{HO-X } (H_2N\text{-X, HOOC}) \longrightarrow \text{Glucose-O-X } (\text{Glucose-O-}\overset{\overset{O}{\|}}{C}\text{-X}) + \text{UDP}$$

Glucosyltransferases, which catalyze the conjugation reaction, are very specific for UDP-Glucose. However, they have a wide range of substrate specificities towards xenobiotics; hence, facilitating the glucosylation of xenobiotics in plant glycosyltransferases [29,30].

Amino acid conjugation

Amino acids are versatile secondary metabolites that play essential roles in enzymatic detoxification process. They can effectively conjugate to a broad range of toxic organic compounds. Acidic herbicides such as phenoxyacetic acids are known to form

amino acid conjugates. The first evidence of amino acid-xenobiotic conjugation was observed in 2,4-dichlorophenoxyacetic acid (2,4-D) with aspartic acid in peas [31]. The evidence revealed that an alpha bond conjugates 2,4-D with the amino acid. Since then, several amino acids including glutamic acid, tryptophan, leucine, phenylalanine, and valine are known to form conjugates with xenobiotics [32]. Aspartic or glutamic acids have been observed to be the most common amino acids conjugated with xenobiotics in plants (Fig. 2) [33].

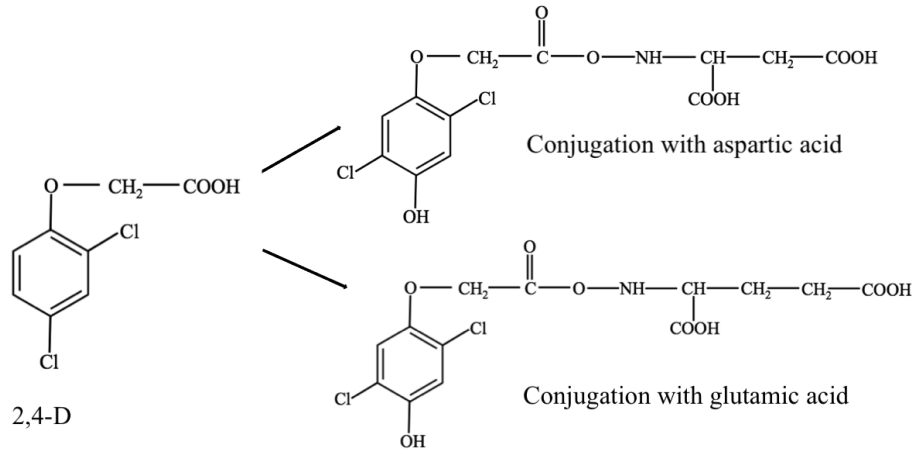

Fig. 2 Conjugation of 2,4-D with aspartic and glutamic acids.

GSH conjugation

The small peptide and major cellular thiol, γ-Glutamyl-Cys-Gly (GSH), plays an essential role in the metabolism of electrophilic chemicals through GSH conjugation reaction (Fig. 3). The conjugation involves a nucleophilic displacement reaction with the GSH anion (GS^-) serving as the nucleophile. The sulfhydryl group (–SH) on the cysteine residue of the GSH is the reactive component required for this reaction. These reactions are catalyzed by glutathione transferases (GSTs) [34]. A simplified form of the reaction is as follows:

$$X\text{-}Z + GS^- \longrightarrow GS\text{-}X + Z^-$$

(X is xenobiotic and Z^- the leaving group)

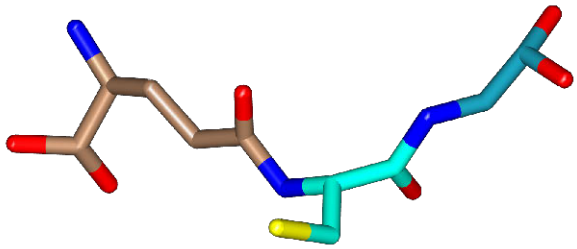

Fig. 3 Stick representation of GSH. The glutamyl, cysteinyl, and glycine part of GSH are colored pale brown, cyan, and dark cyan respectively. The nucleophilic sulfhydryl group is colored yellow.

Due to the fact that most herbicides possess an electrophilic center such as sulfur, nitrogen or carbon, they can directly conjugate with GSH whereas sugar conjugation reactions require hydrolysis or oxidation before the conjugation reaction. Once a water-soluble GSH-conjugate is formed, it can easily be excreted via multidrug resistance protein (MRP) efflux pumps. Furthermore, GSTs are capable of detoxifying free radicals through their glutathione peroxidase activity [35,36].

1.2 GST Overview

Glutathione transferases (GSTs) are an ancient and large superfamily of homo- and hetero-dimeric proteins that mediate the catalytic binding of glutathione to an array of endo- and exobiotic compounds as a general detoxification scheme [3,37]. In addition to their classic role, i.e. GSH conjugation as a general detox reaction, GSTs are involved in a broad range of cellular reactions and processes such as isomerization, reduction, transport, storage, prostaglandin biosynthesis, thiol transfer, stress kinase regulation, cell apoptosis, and peroxidase activity [37-44]. GSTs are currently of great interest in drug discovery, nanotechnology and green biotechnology owing to their involvement in major cellular processes. GSTs are widely distributed in eukaryotes and prokaryotes and it has been reported that they may have evolved in aerobic bacteria to protect cells against oxygen toxicity [45].

Microsomal (MAPEGs), mitochondrial, and cytosolic GSTs comprise the three major groups of GSTs. MAPEGs display peroxidase activity in addition to their role in providing cellular defense against reactive electrophilic compounds. They are reported to protect the endoplasmic reticulum and outer mitochondrial membrane

against oxidative stress [46]. Mitochondrial GSTs are divided into four classes: alpha, kappa, mu, and pi. Kappa class is unique to mitochondrial GSTs [47]. Members of MAPEG family of proteins are MAPEG 1, MAPEG 2, MAPEG 3, 5-lipoxygenase activating protein (FLAP), leukotriene C_4 synthase (LTC_4), and prostaglandin E synthase (PGES) [48]. MAPEGs are structurally distinct compared to mitochondrial and cytosolic GSTs. On the other hand, several previous studies had shown that mitochondrial and cytosolic GSTs possess similar structure and catalytic activity [49-51].

1.2.1 Cytosolic GSTs

Cytosolic GSTs are abundantly found in all aerobic organisms and they have been studied extensively. In recent years, numerous cytosolic GST genes from mammals (15–20 genes), insects (over 10 genes), plants (40–60 genes) and bacteria (10–15 genes), fungi, yeast, and fish have been identified, purified and characterized [4,52,53]. Presently, seven classes of cytosolic GSTs are recognized in mammals (alpha, mu, omega, pi, sigma, theta, and zeta), six classes in insects (delta, epsilon, omega, sigma, theta, and zeta), six classes in plants (DHAR (dehydroascorbate), lambda, phi, tau, theta, and zeta), three classes in fungi (alpha, gamma, and mu), and five classes in bacteria (beta, chi, rho, theta, and zeta) (Table 1; Fig. 4) [54-56]. Cytosolic GSTs share >40% sequence identity within classes but <25% between classes [3].

Table 1. Common and specific cytosolic GSTs classes in major taxa.

Taxa	Common	Specific
Mammals	Omega, Sigma, Theta, Zeta	Alpha, Mu, Pi
Insects	Epsilon, Omega, Sigma, Theta, Zeta	Delta
Fungi	Alpha, Gamma, Mu	-
Plants	Theta, Zeta	DHAR, Lambda, Phi, Tau
Bacteria	Chi, Rho, Theta, Zeta	Beta

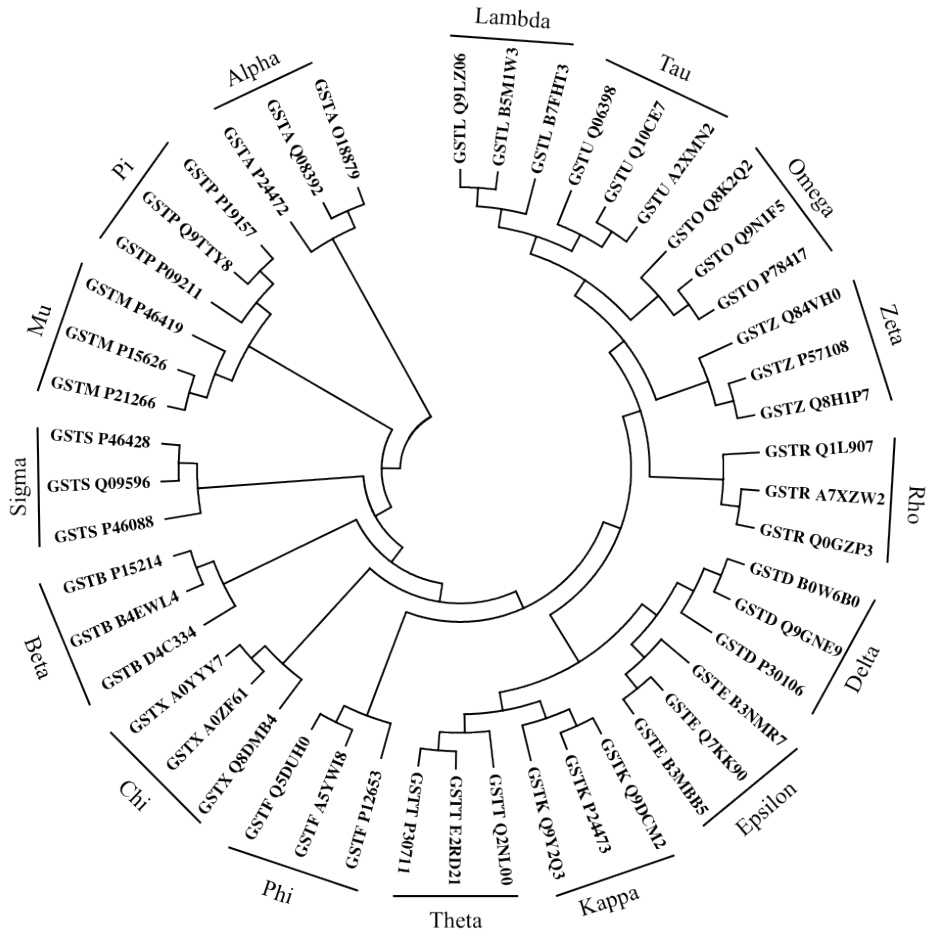

Fig. 4 Topology of the phylogenetic relationship of representative GSTs. The diagram was constructed using representative members from all known mainstream GST classes: GSTA (Alpha), GSTB (Beta), GSTD (Delta), GSTE (Epsilon), GSTF (Phi), GSTK (Kappa), GSTL (Lambda), GSTM (Mu), GSTO (Omega), GSTP (Pi), GSTR (Rho), GSTS (Sigma), GSTT (Theta), GSTU (Tau), GSTX (Chi), and GSTZ (Zeta). Labels of the GSTs in the figure are shown as GSTClass UniProt accession number [57]. Sequence alignments, phylogenetic analysis and tree generation were done using MEGA 6 [58].

1.2.2 Plant GSTs

Since their discovery in the 1970s, many GST isoenzymes have been characterized, cloned, and their activities towards various herbicides described. In both crops and weeds, GSTs catalyze glutathione-dependent reactions with electrophilic herbicides. ATP binding cassette (ABC) transporters tag the *S*-glutathionylated metabolites for vacuolar import; consequently, the GSH conjugates are transported selectively. GSTs

are known to have overlapping substrate specificities despite being distinct in their regulation [59-61].

There are significant differences in the responsiveness of various plants species towards some herbicides. These species-specific differences in metabolism are the result of differences in GST isoenzymes properties. Conjugation of phenylpropanoids such as p-coumaric and trans-cinnamic acid with GSH has been shown in numerous *in vitro* studies; however, the physiological substrates of plant GSTs are yet to be identified [62].

Plant GSTs have been the focus of research in red and green biotechnology for their herbicide detoxification properties [63,64]. Phi and tau class GSTs, which are specific to plants, are responsible for the detoxification of several herbicides. It has been reported that phi class GSTs detoxify chloroacetanilides and thiocarbamate herbicides whilst tau class GSTs detoxify aryloxyphenoxypropinates and diphenylether herbicides, thus showing the class specificities in herbicide preference [65].

As all GSTs, tau class glutathione transferases (GSTUs) are also important in providing protection for plants against oxidative damage that can be induced by various abiotic and biotic stresses [62]. Plant GSTUs have been identified in a range of major crops, including maize, wheat, and soybean. Several tau class GSTs from soybean (*Glycine max; Gm*GSTU) have been identified and characterized. *Gm*GSTUs are active as GSH-dependent peroxidase, e.g. they display high GSH peroxidase activity towards hydroperoxides of linolenic acid and arachidonic acid [66]. Moreover, the role of *Gm*GSTU isoenzymes (e.g. *Gm*GSTU4) in the detoxification of diphenylether and chloroacetanilide herbicides such as fluorodifen and alachor, respectively, has been studied extensively (Fig. 5). The evidence of the functional activity of *Gm*GSTU4 isoenzyme can be used to develop transgenic plants with better biotransformation capabilities [67,68].

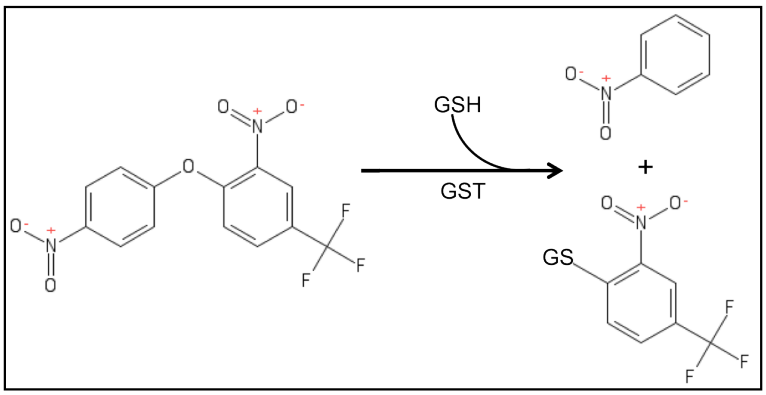

Fig. 5 Fluorodifen conjugation reaction. Fluorodifen is initially degraded by ether cleavage followed by conjugation through GST mediation.

1.2.3 Bacterial GSTs

Bacterial GSTs are found in aerobic prokaryotes such as soil bacteria and plant pathogens. The existence of GST activity was first reported in *Escherichia coli.* GSTs and GSH share the same evolutionary path and the absence of GSH from anaerobic bacteria is notably the evidence [69,70]. Among the five classes of cytosolic bacterial GSTs identified to date (beta, chi, rho, theta, and zeta), beta class GSTs, which are specific to bacteria, have been studied in depth. Beta class GSTs display GSH conjugation activity towards numerous molecules such as 1-chloro-2,4-dinitrobenzene (CDNB) and they play major role in the detoxification of antimicrobial agents. Additionally, bacterial GSTs have demonstrated an ability to promote the degradation of several monocyclic and polycyclic aromatic compounds that are toxic to human health and the environment (Fig. 6) [71-73].

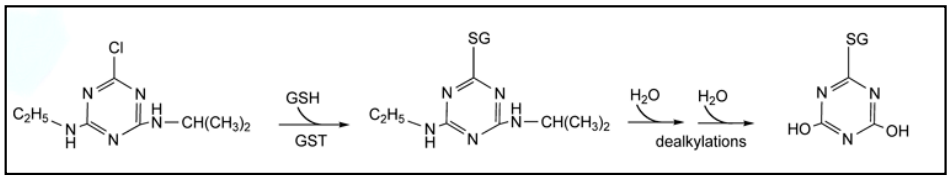

Fig. 6 Proposed pathway of atrazine degradation by bacterial GSTs. GSTs are involved in the initial step of atrazine biodegradation. Firstly, they cleave the chlorine atom (dechlorination) by forming atrazine–GSH conjugate. Then, isopropylamine and ethylamine groups are removed step by step by dealkylation using hydrolase enzymes [74-78].

In recent years, a number of studies are focused on identification and characterization of novel bacterial GSTs of biotechnological and agronomical importance. In particular, GSTs from *Agrobacterium tumefaciens* – a pathogenic soil bacterium, is a good example. *A. tumefaciens* is a novel tumor-inducing factor and causes large tumors (crown gall tumor disease) in a wide range of plants as the culprit behind major costly agronomical losses worldwide affecting more than 93 plant families (Fig. 7) [79-81].

Fig. 7 (Left) *A. tumefaciens* colonizing tobacco leaves. Originally published in de la Riva *et al.* [82]. **(Right)** Crown gall tumor caused by *A. tumefaciens*. Originally published in Escobar *et al.* [83].

Agrobacterium species are soil-borne gram-negative bacteria that are able to transfer genes between kingdoms – interkingdom gene transfer. They are responsible for the initiation of tumor formation on mainly dicotyledonous and monocotyledonous species [84,85]. It has been reported that the tumors do not require the continuous presence of the bacteria for proliferation, confirming the capability of *Agrobacterium* for gene transformation of the plant cells [86]. Evidently, structural studies of *Agrobacterium*-plant interaction will shed light on pathogen strategies for host infection, thereby providing information to the development of stress-resistant transgenic plants.

1.3 GST structure

Cytosolic GSTs exist as dimers and they have a strictly conserved secondary and tertiary structure. The biologically active dimers (homo- or hetero-) are made of subunits of ~25–30 kDa (~200–250 amino acids). The existence of a large inter-subunit cleft of a varying size is a typical characteristic of GSTs. A GST monomer is comprised of a GSH binding domain (N-terminal domain, Domain I) and a hydrophobic substrate-binding domain (C-terminal domain, Domain II). The N-terminal domain exhibits a typical thioredoxin superfamily fold, found also in glutathione peroxidase and glutaredoxin [87,88]. This domain is made up of a β1–α1–β2–α2–β3–β4–α3 structural motif with the four-stranded β-sheet flanked by the two α-helices (α1, α2). Moreover, the presence of an irregular loop (known as *cis*-Pro loop) connecting α2 and β3 is consistent with classical thioredoxins (Fig. 8a). The N-terminal domain provides the GSH binding site [53,89]. Nevertheless, the C-terminal domain is exclusively made up of α-helices and it is composed of five major helices (α4–α8). However, some GST classes (alpha, omega, tau, and theta) can have an extra helix (α9). The C-terminal domain is responsible for defining isoenzymes specificities and recognizing substrates [1,4]. A short linker sequence connects the N-terminal domain with the C-terminal domain. An inter-subunit interaction exists between the N-terminal domain (Domain I) of one subunit and the C-terminal domain (Domain II) of the other subunit. Residues from helices α3 and α4 are involved in these interactions (Fig. 8b). These dynamic lock-and-key types of interactions may enhance protein stability and provide significant catalytic efficiency for the active site [68,89-93].

1.3.1 GSH binding site

The GSH binding site (G-site), which is highly conserved in GSTs, is the essential feature of GSTs' catalytic mechanism. It facilitates the activation of the sulfur of the GSH for nucleophilic attack by primarily defining the GSH binding interactions. The tripeptide GSH that runs antiparallel to the loop preceding strand β3 is bound in an extended conformation and residues from the β3–β4–α3 motif affix it through a network of hydrogen bonds (Fig. 8a). The γ-glutamyl moiety of GSH is found pointing down towards the dimer interface. Helix α2 exhibits significant variations between classes and it hosts the catalytic conserved residues that form hydrogen

interactions with the glycine residue of GSH. These hydrogen bond interactions are class-specific [94].

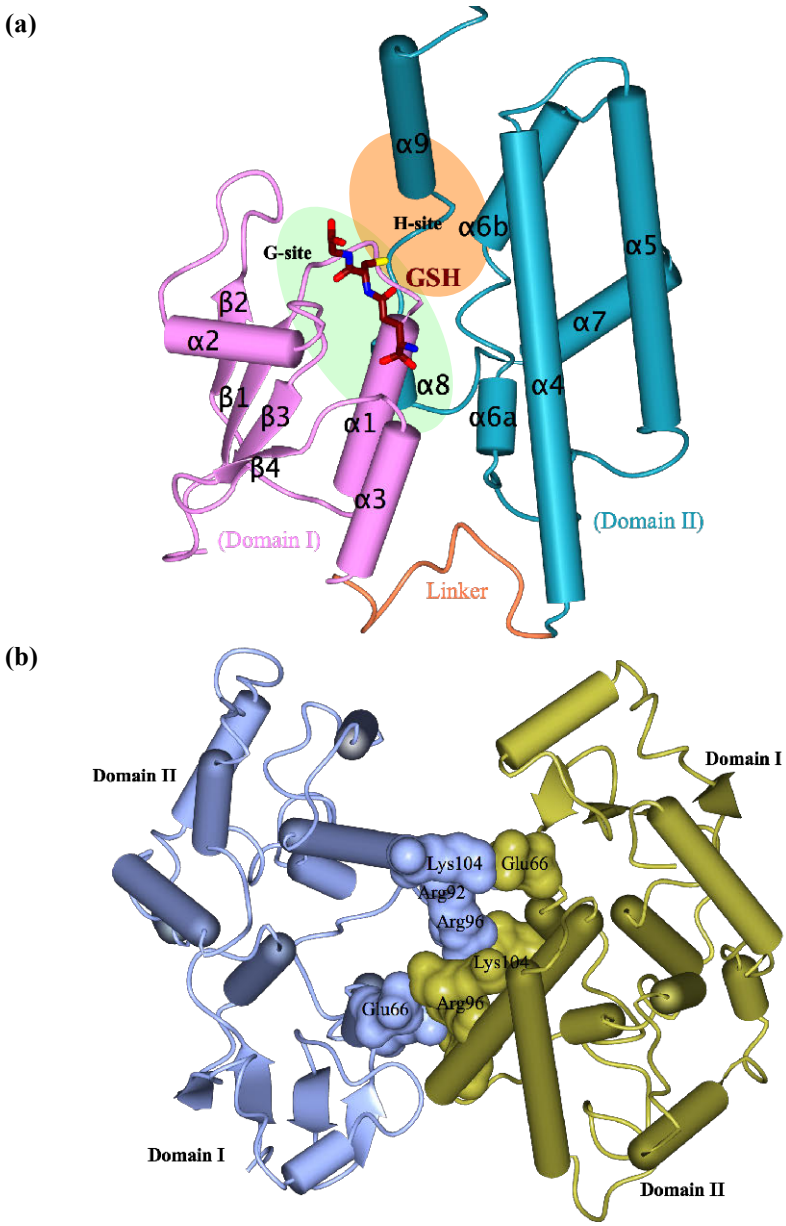

Fig. 8 Tertiary structure of a GST enzyme using *Gm*GSTU4-4 (PDB ID: 4TOP) as an example of the overall architecture in a GST structure. **(a)** 3D structure of the GST monomer, consisting of two domains: Domain I for binding to GSH (colored pink) and Domain II for hydrophobic substrates (colored dark cyan). The light green and light orange spheres depict the G-site and H-site, respectively. The linker is colored coral. **(b)** Inter-subunit interactions (salt bridges) between GST subunits in the GST dimer. Ice blue and gold indicate the two monomers. Interacting residues are shown as surfaces and colored according to their respective subunits. Figures were generated using CCP4mg [95].

Common characteristics of the G-site are the existence of (Fig. 9):

(a) hydrogen bond interactions between the cysteine residue of GSH and residues from the main chain of the enzyme.

(b) hydrogen bond interactions between the γ-glutamyl residue of GSH and two residues from the loop between α3 and β4 of the enzyme. These two residues are either serine/threonine or glutamate/glutamine.

(c) hydrogen bond interaction between the sulfur atom of GSH and a conserved catalytic residue in the protein.

Previous spectroscopic experiments have revealed that the nucleophilic GS^- is bound at the active site instead of GSH [96,97]. The hydrogen bond interactions stabilize the activated GSH (GS^-), hence playing a crucial role in catalysis. Classes of cytosolic GSTs are characterized by the presence of a conserved catalytic residue that forms a hydrogen bond with GS^-: serine in theta, tau and zeta; tyrosine in alpha, mu, pi, and sigma; cysteine in beta and omega class enzymes.

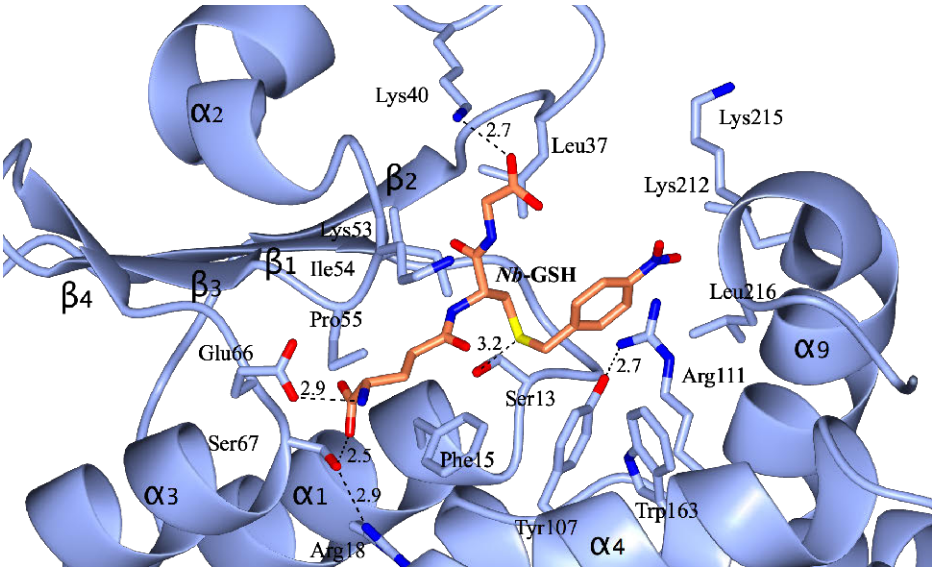

Fig. 9 Close-up view of the GSH and electrophilic substrate binding sites of *Gm*GSTU4-4–*Nb*-GSH (PDB ID: 2VO4) [68]. Hydrogen bonds are shown as dash lines with their corresponding distances (in Å). Interacting side-chains of the protein and *Nb*-GSH are shown as sticks and colored according to atom type. Cutoff distance was 3.5 Å.

1.3.2 Substrate binding site

The hydrophobic substrate-binding site (H-site) is located adjacent to the GSH binding site (G-site). This site shows high variability, and thus it defines the binding site with distinct physicochemical properties, such as hydrophobicity, shape, and size. Residues from the H-site are responsible for binding a hydrophobic substrate through hydrogen bond interactions. Additionally, certain residues from the H-site may contribute (occurring in helix α4) to the GSH binding site. Basically, the loop between the strand β1 and the helix α1, helix α4, and/or the C-terminal tail forms the H-site (Fig. 9). The C-terminal domain shows major variations between classes and it is suggested to be the reason for the differences in substrate specificities, and hence, substrate recognition [3].

1.3.3 Induced-fit mechanism

Conformational changes of the active site upon ligand binding are common features of many GST isoenzymes. Both the G- and H-sites can change conformations to snugly bind the ligand, giving rise to an induced-fit mechanism. When a suitable ligand enters the active site, it forms weak bonds with the protein, thus inducing the protein to fold in a manner to accommodate the ligand. Numerous biochemical and structural studies on GST enzymes confirm the existence of this catalytically important mechanism (e.g. in maize GSTF1, soybean GSTU4-4, human GSTP1-1, and GSTA1-1 enzymes) [68,75,98,99]. Structural comparison of the ligand-bound and apo structures of maize GSTF revealed the movement of a ten-residue loop at the active site upon ligand binding. In addition, similar conformational changes were indicated in human GSTP1-1 at helix H2 (residues 42–50), H4, and C-terminal loop [100,101]. These structural rearrangements upon ligand binding confirm the induced-fit mechanism proposed for GST enzymes.

1.3.4 Ligandin function

GSTs are able to bind non-substrate ligands, called as its ligandin function, besides their widely known function of catalyzing GSH-xenobiotic conjugation for cellular detoxification. It has been postulated that the non-substrate ligand-binding site (L-site) is biologically important in cellular signaling [102]. The L-site may be responsible for the transfer and delivery of non-substrate ligands to specific receptors.

The exact location of this site varies between GSTs and its nature is not fully known. However, recent studies have reported the location of the L-site for tau class GSTs, human P1-1, *Schistosoma japonica* and *Arabidopsis thaliana* enzymes. Non-conserved residues Trp^{11}, Arg^{20}, Tyr^{32}, Leu^{32}, Leu^{199}, and Pro^{200} form the L-site of plant tau class GSTs. It is a hydrophobic pocket that is located on the surface of the enzyme [68,90]. An equivalent hydrophobic surface pocket that is formed by residues Pro^{9}, Val^{10}, Pro^{202}, and Gly^{205} is reported for human P1-1 enzyme [103] whereas, the L-site of human pi class and *S. japonica* GST is located at the H-site and the dimer interface [104,105]. In case of *A. thaliana* GST enzyme, the L-site is located between the side chains of amino acids from helices α3 and α5, which are located next to the G-site [106].

1.4 Applications

Considering the crucial biological roles GST enzymes play, there is a huge research interest for their potential versatile biomedical, biotechnological and nanotechnological applications (Table 2). Recent developments in biosciences and biotechnologies have tremendously increased our knowledge of the 3D structures, functions, evolution and roles of enzymes in a complex network of biological processes. Accordingly, the newly acquired knowledge can be the basis for the development of novel biomolecules of medicinal and industrial importance and creation of transgenic organisms with enhanced and/or novel properties.

The modular structure of GSTs is a suitable platform to bioengineer the enzymes in order to enhance and/or discover novel catalytic activities and hence to discover new drugs and biocatalysts [80]. Another potential area of interest is GST transgenesis. GST transgenic organisms have the potential to be used as “cell factories” to biosynthesize compounds of technological and medicinal applications. In addition, these organisms can also shed light on understanding the functions of a gene in *in vivo* experiments [63].

GSTs are targets of bioengineering research in regards to their ability to inactivate and sequester toxic electrophilic compounds, such as herbicides, explosives, carcinogens, and drugs [44]. Bacterial and plant GSTs are involved in antibiotic resistance and

herbicide resistance, respectively. In plants, biotransformation of herbicides and other environmental pollutants using engineered novel GSTs has shown a great promise. Weed control without harming the crop by selectively killing the weed with several classes of herbicide has been demonstrated for crops such as sorghum (*Sorghum bicolor*), maize (*Zea mays),* rice (*Oryza sativa*), wheat (*Triticum aestivum*), and barley (*Hordeum vulgare*) [107-112]. Bioengineering of vital crops for herbicide tolerance is essential in phytoremediation [53,113].

Meanwhile, in bacterial GSTs, recent studies have been/are being focused on generating strains with the ability to detoxify toxic xenobiotics from soil and water for use in bioremediation. Environmental pollutants that include various toxic organic compounds, radioactive elements, explosives, highly toxic metals, herbicide residues, crude oil, etc. may be degraded and eliminated by utilizing GSH-conjugation oriented detoxification systems [114]. Unfortunately, there is no suitable species both in plants and bacteria to detoxify certain highly toxic xenobiotics. In some cases, suitable plants or bacteria that can be used for bioremediation are available; however, they are not efficient enough in degradation and/or accumulation. These species, therefore, need to be engineered to increase their bioremediation capability significantly through genetic manipulation and forced evolution.

Table 2. Examples of biomedical and biotechnological applications of GSTs

Application	GST	Description
Biosensor	GST I from maize	Fiber-optic portable biosensor to determine a pesticide-like atrazine [115]
	*Gm*GSTU4-4 from Soy	Biosensor for the determination of chloroacetanilide herbicide alachor [116]
	GST I from maize	Potensiometric biosensor for the determination of malathion [117]
	GST-immobilized gel film	Optical biosensor for sensitive detection of the carcinogen captan in water supplies [118,119]
	GST-4 from *C. elegans*	Biosensor for acrylamide detection in starchy foods [120]
Bioassay	*Aa*GSTE2-2 from mosquito	Specific enzyme assay for the determination of DDT [121]
Biomarker	GSTP1-1	Immunohistological marker for gastric cancer and gliomas [122]
Drug and Pro-drug design	GSTA3-3	Production of drug for the treatment of steroid hormone dependent diseases (e.g. cancer) [123,124]
	GSTP1-1	Pro-drug activation (e.g. Telcyta) [125]
	GSTA1-1	Tumor-directed pro-drug activation [126]
Bioremediation	Phi GST from plants	Detoxification of chloroacetanilide and thiocarbamate herbicides [65]
	Tau GSTs from plants	Detoxification of diphenylethers and aryloxyphenoxypropionate herbicides [60,68,127]
Transgenesis	GST I gene from maize	Transgenic tobacco with high tolerance to the herbicide alachlor [128]
	GST-27 gene from maize	Transgenic wheat with high herbicide resistance [129]

2. OBJECTIVES AND AIMS OF THE STUDY

The existence of 3D structural information of potentially important enzymes and their mutants is crucial in understanding the catalytic activities and roles they play in biological processes. Moreover, structural information can be used for the development of enhanced versions of the target enzymes for biomedical and biotechnological applications. In this study, the key aim is to provide detailed structural information for tau class GSTs from *Glycine max* and the newly identified GSTs from *Agrobacterium tumefaciens* C58. The study project is comprised of the following subprojects:

Subproject I. Crystallographic structural studies

- Expression, purification, crystallization and crystal structure determination of a tau class native GST10 from *Glycine max* (*Gm*GSTU10-10) complexed with glutathione sulfenic acid (GSOH).
- Expression, purification, crystallization and crystal structure determination of a chimeric clone 14 of *Gm*GSTU homologues from *Glycine max* (Sh14) complexed with *S*-(*p*-nitrobenzyl)-glutathione (*Nb*-GSH).
- Structural comparative studies to elucidate substrate recognition and binding, and the kinetic and catalytic activities.

Subproject II. Structural bioinformatics studies

- Homology modeling, molecular docking and molecular dynamics simulations studies of GSTU2 from *Glycine max* (*Gm*GSTU2-2) with glutathione and fluorodifen.
- Molecular docking and molecular dynamics simulations studies of a chimeric clone 14 of *Gm*GSTU homologous from *Glycine max* (Sh14) with *Nb*-GSH.
- Homology modeling and structural characterization of eight *Agrobacterium tumefaciens* C58 GSTs (*Atu*GSTs): *Atu*GST1–*Atu*GST8.
- Structural comparative studies of the active site, role of the loops/residues at the active site, and catalytic activities.

3. MATERIALS AND METHODS

3.1 Expression and purification

3.1.1 *Gm*GSTU10-10–GSOH complex

E. coli cells harboring recombinant plasmid were grown at 37 °C in 1 L LB medium containing ampicillin (100 μg/mL). The synthesis of GST was induced by the addition of 1 mM isopropyl 1-thio-β-galactopyranoside when the absorbance at 600 nm reached 0.6. After 4 hours of induction, cells were harvested by centrifugation at 8,000 rpm for 15 min, resuspended in potassium phosphate buffer (0.1 M, pH 6.5), sonicated, and centrifuged at 13,000 rpm for 5 min. The supernatant was loaded to GSH-Sepharose column (1,4-butanediol diglycidyl ether-GSH-Sepharose-CL6B, 1 mL), which was previously equilibrated with potassium phosphate buffer (20 mM, pH 7). Non-adsorbed protein was washed off with 10 mL equilibration buffer. Bound GST was eluted with equilibration buffer containing 10 mM GSH. Protein purity was judged by SDS-PAGE.

3.1.2 Sh14–*Nb*-GSH complex

Sh14 was constructed with DNA shuffling using three *Gm*GSTU homologues: *Gm*GSTU2-2, *Gm*GSTU4-4, and *Gm*GSTU10-10 [68,130]. The preparation for DNA shuffling and construction of *Gm*GSTUs library was performed as described in [131] (PAPER III). Sh14 is derived from *Gm*GSTU4-4 with three segments from *Gm*GSTU10-10 and *Gm*GSTU2-2, resulting in three random mutations–$Arg^{38}Gln$, $Gln^{46}Cys$ and $Ile^{183}Val$ and a point mutation–$Trp^{114}Cys$.

E. coli M15[pREP4] cells harboring recombinant plasmid were grown at 37 °C in 1 L LB medium containing ampicillin (100 μg/mL) and kanamycin (50μg/mL). The synthesis of GST was induced by the addition of 1 mM isopropyl 1-thio-β-galactopyranoside when the absorbance at 600 nm was 0.6. After 4 hours of induction, cells were harvested by centrifugation at 8,000 rpm for 15 min, resuspended in potassium phosphate buffer (0.1 M, pH 6.5), sonicated, and centrifuged at 13,000 rpm for 5 min. Enzyme purification was carried out using

affinity chromatography as described in [68,90]. Non-adsorbed protein was washed off with 10 mL of equilibration buffer. Bound GST was eluted with equilibration buffer containing 10 mM GSH. Protein purity was judged by SDS-PAGE.

3.2 Crystallization

3.2.1 *Gm*GSTU10-10–GSOH complex

Prior to crystallization, purified *Gm*GSTU10-10 was concentrated to ~10 mg/mL in HEPES-NaOH buffer (pH 7.0). INDEX crystallization screen (Hampton Research) was used to establish initial crystallization conditions in the presence of 10 mM spirodiclofen dissolved in acetone solution. Small crystals appeared in Condition No. 82 (0.2 M $MgCl_2$, 0.1 M Bis–Tris, pH 5.5, PEG 3350 25% (w/v)). Subsequently, an optimization process using the hanging drop vapor diffusion method was implemented resulting in crystals of maximum size of 0.4 mm × 0.2 mm × 0.2 mm growing within 2 days (0.2 M $MgCl_2$, 0.1 M Bis–Tris, pH 5.6, PEG 3350 22% (w/v)) (Fig. 10).

3.2.2 Sh14–*Nb*-GSH complex

In order to investigate the effect of the mutations on ligand binding and structure of the protein, a purified Sh14 was concentrated to ~12 mg/mL in 10 mM HEPES buffer (pH 7.0) and mixed with a 100 mM *S*-(*p*-nitrobenzyl)-glutathione (*Nb*-GSH) stock solution (0.1 M Na/K phosphate buffer, pH 7.0), resulting in a final concentration of 10 mM *Nb*-GSH. A hanging drop vapor diffusion method (at 16 °C) with drops of 2.5 µL of reservoir solution (PEG 4K 27.5 % w/v, NH_4OAc 0.2 M in buffer NaOAc 0.1 M, pH 5.0) mixed with an equal volume of the enzyme solution resulted in the appearance of good crystals within 4–5 days (Fig. 10).

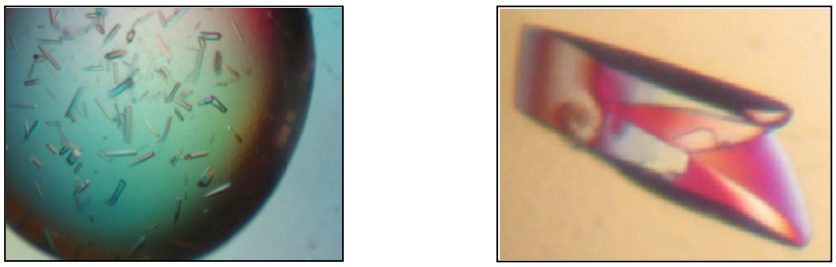

Fig. 10 Crystals of *Gm*GSTU10-10 complexed with GSOH **(left)** and a typical crystal of Sh14 grown in the presence of *Nb*-GSH **(right)**.

3.3 Data collection

3.3.1 *Gm*GSTU10-10–GSOH complex

X-ray diffraction data to 1.6 Å were collected from a single crystal at 100 K on station X13 (λ = 0.8123 Å) at EMBL Hamburg (c/o DESY, Germany) using a 165 mm MARCCD detector (distance 134.8 mm, 0.5° rotation angle per image with an exposure time of 10–15 seconds). During data collection, 20% (v/v) glycerol was used as cryoprotectant.

3.3.2 Sh14–*Nb*-GSH complex

X-ray diffraction data to 1.75 Å resolution were collected on the beamline BW7A (λ = 0.9537 Å) at EMBL Hamburg (c/o DESY, Germany) using a 165 mm MARCCD detector (distance 125.2 mm, 0.5° rotation angle at an exposure time of 3 seconds per image). Prior to data collection the crystal was briefly soaked in 20% (v/v) glycerol for cryoprotection.

3.4 Data processing

XDS [132] was used to autoindex, scale, merge and integrate raw diffraction data.

3.5 Structure determination and refinement

3.5.1 *Gm*GSTU10-10–GSOH complex

The CCP4 program suite [133] was used for structural determination of *Gm*GSTU10-10–GSOH complex. The structure was determined by molecular replacement (MR) using Phaser [134]. As a search model a *Glycine max* GST (PDB ID: 3FHS [68]), which has a 91% sequence identity with *Gm*GSTU10-10, was used. Phenix [135] was used for refinement of the structure. Building and rebuilding of the structure, and inspections of $2|F_o| - |F_c|$, and $|F_o| - |F_c|$ electron density maps generated after each Phenix refinement run were carried out using Coot [136]. Additionally, manual modeling of the structure, such as fitting of ligands, incorporation of solvent molecules, flipping of Asn and Gln residues and alternative conformations were carried out. Water molecules with B-factor higher than 60 $Å^2$ were removed from subsequent refinement runs. The eLBOW program [137] as implemented in Phenix

was used to generate ligand restraints for subsequent refinement and rebuilding rounds.

3.5.2 Sh14–*Nb*-GSH complex

Phaser, implemented within the CCP4 software suite, was used to obtain initial phases of the Sh14–*Nb*-GSH complex by molecular replacement [134]. Rice GST1 structure (PDB ID: 1OYJ, 51 % sequence identity with Sh14 [138] was used as a search model. ARP/wARP [139] was used for automated model building. Refmac [140] at the initial steps of refinement was used followed by Phenix [135] at the final refinement steps. Coot [136] was used for inspection of $2|F_o| - |F_c|$ and $|F_o| - |F_c|$ maps. Summary of data processing and refinement statistics is shown in Table 3.

3.6 Molecular modeling

3.6.1 *Gm*GSTU2-2 and *Atu*GSTs

Discovery Studio 3.0 and 2.5 (http://accelrys.com) were used to build the homology models for *Gm*GSTU2-2 and *Atu*GSTs, respectively. Crystal structures of *Gm*GSTU4-4 (PDB ID: 2VO4) and *Gm*GSTU10-10 (PDB ID: 4CHS) with 91% and 87 % sequence identity, respectively, were used as templates for homology modeling of *Gm*GSTU2-2 [68,130]. For *Atu*GSTs, BLAST searches [141] were carried out using the sequences provided to identify homologous proteins suitable for molecular modeling. MODELLER [142] as implemented within Discovery Studio was used to build the models. Model optimizations were performed using spatial restraints and DOPE energy function, followed by inspections of the models by superimposing the models on their respective templates. Align and Superimpose Proteins protocols that are implemented within Discovery Studio were used for structure superimpositions.

3.7 Molecular docking

3.7.1 Sh14

ZDOCK 3.0, a protein-protein docking suite that uses a rigid-docking algorithm, was used for molecular docking [143]. The docking predictions produced by ZDOCK were refined with RDOCK [144]. The molecular dynamics program CHARMM [145]

that is implemented in RDOCK is used for energy minimization in three-steps: removing clashes, optimizing polar charges and optimizing charged interactions. In the first two steps, RDOCK kept the ionic side-chains neutral. In the last step, it returned the side-chains into their full charge states. Finally, the results from RDOCK were ranked according to their re-calculated binding free energies.

3.7.2 *Gm*GSTU2-2, GSH and fluorodifen complex

USCF Chimera was used for molecular docking [146]. The program DockPrep was used to prepare the ligands (glutathione and fluorodifen) and the protein (*Gm*GSTU2-2 model) prior to docking calculations. The binding space (volume) was manually set and AutoDock as implemented in USCF Chimera was employed for rigid-body docking using a geometric matching algorithm to superimpose the ligand onto a negative image of the binding pocket [147,148]. ViewDock program was used to interactively analyze the receptor-ligands docking results and select the best pose of the ligands in the context of the binding site [149].

3.8 Molecular dynamics simulations and energy minimizations

3.8.1 Sh14

Molecular dynamic simulations (MDs) were performed using GROMACS 4.5.5 [150]. The forcefield CHARMM was used for energy minimizations of both native and mutant structures [145]. Generalized Born (GB) model was used during energy minimization steps to include thermodynamic aspects of aqueous solvation into the simulations and to speed up the energy minimization process [151]. MDs of the top results from the molecular docking experiments were performed in a single point charge (SPC) water-solvated, periodic environment.

3.9 Validation and quality of the 3D structures and models

3.9.1 *Gm*GSTU10-10–GSOH complex and Sh14–*Nb*-GSH complex

Stereochemical analyses and the quality of the final structures were evaluated with validation tools in Coot [136], MolProbity [152], and Procheck [153]. During refinement, 5% of the total number of reflections were set aside for cross-validation analysis using R_{free} value to monitor the refinement progress and structure improvement [154].

3.9.2 *Gm*GSTU2-2 and *Atu*GSTs

MODELLER PDF energies and DOPE scores were used to validate models [142,155]. Profiles-3D was used to verify and evaluate the homology models by comparing their structural environments with the preferred environments of amino acids [156].

3.10 Bioinformatics and structural analysis

Superimpositions of structures were performed with Secondary Structure Matching (SSM) that is implemented in PDBeFold server at EBI (http://www.ebi.ac.uk/msd-srv/ssm) [157,158]. The program Contact, which is implemented in CCP4, was used to measure distances [133]. The PDBePISA server (http://www.ebi.ac.uk/msd-srv/prot_int) was used for interface analysis [159]. Areas and volumes calculations of enzymes' active sites were carried out using CASTp [160,161]. ClustalW 2.0 [162] at EBI was used for multiple sequence alignments. The secondary structure assignments were performed with DSSP [163]. Figures were generated using CCP4mg [95], ESPript 3.0 [164], and LigPlot+ [165].

Table 3. Data processing and refinement statistics

Statistic	***Gm*GSTU10-10**	**Sh14**
Data Processing		
PDB ID	4CHS	5AGY
Space group	$P2_12_12_1$	$P2_12_12_1$
Unit cell dimensions (Å)	a = 47.7 b = 90.9 c = 112.9	a = 62.2 b = 77.6 c = 99.1
No. molecules	2	2
Resolution range (Å)	20.0–1.6 (1.7–1.6)[&]	19.79–1.75 (1.85–1.75)[&]
No. measured reflections	226,980	339,635
Unique reflections	65,115	48,837
Completeness (%)	99.5 (98.9)	99.3 (97.7)
Mosaicity (°)	0.5	0.375
$<I/\sigma(I)>$	17.8 (2.6)	40.1 (4.1)
R_{merge} (%)	4.5 (45.5)	5.2 (44.7)
R_{meas} (%)	5.3 (54.3)	5.7 (48.5)
Wilson B-factor (Å^2)	26.7	29.1
Refinement		
Reflections (working/test)	61,859/3,256	46,806/2,000
R_{cryst} (%)/ R_{free} (%)	16.2/19.4	16.8/18.7
No. protein atoms	3,561	3,685
No. water molecules	764	412
Bond lengths (Å)	0.012	0.009
Bond angles (°)	1.39	1.2
Residues in most favorable regions (%)	97.4	93.3
Residues in additionally allowed regions (%)	2.1	5.2
Average B-factors (Å^2)		
Protein	19.7	24.2
Waters	34.2	35.7
GSOH	17.3[*], 16.4[#]	-
Nb-GSH	-	28.2[*], 22.7[#]
4NM	-	27.5[**], 39.1[*], 51.2[#]
Phosphate	-	29.5
Acetate	-	42.5

[&]Numbers in parentheses correspond to the highest resolution shell (1.7 – 1.6 Å) and (1.85 – 1.75 Å) in *Gm*GSTU10-10 and Sh14, respectively. [*] in subunit A, [#] in subunit B and [**] on the surface of subunit A.

4. RESULTS AND DISCUSSION

4.1 Crystal structures of *Gm*GSTU10-10 and Sh14 (Papers II and III)

4.1.1 Structure determination

The crystal structures of *Gm*GSTU10-10 and Sh14 were determined using the MR method with the program Phaser [134]. Initial phases for *Gm*GSTU10-10 and Sh14 were obtained with a polyalanine search model based on a monomer of a *Glycine max* GST structure (PDB ID: 3FHS) and rice GST1 structure (PDB ID: 1OYJ), respectively. Both structures were found to belong to the orthorhombic $P2_12_12_1$ space group containing a dimer per asymmetric unit. The cell dimensions were a = 47.7 Å, b = 90.9 Å, c = 112.9 Å, for *Gm*GSTU10-10 and a = 62.2 Å, b = 77.6 Å, c = 99.1 Å for Sh14.

The final models for both *Gm*GSTU10-10 and Sh14 structures were obtained following iterative cycles of model building and inspection in Coot [136] and structure refinement in Phenix [135] with simulated annealing. At the initial stage, model building and refinement was performed with Arp/wARP and Refmac for Sh14 structure [139,140]. *Gm*GSTU10-10 was refined to R_{cryst} 16.2% (R_{free} 19.4%) at 1.6 Å, while Sh14 was refined to R_{cryst} 16.8% (R_{free} 18.7%) at 1.75 Å.

All main chain atoms of the residues in both subunits of *Gm*GSTU10-10 were visible with the exception of Met^1, Val^{217}, Ile^{218} and Glu^{219} of subunit A and Met^1, Thr^2, Asp^3 and Glu^{219} of subunit B. The listed residues were not visible in the $2|F_o| - |F_c|$ maps owing to lack of clear electron density map. In Sh14 structure, all main chain and side chain atoms for both subunits, except Met^1 and Gln^2 of subunit B, were visible in $2|F_o| - |F_c|$ maps.

The main chain and side chain stereochemistry of the refined final models for *Gm*GSTU10-10 and Sh14 are excellent and 99.5% and 98.5% of main chain φ/ψ angles lie within the allowed region of the Ramachandran plot, respectively [166]. Details of the quality of the structures are shown in Table 3. The refined atomic

coordinates and structure factors have been deposited in the Protein Data Bank under the accession numbers 4CHS and 5AGY for *Gm*GSTU10-10 and Sh14, respectively.

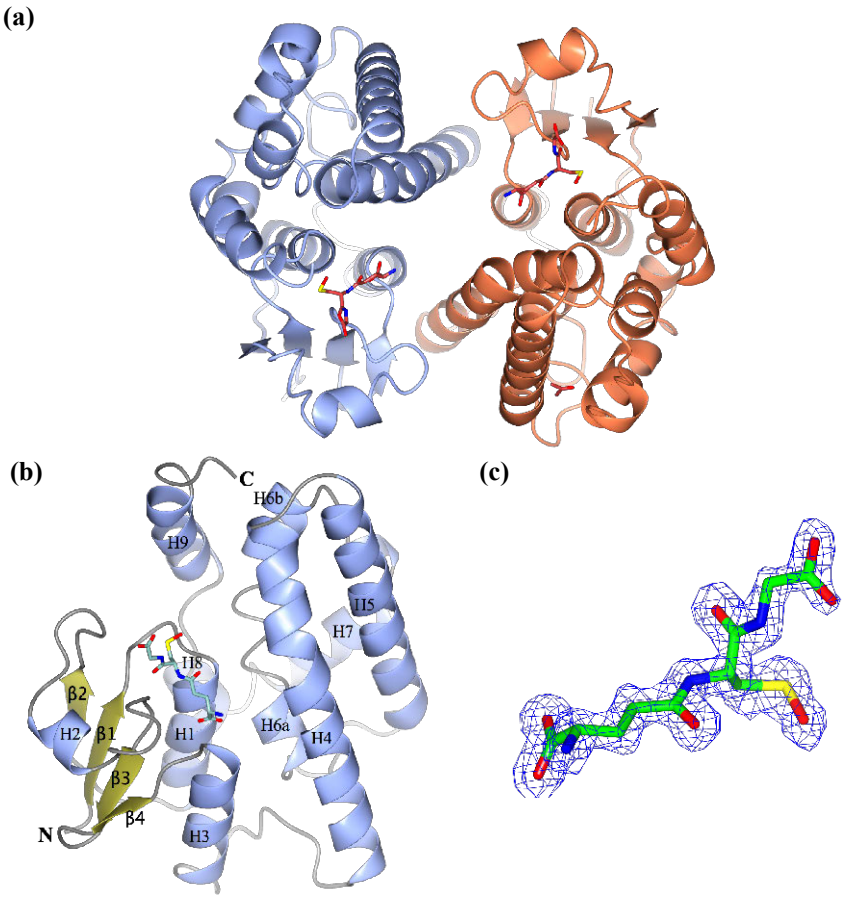

Fig. 11 (a) Ribbon diagram of the *Gm*GSTU10-10 dimer. Subunit A and B are colored ice blue and coral, respectively. GSOH and acetone are represented as sticks and are colored according to atom type. (**b)** Ribbon diagram of the *Gm*GSTU10-10 monomer. GSOH bound at the active site is represented as stick and is colored according to atom type. The helices are colored ice blue, strands in gold and turns in grey. **(c)** σ_A-weighted $2|F_o| - |F_c|$ electron density for the GSOH molecule bound to subunit B of *Gm*GSTU10-10 contoured at 1σ [167].

4.1.2 Structural features

4.1.2.1 Overall structure of *Gm*GSTU10-10–GSOH

The final refined model of *Gm*GSTU10-10 has two subunits in the asymmetric unit, forming a dimer as in the case of all cytosolic GSTs. The structure is globular in shape and displays the classical GST topology at the active sites (Fig. 11a). The two subunits of the homodimer are related by a two-fold symmetry axis. The highly

conserved and α/β-containing N-terminal domain hosts the GSH binding site (G-site) while the more variable and all α-helical C-terminal domain hosts the electrophilic binding site (H-site) (Fig. 11b). The dimer has an open configuration with a buried surface area of 1054 $Å^2$ between the subunits.

4.1.2.2 The G-site of *Gm*GSTU10-10

Despite the fact that spirodiclofen was used in the crystallization mixture, it was not present in the 3D structure of *Gm*GSTU10-10. Instead, only a clear density for a glutathione sulfenic acid molecule (GSOH) was observed at the G-site of each subunit of *Gm*GSTU10-10 (Fig. 11c). It is plausible that GSOH was created as an intermediate compound during crystallization after the oxidation of the –SH group of GSH during the catalytic reaction of GSH with hydroperoxides [36]. GSOH is found in the same orientation and conformation in both subunits.

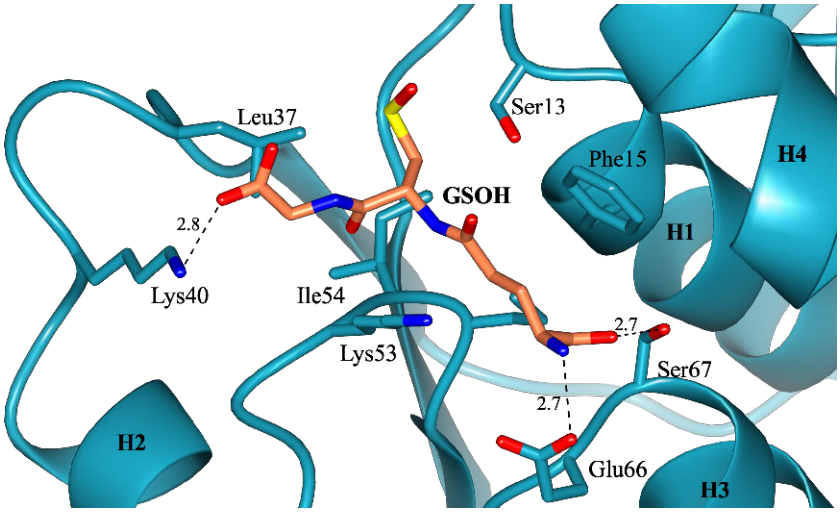

Fig. 12 Active site of *Gm*GSTU10-10. Hydrogen bonds are shown as dash lines with their corresponding distances (in Å). Interacting side-chains of the protein and GSOH are depicted as sticks and colored according to atom type. Cutoff distance was 3.5 Å.

The conserved catalytic Serine (Ser^{13}) stabilizes the Cys moiety of GSOH by forming a hydrogen bond with the sulfur of GSOH, hence, lowering the pKa for the thiol in the GST-glutathione complex [68,104,168]. The hydroxyl group of Ser^{13} is 4.6 Å away from the sulfur group of the GSOH. The oxygen and nitrogen of the cysteinyl moiety of GSOH forms a hydrogen bond with the main chain oxygen and nitrogen of Ile^{54} (Fig. 12). The γ-Glu moiety of GSOH faces the inter-domain cleft and is stabilized by hydrogen bonds with Glu^{66} and the hydroxyl group of Ser^{67}. Glu^{66} and

Ser[67] are conserved in plant tau class of GSTs. Helices H1 [Pro[14]–Lys[26]], H2 [Pro[42]–Met[47]], and H3 [Ser[67]–Val[77]], which partly make the N-terminal domains of *Gm*GSTU10-10 subunits, form a polar pocket that binds the glycine moiety of GSOH. In addition, the glycyl residue in GSOH forms hydrogen bonds with Lys[40].

4.1.2.3 Overall structure of Sh14

Sh14 has two subunits in the asymmetric unit and shares all the structural features of the native *Gm*GSTU4-4 (PDB ID: 2VO4) (Fig. 13). The dimer has an open configuration with a buried surface area of 1080 $Å^2$ between the subunits.

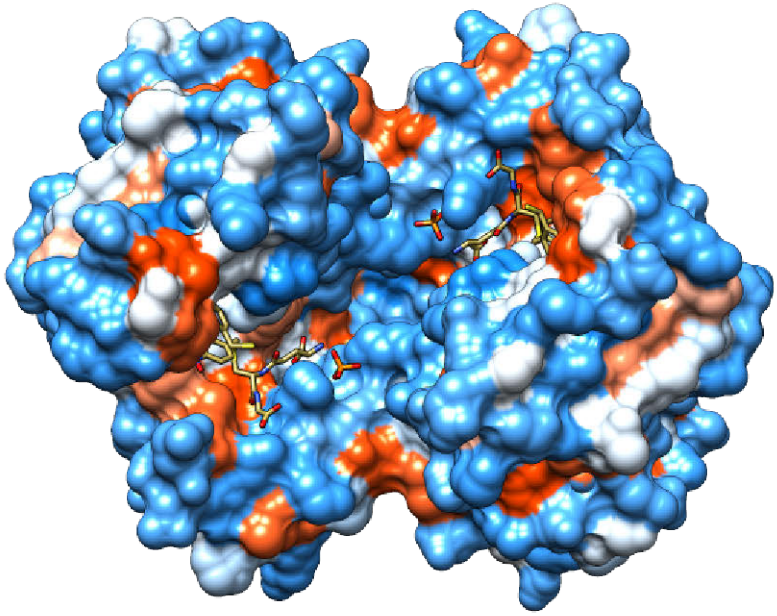

Fig. 13 Surface representation of Sh14 dimer. Residues are colored according to the hydrophobicity scale of Kyte and Doolittle [169]. *Nb*-GSH, 4NM, and PO_4^{-3} are represented as sticks and colored according to atom type. Figure was generated using UCSF Chimera [146].

Nevertheless, the overall structure of the N-terminal domain of Sh14 is very similar to the native structure (RMSD of 0.18 Å), significant differences were observed in helices H4, H5 and H9. Superimposition of Sh14 (subunit B) onto the native (subunit A) revealed an RMSD of 0.73 Å. Moreover, the replacement of Cys[114] with Trp[114] created more room at the active site to accommodate a substrate because the lid that blocked access to the site was eliminated.

4.1.2.4 The G-site of Sh14

Each subunit of Sh14 hosts a *Nb*-GSH molecule at the G-site. Additionally, two phosphate molecules (PO_4^{-3}) are present in close proximity to this site, in the solvent filled amphipathic cleft between the subunits. *Nb*-GSH was co-crystallized with Sh14, while phosphate was used to prepare the stock solution. The conformation of *Nb*-GSH and its interaction with Sh14 is very similar to the conformation and interactions exhibited in the native structure (*Gm*GSTU4-4–*Nb*-GSH; PDB ID: 2VO4) [68]. The rotation of the aromatic ring of 4-nitrobenzene by 12° is the only difference observed. The hydroxyl group of the catalytically active residue Ser^{13} is 3.1 Å away from the sulfur group of the *Nb*-GSH. The oxygen of the cysteinyl moiety of *Nb*-GSH forms a hydrogen bond with the main chain nitrogen of Ile^{54} (2.9 Å) (Fig. 14). The γ-Glu moiety of *Nb*-GSH points to the inter-domain cavity and forms hydrogen bonds with Glu^{66} (2.7 Å) and the hydroxyl group of Ser^{67} (2.7 Å). The glycyl residue in *Nb*-GSH forms hydrogen bonds with Lys^{40} (2.6 Å). Additionally, His^{51}, Pro^{55}, Leu^{68} and Asp^{103} further stabilize the binding of *Nb*-GSH through water-mediated hydrogen bonds.

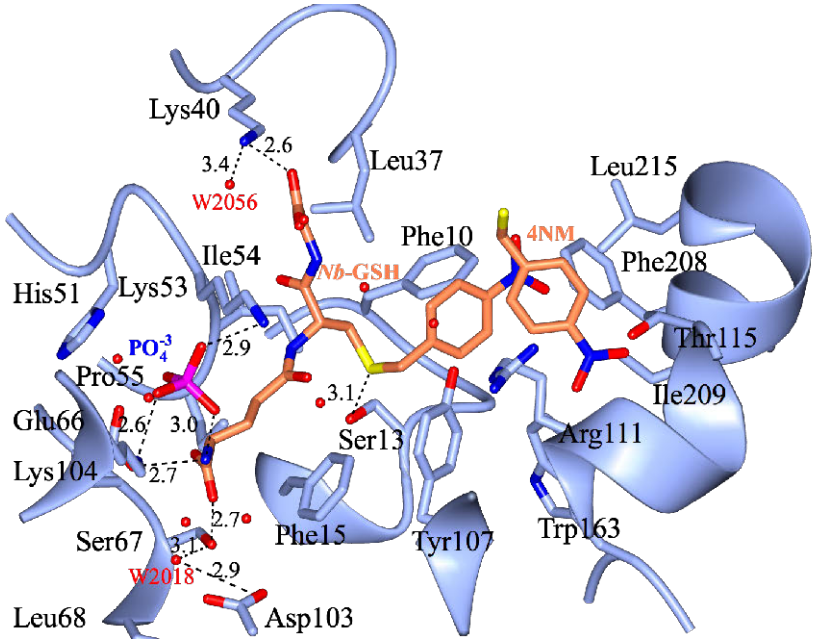

Fig. 14 Active site of Sh14. *Nb*-GSH, 4NM, and PO_4^{-3} are depicted as sticks and colored according to atom type. Red spheres represent water molecules. Hydrogen bonds are shown as dashed lines with their corresponding distances (in Å). Interacting side-chains of the protein are depicted as sticks and colored according to atom type. Cutoff distance was 3.5 Å.

4.1.2.5 The H-site of Sh14

The H-site is a deep cleft mainly formed by hydrophobic residues at the C-terminal domain. This site is responsible for ligand recognition and defines substrate specificities. It is observed that the residues that form this cleft are non-conserved; hence, the structure of the site varies widely across members of the tau class GSTs. Residues His^{107}, Lys^{111}, Trp^{114}, Trp^{163}, Phe^{208}, and Leu^{212} form the H-site for *Gm*GSTU10-10 whereas Tyr^{107}, Trp^{163}, and Leu^{212} for Sh14. Additionally, Phe^{10} that is located at the N-terminal domains on helix H1 is involved in forming the H-site in both structures. Interestingly, a molecule of (4-nitrophenyl)methanethiol (4NM) is present in each subunit of Sh14 – unlike the native enzyme (*Gm*GSTU4-4–*Nb*-GSH) [68]. 4NM resides in a hydrophobic pocket formed at the H-site by residues Tyr^{107}, Cys^{114}, Thr^{115}, and Trp^{163}.

4.2 Comparison of *Gm*GSTU10-10–GSOH, *Gm*GSTU4-4–GSH and *Gm*GSTU4-4–*Nb*-GSH

Structural comparison of *Gm*GSTU10-10 and *Gm*GSTU4-4 (PDB ID: 2VO4) revealed no major differences in tertiary structures (RMSD = 0.5 Å for main chain atoms) owing to the higher sequence identity (88%). However, a considerable structural variation was observed at the C-terminal part, most notably in helix H9 (Fig. 15). Lys^{215} in *Gm*GSTU10-10 is pointing away from the ligand-binding cavity, thus eliminating a lid over the active site pocket in *Gm*GSTU4-4–*Nb*-GSH. Consequently, a more open conformation is obtained for the active site in the absence of a ligand. The presence of a more accessible active site leads to different binding properties. Furthermore, a striking reduction in the size of *Gm*GSTU10-10's active site cavity as compared to *Gm*GSTU4-4–*Nb*-GSH is noted. Thus the higher affinity and compact binding of GSH in *Gm*GSTU10-10 is inline with the smaller binding cavity.

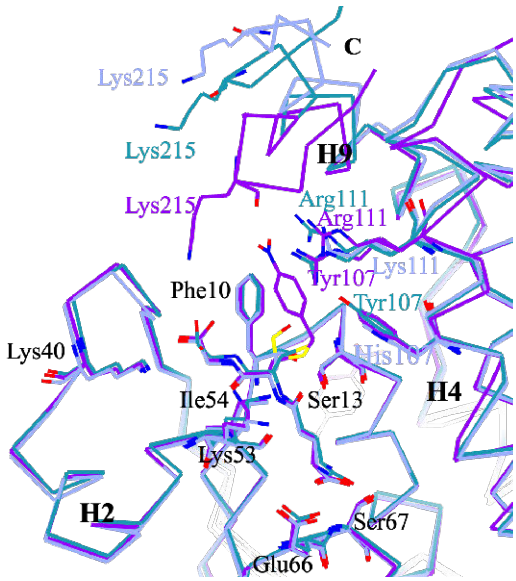

Fig. 15 Superimposition of *Gm*GSTU10-10 (ice blue), *Gm*GSTU4-4 (dark cyan), and *Gm*GSTU4-4–*Nb*-GSH (purple) at the active site. GSH, GSOH and *Nb*-GSH are shown as fat-bonds and colored according to atom type. Labels for Tyr^{107}, His^{107}, Arg^{111}, Lys^{111} and Lys^{215} are colored according to the respective structure.

Moreover, the replacement of Tyr^{107} by His^{107} at H4 of *Gm*GSTU10-10 resulted in eliminating a sharp kink that was observed in *Gm*GSTU4-4–*Nb*-GSH. Tyr^{107} strongly interacts with the nitrobenzyl moiety of *Nb*-GSH of *Gm*GSTU4-4–*Nb*-GSH; however, this interaction is absent in *Gm*GSTU10-10 due to His^{107} substitution. Evidently, these variations suggest a conformational variation at the H-site upon glutathione and substrate binding.

4.3 Comparison of native *Gm*GSTU4-4–GSH and Sh14–*Nb*-GSH

To identify the residues that are responsible for the altered kinetics of Sh14 site-directed mutagenesis was performed on each of the four altered amino acids – $Arg^{38}Gln$, $Gln^{46}Lys$, $Trp^{114}Cys$, and $Ile^{183}Val$. Apparently, the mutation of the conserved residue Trp^{114} to Cys^{114} was the only mutation of structural importance because it resulted in a striking allosteric effect on the active site cavity of Sh14. In contrast to the native enzyme (PDB ID: 4TOP), Sh14 displayed a different local helix-loop-helix fold that may be explained by the substitution of the bulky residue Trp^{114} with Cys^{114}. Moreover, there are two stabilizing salt bridges from Glu^{66} to

Lys^{104} of the second subunit in Sh14 in contrast to a single, weaker, salt bridge in the native. Furthermore, comparison of the active site cavities revealed a considerable size reduction (100 $Å^3$) in Sh14 as compared to the native. All the aforementioned structural variations, in addition to the restricted conformational flexibility of Lys^{104} as evidenced by the domain motion studies, suggest the existence of intersubunit communication as the reason for the differences in the catalytic efficiency and kinetic properties of the native and Sh14 enzymes.

4.4 Structural characterization of *Gm*GSTU2-2 and *Atu*GSTs (Papers I and IV)

4.4.1 Structural homology modeling

Structural modeling of *Atu*GSTs and *Gm*GSTU2-2 was performed in a number of steps. Firstly, all available sequence and structural information were collected from literature, protein sequences and structures databases. Next, the sequences of the related homologous proteins were used to create sequence alignments and the best and suitable structures were used to create the final 3D model of the proteins. Finally, the 3D models were validated with visual inspection and available validation softwares. Additionally, the information was cross-validated with the currently available information in literature.

Protein sequences were obtained from the genome database at NCBI (www.ncbi.nlm.nih.gov). Furthermore, a special database for *Glycine max* genes – SoyBase (www.soybase.org) – was used. For each protein sequence, all relevant available information was collected, including information on residues and biological functions of the proteins.

Templates for structure modeling of *Atu*GSTs were obtained from the Protein Data Bank (www.rcsb.org), which is a publicly available database that stores atomic coordinates of experimentally determined 3D structures of biological macromolecules. The sequence identity between the available templates used for modeling and *Atu*GST proteins was low (22%–27%); only *Atu*GST6, *Atu*GST7 and *Atu*GST8 have a sequence identity >30% with the templates used for modeling.

Meanwhile, crystal structures of *Gm*GSTU4-4 (PDB ID: 2VO4) and *Gm*GSTU10-10 (PDB ID: 4CHS) that were solved in our laboratory were used for modeling *Gm*GSTU2-2. Both templates have a sequence identity of more than 88% with *Gm*GSTU2-2.

Structural models of *Atu*GSTs and *Gm*GSTU2-2 were created with the program MODELLER [142] as implemented in Discovery Studio (www.accelrys.com). MODELLER generates a protein model based on a sequence alignment of the template, target protein sequences, and a template protein structure. The final protein models were further validated with Profile-3D protocol (within Discovery Studio) to verify their compatibility with the sequence of residues they contain [156]. It is noted that few residues at the N-terminal of *Atu*GST4 (Met^{1}–Thr^{20}) and *Atu*GST8 (Met^{1}–Tyr^{8}), and at the C-terminal of *Atu*GST1 (Leu^{204}–Gly^{208}) and *Atu*GST2 (Val^{200}–Ala^{203}) were not modeled because of lack of suitable templates.

4.4.2 Structural characterization of *A. tumefaciens* C58 GSTs

*Atu*GSTs have not been previously structurally characterized and their biological functions are not clearly known. Models of *Atu*GSTs were constructed to characterize and provide structural information at the atomic level since there were no experimentally determined 3D structures of *Atu*GST enzymes at the time of the study. However, currently the crystal of *Atu*GST4 is available and it was determined in our laboratory at a high resolution (1.4 Å) [170].

Identification of GST-like protein sequences in *Agrobacterium* genomic sequence was performed with a BLAST search and eight sequences (*Atu*GST1–*Atu*GST8) were found. These *Atu*GST genes have low sequence identity (18–28%) between them. Further analysis of the *Atu*GST proteins revealed that they share a considerable sequence identity (60–68%) with proteins from soil bacteria. Although genetic relationship between GSTs is complex, phylogenetic analysis showed that *Atu*GSTs belong to different classes. *Atu*GST1, *Atu*GST6, *Atu*GST7 and *Atu*GST8 belong to beta class while *Atu*GST2 to theta class. On the other hand, *Atu*GST3, *Atu*GST4 and *Atu*GST5 do not belong to any currently characterized class, suggesting that they may belong to new classes of GSTs.

4.4.2.1 Insights into the structural models of *Atu*GSTs

All models of *Atu*GST proteins displayed a typical GST fold and domain organization. Nevertheless, models of *Atu*GST2–*Atu*GST8 were superimposed onto *Atu*GST1 and significant structural variations were observed, particularly at the second half of helix H5 (H5b) in the C-terminal domain. This variation among *Atu*GSTs models may suggest that the *Atu*GSTs have different modes of substrate recognition and specificities. However, the G-site is conserved among all *Atu*GSTs models and exhibits the characteristic positive electrostatic field [68,171].

Furthermore, comparative sequence and structural analysis of the active site were done and potential catalytic residues were proposed: Ser for *Atu*GST1, *Atu*GST2, *Atu*GST3 and *Atu*GST5; Cys for *Atu*GST6; Ser/Cys for *Atu*GST7; Arg/Glu for *Atu*GST4 and *Atu*GST8. The absence of a characteristic cytosolic GST catalytic residues such as Cys, Ser and Tyr in *Atu*GST4 and *Atu*GST8 makes them special. Indeed, the recently experimentally determined 3D structure of *Atu*GST4 confirms our proposal as Arg^{34} was found to play a significant role in catalysis following mutagenesis studies [170]. It was therefore suggested that *Atu*GST4 is the first member of a new GST class, coined eta. Moreover, a hydrophobic cavity overlapping the G-site and H-site that can accommodate a larger molecule (e.g. antibiotic molecule) is present in *Atu*GST1–*Atu*GST6.

Additionally, the absence of the SNAIL/TRAIL motif at the N-terminal domain is observed. This motif is conserved in most GST classes and provides polar functional groups to the G-site [172]. Nonetheless, putative SNAIL/TRAIL-like motifs with few modifications were present in *Atu*GST2, *Atu*GST3, *Atu*GST4, *Atu*GST7 and *Atu*GST8 (residues 69–73 according to *Atu*GST1 numbering) (Fig. 16).

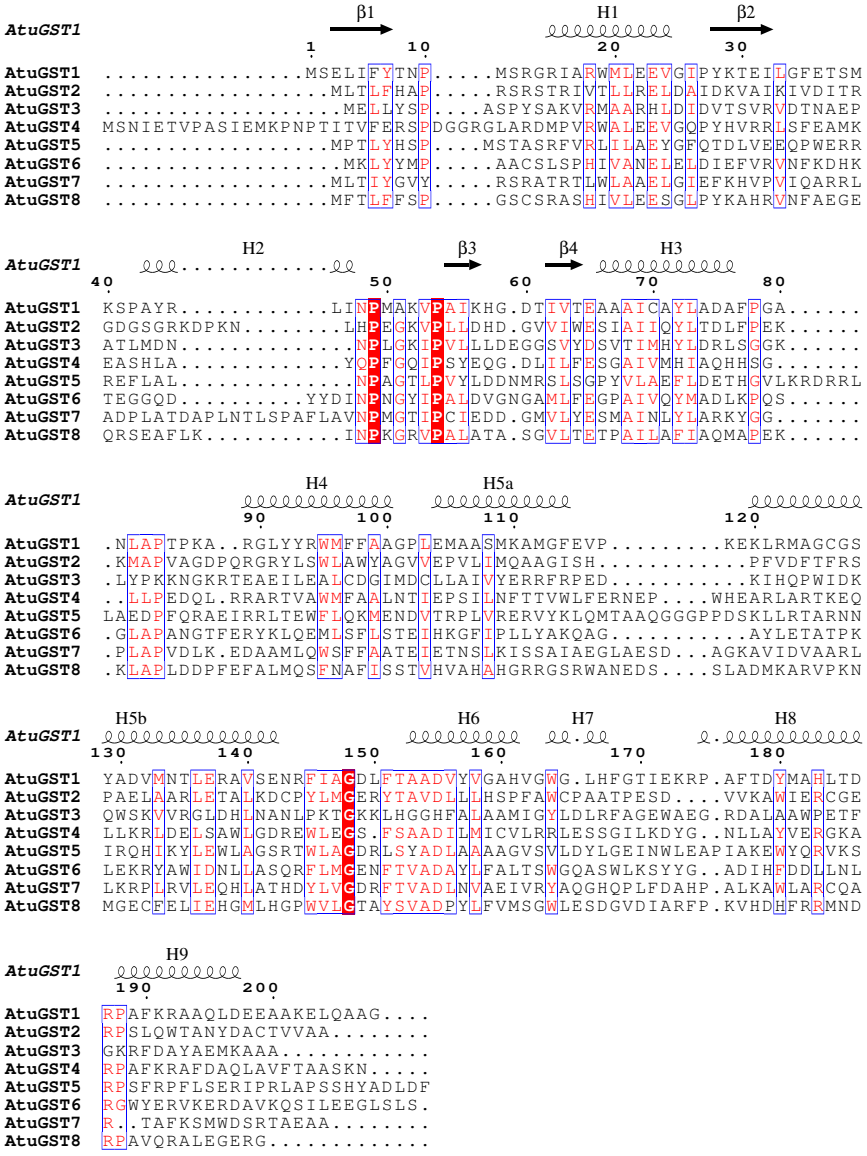

Fig. 16 Structure-based multiple sequence alignment of *Atu*GST proteins. Conserved residues are shown in red shaded rectangles and residues that share physicochemical properties more than 70% are framed in blue rectangles. Secondary structure elements of *Atu*GST1 are displayed on top of the alignment. The multiple sequence alignment was done using ClustalW [173]. The figure was produced using ESPript [164].

The presence of Pro^{54} preceded by Ile/Val in all *Atu*GSTs suggests the existence of a hydrogen bond between Val^{53} and GSH. It is proposed that main chain atoms of O and N of Val^{53} will interact with N2 and O2 of GSH, respectively. Furthermore, as in all GST family members a hydrogen bond between Glu^{65} and N1 of GSH will exist

owing to the fact that Glu^{65} is a conserved residue. It is noteworthy that Glu^{65} is replaced by Asp in *Atu*GST3 and a Gly inserted is preceded by Ser in *Atu*GST5, suggesting altered binding characteristics at this site.

Although a 'lock-and-key' motif is conserved and very crucial in forming a stable and biologically functional dimer in several classes of GSTs, *Atu*GSTs achieve this through different modes of interactions [98,174]. A closer analysis of *Atu*GSTs revealed the presence of a 'lock-and-key' motif in *Atu*GST1, *Atu*GST3, *Atu*GST4 and *Atu*GST8. In contrast, mainly polar interactions are present at the dimer interface of *Atu*GST2, *Atu*GST5, *Atu*GST6 and *Atu*GST7.

Moreover, structural analysis showed the presence of a strictly conserved N-capping box motif in all *Atu*GSTs (except *Atu*GST3). This motif, Ser^{152}/Thr^{152}–Xaa–Xaa–Asp^{155}, is known to be responsible for the formation of α_6-helix during GST folding [175,176]. Interestingly, a sequence His^{152}–Gly^{153}–Gly^{154}–His^{155} is present in *Atu*GST3 though its structural or functional role is unknown.

4.4.3 Structural features of *Gm*GSTU2-2

The structural model for *Gm*GSTU2-2 was constructed because no experimentally determined 3D structure is yet available and crystallizations failed to produce suitable crystals for structural analysis. Investigation of the model for structural variations and similarities with previously determined tau class GSTs from *Glycine max* provided a better understanding of its distinct function.

*Gm*GSTU2-2 model was constructed using two phylogenetically close isoenzymes, namely *Gm*GSTU4-4 (PDB ID: 2VO4) and *Gm*GSTU10-10 (PDB ID: 4CHS) that share sequence identity of 91% and 88% with *Gm*GSTU2-2, respectively. As expected, existence of high sequence identity with known GSTs as well as structural superimpositions against previously solved structures of GSTs showed that *Gm*GSTU2-2 shares the overall fold of GST enzymes and exhibits the typical two-domain architecture of GSTs. Comparatively, the active site is well conserved and dimerization occurs in a similar fashion as in other tau class GSTs (e.g. *Gm*GSTU10-10). GSH binds to *Gm*GSTU2-2 in a similar conformation and orientation as compared to both *Gm*GSTU4-4 and *Gm*GSTU10-10.

Nevertheless, a crucial difference was identified at the C-terminal domain. Residues 208–219 at H9 of *Gm*GSTU2-2 and *Gm*GSTU10-10 form a more open conformation as compared to *Gm*GSTU4-4. Particularly, Lys^{215} of *Gm*GSTU10-10 and the corresponding Lys^{212} of *Gm*GSTU2-2 are observed to point away from the active site thus making the active site available for binding unlike Lys^{215} of *Gm*GSTU4-4 that forms a lid and makes the active site inaccessible (Fig. 17). Interestingly, residues 202–211 at H9 of *Gm*GSTU4-4–*Nb*-GSH also create a lid by folding back over the top of the N-terminal domain thereby blocking access to the active site. Moreover, volume/area calculations at the active site revealed a comparable active site cavity for *Gm*GSTU2-2 and *Gm*GSTU10-10, whereas a considerably larger cavity for *Gm*GSTU4-4 was identified. All these structural variations will plausibly affect binding and xenobiotic recognition and binding properties of *Gm*GSTU2-2.

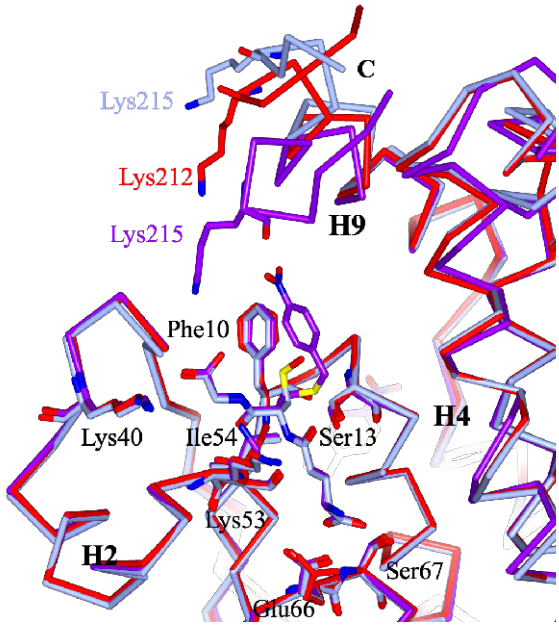

Fig. 17 Structural superimposition of the side chains of *Gm*GSTU10-10 (ice blue), *Gm*GSTU2-2 (red), and *Gm*GSTU4-4–*Nb*-GSH (purple) at the active site. GSOH and *Nb*-GSH are shown as sticks and colored according to atom type. Lys^{215} of *Gm*GSTU4-4–*Nb*-GSH and *Gm*GSTU10-10 and Lys^{212} of *Gm*GSTU2-2 are colored according to the respective structure.

Additionally, *Gm*GSTU2-2 may have a ligandin function besides the classic catalytic function as other tau class GSTs. Although the exact location of ligand-binding site varies, it is proposed that residues Trp^{11}, Arg^{20}, Tyr^{30}, and Tyr^{32} form a hydrophobic

pocket on the surface of the protein [68,90]. Tyr^{32} is replaced by Ser^{32} in *Gm*GSTU2-2, hence it can be suggested that *Gm*GSTU2-2 could perform a ligandin function with a different ligand recognition mode owing to the size reduction by Ser^{32}.

4.4.3.1 Fluorodifen docking

Following docking calculations, fluorodifen appears to bind in a hydrophobic cavity of *Gm*GSTU2-2 formed by Phe^{10}, Pro^{12}, Phe^{15}, Leu^{37}, Phe^{107}, Trp^{114}, Trp^{163}, Phe^{208}, Ile^{212}, and Phe^{216} (Fig. 18). This optimal mode of ligand binding is stabilized by hydrogen bond interactions between the side chains of Arg^{111} and Lys^{215} with the $–NO_2$ groups of the ligand.

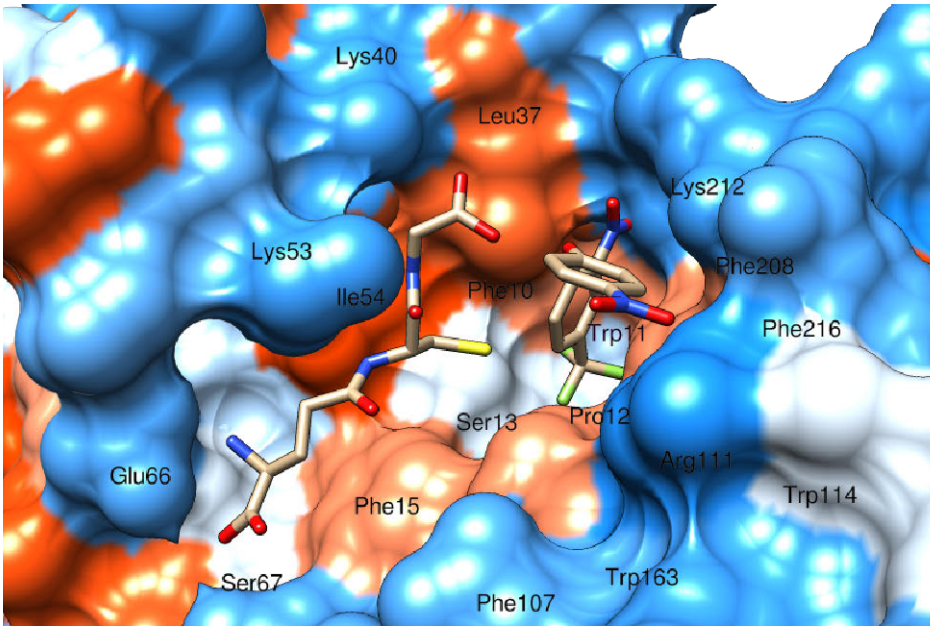

Fig. 18 Surface representation of the predicted mode of interaction between fluorodifen and *Gm*GSTU2-2. Fluorodifen and GSH are shown as sticks and colored according to atom type. Residues of *Gm*GSTU2-2 that are predicted to interact with fluorodifen and GSH are labeled. Residues are colored according to the hydrophobicity scale of Kyte and Doolittle [169]. Figure was generated using UCSF Chimera [146].

5. CONCLUSIONS AND FUTURE PERSPECTIVES

The primary objective of this study was to provide structural and functional insights into cytosolic GST enzymes from *Glycine max* and *Agrobacterium tumefaciens* C58. X-ray crystal structures of *Gm*GSTU10-10 and Sh14 complexed with GSOH and *Nb*-GSH, respectively, were determined. Moreover, structural models of *Gm*GSTU2-2 and *Atu*GSTs were constructed. Ligand recognition and binding were studied by experimental and *in silico* techniques that included X-ray crystallography, bioinformatics, molecular modeling and docking as well as molecular dynamics simulation. The analyses were combined with kinetic experiments and specificities data.

The main conclusions of the study are:

1. *Gm*GSTU10-10–GSOH complex

- The crystal structure of *Gm*GSTU10-10 in complex with GSOH was solved at 1.6 Å resolution. *Gm*GSTU10-10 was shown to share the same overall and domain organization with previously solved cytosolic tau class GSTs. It consists of two subunits, each folding into two distinct domains: the N- and C-terminal domains. The N-terminal domain (residues 1–77) exhibits a thioredoxin-like fold, whereas, the C-terminal domain (89–219) is entirely made up of α-helices positioned downstream the thioredoxin-like domain structure and connected to the N-terminal domain by a short linker sequence.

- The conformation of GSOH bound to *Gm*GSTU10-10 was resolved. The Cys and cysteinyl moieties of GSOH are stabilized by hydrogen bonds with Ser^{13} and Ile^{54}, respectively. Residues at the N-terminal domain Pro^{14}–Lys^{26}, Pro^{42}–Met^{47} and Ser^{67}–Val^{77} from helices H1, H2, and H3, respectively, form a polar cavity that binds the glycine moiety of GSOH. The γ-Glu moiety of GSOH face the inter-domain cleft and is stabilized by hydrogen bonds with Glu^{66} and Ser^{67}. Lys^{40} further stabilizes the glycyl residue in GSOH through hydrogen bonds.

- Absence of a sharp kink in *Gm*GSTU10-10 structure at the upper part of helix H4 formed by residues $D^{103}KKIY^{107}$ in *Gm*GSTU4-4 structures coupled with the replacement of Tyr^{107} with His^{107} is suggested to alter substrate recognition and catalytic mechanism.

- A more open conformation at the C-terminal, particularly at H9, that makes access to the active site pocket easy was revealed.

- A comparatively smaller active site pocket that can tightly bind GSH with high affinity was identified that can be explained by the aforementioned structural variation at helices H4 and H9.

2. Sh14

- The structure of Sh14, the first GSTU enzyme produced through directed evolution, complexed with *S*-(*p*-nitrobenzyl)-glutathione was solved at 1.75 Å resolution.

- $Trp^{114}Cys$ point mutation significantly rearranged the local helix–loop–helix fold as compared to the native *Gm*GSTU4-4 enzyme owing to the replacement of a bulky Trp^{114} with a smaller Cys^{114}. This mutation allowed the helix–loop–helix motif to move towards the inner side of the structure, consequently enhancing the catalytic properties of Sh14.

- Two strong salt bridges between Glu^{66} and Lys^{104} that stabilize the dimer and trigger an allosteric effect between the subunits were identified at the subunit interface.

- Lys^{104} at the H4 exhibits a more restricted conformational flexibility as compared to the native enzyme and this stability allows communication between the H-sites of the subunits.

3. *Atu*GSTs

- *Atu*GSTs share the same overall fold and domain organization with eukaryotic cytosolic GSTs. However, sequence and phylogenetic analysis showed that the *Atu*GSTs the belong to different classes: *Atu*GST1, *Atu*GST6, *Atu*GST7 and

*Atu*GST8 to beta; *Atu*GST2 to theta; whereas, *Atu*GST3, *Atu*GST4 and *Atu*GST5 belong to new classes of GSTs as they do not belong to any currently characterized classes.

- Interactions at the G-site appear conserved. Nonetheless, considerable variations between *Atu*GST structures at their respective active site, C-terminal domain, and the linker connecting the N- and C-terminal domains were revealed. Particularly, the H5b region that forms the H-site displays significant variation among the *Atu*GST structures. This region recognizes the electrophilic moiety of the xenobiotic, thus providing different specificities.

- Functional *Atu*GST dimers are formed through diverse modes. Hydrophobic 'lock–and–key'-like motifs form *Atu*GST1, *Atu*GST3, *Atu*GST4 and *Atu*GST8 dimers. However, polar interactions dominate the dimer interface of *Atu*GST2, *Atu*GST5, *Atu*GST6 and *Atu*GST7.

- Although the SNAIL/TRAIL motif at the N-terminal domain is conserved in other GSTs, it is absent in all *Atu*GSTs. Nevertheless, slightly modified putative SNAIL/TRAIL-like motifs were present in *Atu*GST2, *Atu*GST3, *Atu*GST4, *Atu*GST7 and *Atu*GST8.

- The H-site of all *Atu*GSTs is large enough to accommodate bulky ligands. Strikingly, the hydrophobic cavity at the active site of *Atu*GST1, *Atu*GST2, *Atu*GST3, *Atu*GST4, *Atu*GST5, and *Atu*GST6 can host a molecule as large as an antibiotic.

- Based on in-depth sequence and structural analysis of the active site candidates for the catalytic amino acids were proposed: Ser for *Atu*GST1, *Atu*GST2, *Atu*GST3 and *Atu*GST5; Cys for *Atu*GST6; Ser/Cys for *Atu*GST7; Arg/Glu for *Atu*GST4 and *Atu*GST8.

4. *Gm*GSTU2-2

- *Gm*GSTU2-2 shares same structural features with other tau class GSTs and forms a dimer in a similar way as suggested by the presence of dimer forming residues at its dimer interface; namely, Arg^{92}, Arg^{96}, Glu^{76}, and Gln^{72}.

- Absence of a lid block at the entrance of the active site makes the active site cavity of *Gm*GSTU2-2 more accessible as compared to *Gm*GSTU4-4.

- *Gm*GSTU2-2 has a smaller active cavity compared to *Gm*GSTU4-4 because of the differences at the C-terminal domain, particularly residues 202–211 at H9.

- The binding of fluorodifen at the active cavity of *Gm*GSTU2-2 was resolved with molecular docking and essential amino acid residues that could form the binding cavity were identified: Phe^{10}, Pro^{12}, Phe^{15}, Leu^{37}, Phe^{107}, Trp^{114}, Trp^{163}, Phe^{208}, Ile^{212}, and Phe^{216}. Moreover, side chains of Arg^{111} and Lys^{215} could provide stability by forming hydrogen bonds with the $–NO_2$ groups of fluorodifen.

- Ser^{32} in *Gm*GSTU2-2 replaces Tyr^{32}, an important residue that forms the L-site in other GSTs, and this may alter the enzyme's xenobiotic recognition and binding properties.

The structural, catalytic and evolutionary aspects of *Agrobacterium tumefaciens* C58 and *Glycine max* tau class GSTs were studied in this work. However, there is still more to be answered in future studies since our understanding of the role of GSTs in metabolism with respect to the degradation of toxic pollutants is not fully understood.

Recent meteoric advances in biotechniques and the availability of a vast amount of biochemical knowledge have resulted in potentially several applications for enzymes in nanotechnology, biotechnology, and biomedicine. However, natural enzymes often need optimization and enhancement for real life applications to meet the specific objectives. It is possible to utilize the power of enzymes with protein engineering techniques to further potential applications by shedding light on the complexity that characterizes the enzymes.

From protein redesigning/engineering point of view, GSTs are highly promising candidate enzymes since they display promiscuous functions that can potentially be enhanced/altered for novel functions. It is known that polymorphisms and gene deletions alter substrate selectivity hence susceptibility towards a wide range of toxic

chemicals including pollutants and pharmaceuticals. Hence, the study and understanding of GST enzymatic functions and their structures could pave the way for protein engineering experiments so that GSTs can be employed in biotechnological and biomedical applications.

Bioremediation is a crucial treatment method that is capable of cleaning the environment by degrading toxic pollutants into less toxic or non-toxic substances. Because bioremediation uses biological organisms, including bacteria, it is clearly inexpensive and less labor intensive as compared to traditional remediation technologies like thermal desorption, stabilization, soil vapor extraction, incineration and vitrification. In light of recent scientific developments, it is possible to engineer bacteria to produce specific enzymes that can be used as biodegrading agents, hence removing toxic organic pollutants from the environment. Since GSTs play a significant role in detoxification of toxic pollutants, they are undoubtedly amongst the top candidates to be genetically engineered and used for bioremediation.

ACKNOWLEDGEMENTS

This work was carried out in the Protein Crystallography Group at Turku Centre for Biotechnology (Biocity). Financial support from Finnish Cultural Foundation, University of Turku, and Turku Centre for Systems Biology are acknowledged.

My deepest gratitude to my supervisor Adj. Prof. Tassos Papageorgiou for introducing me to the world of structural biology, especially X-ray protein crystallography, and for excellent guidance throughout my studies and giving me the opportunity to work in a magnificent laboratory and research facilities. Additionally, he is thanked for never-ending positive attitude, for patience, and for helping me to overcome many crisis situations throughout my studies and finish this dissertation.

Besides my advisor, I would like to thank my research director, Prof. Jyrki Heino for his encouragement, Prof. Tiina Salminen and Dr. David Reverter for critically pre-examining the thesis and for their insightful constructive comments.

I would like to acknowledge all my co-authors from the University of Athens for the excellent collaboration and the passion to reach the same goal from a different point of view.

Sincere gratitude to Natalia Pakharukova, Sumera Perveen and Yonatan Mideksa for patiently proofreading the manuscript and your true friendship.

Warm thanks to Sirkku Grönroos for helping me with small and big administrative issues of all sorts. Mikael Wasberg, Mårten Hedman, Petri Vahakoski, Juha Strandén, and Pasi Viljakainen are thanked for keeping the laboratory working and making it possible for us to concentrate on science.

Many thanks to all past and present labmates. It has been a pleasure to get to know all these great persons from different cultures and backgrounds: Pratusha Dhavala, Teemu Hakarainen, Bishwa Subedi, Carlos Prieto, David Hernandez, Kirti Sonkar,

Maheshkumar Kharat and all the others, you have made the working environment extremely pleasant with your stimulating discussions and all the fun we had outside the lab.

Last but not least, I want to thank my loving and caring parents, Eng. Worku Mulleta and Tsehay Mulugeta, for teaching me to be curious and sincere to life, and above all, for the unflagging and unconditional support throughout my life and studies; my siblings Jote, Lensa, Beri, Diriba and Iftu for their love and encouragement; my big sister Yeshimebet Tigabu for your tremendous help on the long and sometime bumpy road of research; my dearest wife Hermi, for your immeasurable patience, for always supporting me in my academic pursuits and for the wonderful life that we share together. Words cannot express my appreciation and love for my daughter Feenet, for being cute, making me laugh and being the great joy that you are.

Turku, August 2016
Abdi Worku Muleta

REFERENCES

[1] **Oakley AJ.** Glutathione transferases: new functions. *Curr Opin Struct Biol.* 2005; 15: 716-23.

[2] **Oakley AJ.** Glutathione transferases: a structural perspective. *Drug Metab Rev.* 2011; 43: 138-51.

[3] **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: 1-16.

[4] **Wilce MC, Parker MW.** Structure and function of glutathione S-transferases. *Biochim Biophys Acta.* 1994; 1205: 1-18.

[5] **Atkinson HJ, Babbitt PC.** Glutathione transferases are structural and functional outliers in the thioredoxin fold. *Biochemistry.* 2009; 48: 11108-16.

[6] **Hvidsten TR, Laegreid A, Kryshtafovych A, *et al.*** A comprehensive analysis of the structure-function relationship in proteins based on local structure similarity. *PLoS One.* 2009; 4: e6266.

[7] **Zhang C, Kim SH.** Overview of structural genomics: from structure to function. *Curr Opin Chem Biol.* 2003; 7: 28-32.

[8] **Nagaraj SH, Deshpande N, Gasser RB, Ranganathan S.** ESTExplorer: an expressed sequence tag (EST) assembly and annotation platform. *Nucleic Acids Res.* 2007; 35: W143-7.

[9] **Chandonia JM, Brenner SE.** The impact of structural genomics: expectations and outcomes. *Science.* 2006; 311: 347-51.

[10] **Bourgeois D, Royant A.** Advances in kinetic protein crystallography. *Curr Opin Struct Biol.* 2005; 15: 538-47.

[11] **Sies H.** Glutathione and its role in cellular functions. *Free Radic Biol Med.* 1999; 27: 916-21.

[12] **Liska DJ.** The detoxification enzyme systems. *Altern Med Rev.* 1998; 3: 187-98.

[13] **Ioannides C.** Enzyme systems that metabolise drugs and other xenobiotics. New Jersey, USA: John Wiley and Sons Ltd.; 2001.

[14] **Keller W.** On the conversion of benzoic acid into hippuric acid. *Chem Pharm.* 1842; 43: 108.

[15] **Meister A.** On the discovery of glutathione. *Trends Biochem Sci.* 1988; 13: 185-8.

[16] **Bode M, Haas M, Faymonville T, *et al.*** Biotransformation of metamitron by human P450 expressed in transgenic tobacco cell cultures. *J Environ Sci Health B.* 2006; 41: 201-22.

[17] **Schmidt B, Faymonville T, Gembe E*, et al*.** Comparison of the biotransformation of the 14C-labelled insecticide carbaryl by non-transformed and human CYP1A1-, CYP1A2-, and CYP3A4-transgenic cell cultures of *Nicotiana tabacum*. *Chem Biodivers.* 2006; 3: 878-96.

[18] **Schmidt B, Joussen N, Bode M, Schuphan I.** Oxidative metabolic profiling of xenobiotics by human P450s expressed in tobacco cell suspension cultures. *Biochem Soc Trans.* 2006; 34: 1241-5.

[19] **Anzenbacher P, Anzenbacherova E.** Cytochromes P450 and metabolism of xenobiotics. *Cell Mol Life Sci.* 2001; 58: 737-47.

[20] **Guengerich FP.** Uncommon P450-catalyzed reactions. *Curr Drug Metab.* 2001; 2: 93-115.

[21] **Ji L, Faponle AS, Quesne MG*, et al*.** Drug metabolism by cytochrome P450 enzymes: what distinguishes the pathways leading to substrate hydroxylation over desaturation? *Chemistry.* 2015; 21: 9083-92.

[22] **Jakoby WB, Ziegler DM.** The enzymes of detoxication. *J Biol Chem.* 1990; 265: 20715-8.

[23] **King CD, Rios GR, Green MD, Tephly TR.** UDP-glucuronosyltransferases. *Curr Drug Metab.* 2000; 1: 143-61.

[24] **Negishi M, Pedersen LG, Petrotchenko E*, et al*.** Structure and function of sulfotransferases. *Arch Biochem Biophys.* 2001; 390: 149-57.

[25] **Homolya L, Varadi A, Sarkadi B.** Multidrug resistance-associated proteins: Export pumps for conjugates with glutathione, glucuronate or sulfate. *Biofactors.* 2003; 17: 103-14.

[26] **Parvez H, Reiss C.** Molecular responses to xenobiotics. Paris: Elsevier; 2001.

[27] **Hatzios KK.** Regulation of enzymatic systems detoxifying xenobiotics in plants, a brief overview and directions for future research. Netherlands: Springer; 1997.

[28] **Williams GJ, Thorson JS.** Natural product glycosyltransferases: properties and applications. *Adv Enzymol Relat Areas Mol Biol.* 2009; 76: 55-119.

[29] **Leah JM, Worrall TL, Cobb AH.** Isolation and characterization of two glucosyltransferases from *Glycine max* associated with bentazone metabolism. *Pestic Sci.* 1992; 34: 81-7.

[30] **Lamoureux GL, Shimabukuro RH, Frear DS.** Glutathione and glucoside conjugation in herbicide selectivity. In: Caseley JC, Cussans GW, Atkin RK, editors. Herbicide Resistance in Weeds and Crops. Oxford, UK: Butterworth-Heinemann; 1991. p. 227-261.

[31] **Andrae WA, Good NE.** Studies on 3-indoleacetic acid metabolism. IV. Conjugation with aspartic acid and ammonia as processes in the metabolism of carboxylic acids. *Plant Physiol.* 1957; 32: 566-72.

[32] **Peterson LA.** Reactive metabolites in the biotransformation of molecules containing a furan ring. *Chem Res Toxicol.* 2013; 26: 6-25.

[33] **Chkanikov DI, Makeyev AM, Pavlova NN*, et al.*** Variety of 2,4-D metabolic pathways in plants; its significance in developing analytical methods for herbicides residues. *Arch Environ Contam Toxicol.* 1976; 5: 97-103.

[34] **Monks TJ, Anders MW, Dekant W*, et al.*** Glutathione conjugate mediated toxicities. *Toxicol Appl Pharmacol.* 1990; 106: 1-19.

[35] **Brigelius-Flohé R, Maiorino M.** Glutathione peroxidases. *Biochimica Biophysica Acta.* 2013; 1830: 3289-303.

[36] **Prohaska JR.** The glutathione peroxidase activity of glutathione S-transferases. *Biochim Biophys Acta.* 1980; 611: 87-98.

[37] **Hayes JD, Flanagan JU, Jowsey IR.** Glutathione transferases. *Annu Rev Pharmacol Toxicol.* 2005; 45: 51-88.

[38] **Sun Y, Li H, Huang JR.** Arabidopsis TT19 functions as a carrier to transport anthocyanin from the cytosol to tonoplasts. *Mol Plant.* 2012; 5: 387-400.

[39] **Litwack G, Ketterer B, Arias IM.** Ligandin: a hepatic protein which binds steroids, bilirubin, carcinogens and a number of exogenous organic anions. *Nature.* 1971; 234: 466-7.

[40] **Kanaoka Y, Ago H, Inagaki E*, et al.*** Cloning and crystal structure of hematopoietic prostaglandin D synthase. *Cell.* 1997; 90: 1085-95.

[41] **Eklund BI, Gunnarsdottir S, Elfarra AA, Mannervik B.** Human glutathione transferases catalyzing the bioactivation of anticancer thiopurine prodrugs. *Biochem Pharmacol.* 2007; 73: 1829-41.

[42] **Desmots F, Loyer P, Rissel M*, et al.*** Activation of C-Jun N-terminal kinase is required for glutathione transferase A4 induction during oxidative stress, not during cell proliferation, in mouse hepatocytes. *FEBS Lett.* 2005; 579: 5691-6.

[43] **Board PG, Menon D.** Glutathione transferases, regulators of cellular metabolism and physiology. *Biochim Biophys Acta.* 2013; 1830: 3267-88.

[44] **Abhilash PC, Jamil S, Singh N.** Transgenic plants for enhanced biodegradation and phytoremediation of organic xenobiotics. *Biotechnol Adv.* 2009; 27: 474-88.

[45] **Mannervik B, Danielson UH.** Glutathione transferases--structure and catalytic activity. *CRC Crit Rev Biochem.* 1988; 23: 283-337.

[46] **Jakobsson PJ, Morgenstern R, Mancini J*, et al.*** Membrane-associated proteins in eicosanoid and glutathione metabolism (MAPEG). A widespread protein superfamily. *Am J Respir Crit Care Med.* 2000; 161: 20-4.

[47] **Morle FP, Rauch C, Petit E*, et al.*** Gene and protein characterization of the human glutathione S-transferase kappa and evidence for a peroxisomal localization. *J Biol Chem.* 2004; 279: 16246-53.

[48] **Zimniak P, Singh SP.** Families of glutathione transferases. In: Taylor Awasthi YC, editor. Toxicology of glutathione transferases. Boca Raton, Fl,USA: CRC Press; 2006. p. 11-26.

[49] **Raza H, Robin MA, Fang JK, Avadhani NG.** Multiple isoforms of mitochondrial glutathione S-transferases and their differential induction under oxidative stress. *Biochem J.* 2002; 366: 45-55.

[50] **Addya S, Mullick J, Fang JK, Avadhani NG.** Purification and characterization of a hepatic mitochondrial glutathione S-transferase exhibiting immunochemical relationship to the alpha-class of cytosolic isoenzyme. *Arch Biochem Biophys.* 1994; 310: 82-8.

[51] **Goto S, Kawakatsu M, Izumi S, *et al.*** Glutathione S-transferase pi localizes in mitochondria and protects against oxidative stress. *Free Radic Biol Med.* 2009; 46: 1392-403.

[52] **Droog F.** Plant glutathione S-transferases, a tale of theta and tau. *J Plant Growth Regul.* 1997; 16: 95-107.

[53] **Frova C.** Glutathione transferases in the genomics era: new insights and perspectives. *Biomol Eng.* 2006; 23: 149-69.

[54] **Snyder MJ, Maddison DR.** Molecular phylogeny of glutathione-S-transferases. *DNA Cell Biol.* 1997; 16: 1373-84.

[55] **Ranson H, Rossiter L, Ortelli F, *et al.*** Identification of a novel class of insect glutathione S-transferases involved in resistance to DDT in the malaria vector *Anopheles gambiae*. *Biochem J.* 2001; 359: 295-304.

[56] **McGoldrick S, O'Sullivan SM, Sheehan D.** Glutathione transferase-like proteins encoded in genomes of yeasts and fungi: insights into evolution of a multifunctional protein superfamily. *FEMS Microbiol Lett.* 2005; 242: 1-12.

[57] **Apweiler R, Bairoch A, Wu CH.** Protein sequence databases. *Curr Opin Chem Biol.* 2004; 8: 76-80.

[58] **Tamura K, Stecher G, Peterson D, *et al.*** MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Mol Biol Evol.* 2013; 30: 2725-9.

[59] **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.

[60] **Jepson I, Lay VJ, Holt DC, *et al.*** Cloning and characterization of maize herbicide safener-induced cDNAs encoding subunits of glutathione S-transferase isoforms I, II and IV. *Plant Mol Biol.* 1994; 26: 1855-66.

[61] **Dixon DP, Skipsey M, Edwards R.** Roles for glutathione transferases in plant secondary metabolism. *Phytochemistry.* 2010; 71: 338-50.

[62] **Dean JD, Goodwin PH, Hsiang T.** Induction of distinct glutathione transerases of wheat by xenobiotics and by pathogen attack. *Plant Physiol.* 2005; 102: 1193-201.

[63] **Chronopoulou EG, Labrou NE.** Glutathione transferases: emerging multidisciplinary tools in red and green biotechnology. *Recent Pat Biotechnol.* 2009; 3: 211-23.

[64] **Dixon DP, Cole DJ, Edwards R.** Cloning and characterisation of plant theta and zeta class GSTs: Implications for plant GST classification. *Chem Biol Interact.* 2001; 1338: 33-6.

[65] **Edwards R, Dixon DP.** The role of glutathione transferases in herbicide metabolism. In herbicides and their mechanisms of action. In: Anonymous Sheffield. UK: Sheffield Academic Press; 2000. p. 38-71.

[66] **Flury T, Wagner E, Kreuz K.** An inducible glutathione S-transferase in soybean hypocotyl is localized in the apoplast. *Plant Physiol.* 1996; 112: 1185-90.

[67] **Benekos K, Kissoudis C, Nianiou-Obeidat I, *et al.*** Overexpression of a specific soybean *Gm*GSTU4 isoenzyme improves diphenyl ether and chloroacetanilide herbicide tolerance of transgenic tobacco plants. *J Biotech.* 2010; 150: 195-201.

[68] **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.

[69] **Piccolomini R, Aceto A, Allocati N, *et al.*** Purification of a GSH-affinity binding protein from *Bacteroides fragilis* devoid of glutathione transferase activity. *FEMS Microbiol Lett.* 1991; 82: 101-6.

[70] **Shishido T.** Glutathione S-transferase from *Escherichia coli*. *Agric Biol Chem.* 1981; 45: 2951-3.

[71] **Cavalca L, Guerrieri N, Colombo M, *et al.*** Enzymatic and genetic profiles in environmental strains grown on polycyclic aromatic hydrocarbons. *Antonie Van Leeuwenhoek.* 2007; 91: 315-25.

[72] **Hofer B, Backhaus S, Timmis KN.** The biphenyl/polychlorinated biphenyl-degradation locus (bph) of *Pseudomonas sp.* LB400 encodes four additional metabolic enzymes. *Gene.* 1994; 144: 9-16.

[73] **Wang Y, Lau PC, Button DK.** A marine oligobacterium harboring genes known to be part of aromatic hydrocarbon degradation pathways of soil pseudomonads. *Appl Environ Microbiol.* 1996; 62: 2169-73.

[74] **Abel EL, Opp SM, Verlinde CL, *et al.*** Characterization of atrazine biotransformation by human and murine glutathione S-transferases. *Toxicol Sci.* 2004; 80: 230-8.

[75] **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. *Structure.* 1998; 6: 1445-52.

[76] **Labrou NE, Karavangeli M, Tsaftaris A, Clonis YD.** Kinetic analysis of maize glutathione S-transferase I catalysing the detoxification from chloroacetanilide herbicides. *Planta.* 2005; 222: 91-7.

[77] **LaBaron HM, McFarland JE, Burnside OC.** The triazine herbicides. 1st ed. Amsterdam, Netherlands: Elsevier; 2008.

[78] **Smith D, Crowley DE.** Contribution of ethylamine degrading bacteria to atrazine degradation in soils. *FEMS Microbiol Ecol.* 2006; 58: 271-7.

[79] **Smith EF, Townsend CO.** A plant-tumor of bacterial origin. *Science.* 1907; 25: 671-3.

[80] **Kosloff M, Han GW, Krishna SS, *et al.*** Comparative structural analysis of a novel glutathione S-transferase (ATU5508) from *Agrobacterium tumefaciens* at 2.0 Å resolution. *Proteins.* 2006; 65: 527-37.

[81] **Jeon GA, Eum J, Sim WS.** The role of inverted repeat (IR) sequence of the virE gene expression in *Agrobacterium tumefaciens* pTiA6. *Mol Cells.* 1998; 8: 49-53.

[82] **de la Riva GA, González-Cabrera J, Vázquez-Padrón R, Ayra-Pardo C.** *Agrobacterium tumefaciens*: a natural tool for plant transformation. *EJB Journal of biotechnology.* 1998; 1: 1-16.

[83] **Escobar MA, Dandekar AM.** *Agrobacterium tumefaciens* as an agent of disease. *Trends Plant Sci.* 2003; 8: 380-6.

[84] **DeCleene M, DeLay J.** The host range of crown gall. *Bot Rev.* 1976; 42: 389-466.

[85] **Bourras S, Rouxel T, Meyer M.** *Agrobacterium tumefaciens* gene transfer: How a plant pathogen hacks the nuclei of plant and nonplant organisms. *Phytopathology.* 2015; 105: 1288-301.

[86] **White,P.R., Braun,A.C.** A cancerous neoplasm of plants: autonomous bacteria-free crown gall tissue. *Cancer Res.* 1942; 2: 597-617.

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

[88] **Foloppe N, Sagemark J, Nordstrand K, *et al.*** Structure, dynamics and electrostatics of the active site of glutaredoxin 3 from *Escherichia coli*: comparison with functionally related proteins. *J Mol Biol.* 2001; 310: 449-70.

[89] **Armstrong RN.** Glutathione transferases. In: Guengerich FP, editor. Comprehensive Toxicology. 2nd ed.: Elsevier Science; 2010. p. 295-321.

[90] **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.

[91] **Sayed Y, Wallace LA, Dirr HW.** The hydrophobic lock-and-key intersubunit motif of glutathione transferase A1-1: implications for catalysis, ligandin function and stability. *FEBS Lett.* 2000; 465: 169-72.

[92] **Vargo MA, Colman RF.** Heterodimers of wild-type and subunit interface mutant enzymes of glutathione S-transferase A1-1: interactive or independent active sites? *Protein Sci.* 2004; 13: 1586-93.

[93] **Vargo MA, Nguyen L, Colman RF.** Subunit interface residues of glutathione S-transferase A1-1 that are important in the monomer-dimer equilibrium. *Biochemistry.* 2004; 43: 3327-35.

[94] **Nishida M, Harada S, Noguchi S, *et al.*** Three-dimensional structure of *Escherichia coli* glutathione S-transferase complexed with glutathione sulfonate: catalytic roles of Cys10 and His106. *J Mol Biol.* 1998; 281: 135-47.

[95] **McNicholas S, Potterton EA, Wilson KS, Noble M.** Presenting your structures: the CCP4mg molecular-graphics software. *Acta Cryst.* 2011; 67: 386-94.

[96] **Adang AE, Brussee J, Van der Gen A, Mulder GJ.** The glutathione-binding site in glutathione S-transferases. Investigation of the cysteinyl, glycyl and gamma-glutamyl domains. *Biochem J.* 1990; 269: 47-54.

[97] **Douglas KT.** Mechanism of action of glutathione-dependent enzymes. *Adv Enzymol Relat Areas Mol Biol.* 1987; 59: 103-67.

[98] **Hegazy UM, Mannervik B, Stenberg G.** Functional role of the lock and key motif at the subunit interface of glutathione transferase p1-1. *J Biol Chem.* 2004; 279: 9586-96.

[99] **Sinning I, Kleywegt GJ, Cowan SW, *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.

[100] **Stella L, Nicotra M, Ricci G, *et al*.** Molecular dynamics simulations of human glutathione transferase P1-1: analysis of the induced-fit mechanism by GSH binding. *Proteins.* 1999; 37: 1-9.

[101] **Neuefeind T, Huber R, Dasenbrock H, *et al*.** Crystal structure of herbicide-detoxifying maize glutathione S-transferase-I in complex with lactoylglutathione: evidence for an induced-fit mechanism. *J Mol Biol.* 1997; 274: 446-53.

[102] **Laborde E.** Glutathione transferases as mediators of signaling pathways involved in cell proliferation and cell death. *Cell Death Differ.* 2010; 17: 1373-80.

[103] **Hayeshi R, Chinyanga F, Chengedza S, Mukanganyama S.** Inhibition of human glutathione transferases by multidrug resistance chemomodulators *in vitro*. *J Enzyme Inhib Med Chem.* 2006; 21: 581-7.

[104] **McTigue MA, Williams DR, Tainer JA.** Crystal structures of a schistosomal drug and vaccine target: glutathione S-transferase from *Schistosoma japonica* and its complex with the leading antischistosomal drug praziquantel. *J Mol Biol.* 1995; 246: 21-7.

[105] **Oakley AJ, Lo Bello M, Nuccetelli M, *et al*.** 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.

[106] **Reinemer P, Prade L, Hof P, *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.

[107] **Cummins I, Cole DJ, Edwards R.** Purification of multiple glutathione transferases involved in herbicide detoxification from wheat (*Triticum aestivum* L.) treated with the safener fenchlorazole-ethyl. *Biochem Physiol.* 1997; 59: 35-49.

[108] **Scalla R, Roulet A.** Cloning and characterization of a glutathione S-transferase induced by a herbicide safener in barley (*Hordeum vulgare*). *Physiol Plant.* 2002; 116: 336-44.

[109] **Wu JR, Cramer CL, Hatzios KK.** Characterization of two cDNAs encoding glutathione S-transferases in rice and induction of their transcripts by the herbicide safener fenclorim. *Physiol Plant.* 1999; 105: 102-8.

[110] **Dean JV, Gronwald JW, Eberlein CV.** Induction of glutathione S-transferase isozymes in sorghum by herbicide antidotes. *Plant Physiol.* 1990; 92: 467-73.

[111] **Dixon D, Cole DJ, Edwards R.** Characterisation of multiple glutathione transferases containing the GST I subunit with activities toward herbicide substrates in maize (*Zea mays*). *Pestic Sci.* 1997; 50: 72-82.

[112] **Gronwald JW, Plaisance KL.** Isolation and characterization of glutathione S-transferase isozymes from sorghum. *Plant Physiol.* 1998; 117: 877-92.

[113] **Van Aken B.** Transgenic plants for phytoremediation: helping nature to clean up environmental pollution. *Trends Biotechnol.* 2008; 26: 225-7.

[114] **Suresh B, Ravishankar GA.** Phytoremediation--a novel and promising approach for environmental clean-up. *Crit Rev Biotechnol.* 2004; 24: 97-124.

[115] **Andreou VG, Clonis YD.** Novel fiber-optic biosensor based on immobilized glutathione S-transferase and sol-gel entrapped bromocresol green for the determination of atrazine. *Anal Chim Acta.* 2002; 460: 151-60.

[116] **Fragoulaki MN, Axarli IA, Labrou NE, Clonis YD.** Recombinant glutathione S-transferase for the determination of the herbicide alachor: The foundations of an optical biosensor. *1st UK-US Conference on Chemical and Biological Sensors and Detectors.* London, UK. 2007.

[117] **Kapoli P, Axarli IA, Platis D, *et al.*** Engineering sensitive glutathione transferase for the detection of xenobiotics. *Biosens Bioelectron.* 2008; 24: 498-503.

[118] **Choi JW, Kim YK, Song SY, *et al.*** Optical biosensor consisting of glutathione-S-transferase for detection of captan. *Biosens Bioelectron.* 2003; 18: 1461-6.

[119] **Choi JW, Kim YK, Oh BK, *et al.*** Optical biosensor for simultaneous detection of captan and organophosphorus compounds. *Biosens Bioelectron.* 2003; 18: 591-7.

[120] **Hasegawa K, Miwa S, Tajima T, *et al.*** A rapid and inexpensive method to screen for common foods that reduce the action of acrylamide, a harmful substance in food. *Toxicol Lett.* 2007; 175: 82-8.

[121] **Morou E, Ismail HM, Dowd AJ, *et al.*** A dehydrochlorinase-based pH change assay for determination of DDT in sprayed surfaces. *Anal Biochem.* 2008; 378: 60-4.

[122] **Fan KC, Huang YC, Li CH.** Radioimmunoassay for plasma glutathione S-transferase-pi and its clinical application in gastrointestinal cancer. *Cancer.* 1995; 76: 1363-7.

[123] **Johansson AS, Mannervik B.** Human glutathione transferase A3-3, a highly efficient catalyst of double-bond isomerization in the biosynthetic pathway of steroid hormones. *J Biol Chem.* 2001; 276: 33061-5.

[124] **Johansson AS, Mannervik B.** Active-site residues governing high steroid isomerase activity in human glutathione transferase A3-3. *J Biol Chem.* 2002; 277: 16648-54.

[125] **Morgan AS, Ciaccio PJ, Tew KD, Kauvar LM.** Isozyme-specific glutathione S-transferase inhibitors potentiate drug sensitivity in cultured human tumor cell lines. *Cancer Chemother Pharmacol.* 1996; 37: 363-70.

[126] **Axarli I, Labrou NE, Petrou C, *et al.*** Sulphonamide-based bombesin prodrug analogues for glutathione transferase, useful in targeted cancer chemotherapy. *Eur J Med Chem.* 2009; 44: 2009-16.

[127] **Thom R, Cummins I, Dixon DP, *et al.*** Structure of a tau class glutathione S-transferase from wheat active in herbicide detoxification. *Biochemistry.* 2002; 41: 7008-20.

[128] **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.

[129] **Milligan AS, Daly A, Parry MA, *et al.*** The expression of a maize glutathione S-transferase in transgenic wheat confers herbicide tolerance, both in planta and *in vivo*. *Mol Breed.* 2001; 7: 301-5.

[130] **Skopelitou K, Muleta AW, Papageorgiou AC, *et al.*** 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.* 2015; 1854: 166-77.

[131] **Axarli I, Muleta AW, Vlachakis D, *et al.*** Directed evolution of tau class glutathione transferases reveals a site that regulates catalytic efficiency and masks cooperativity. *Biochem J.* 2016; 473(5):559-570.

[132] **Kabsch W.** Xds. *Acta Crystallogr D Biol Crystallogr.* 2010; 66: 125-32.

[133] **Winn MD, Ballard CC, Cowtan KD, *et al.*** Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr.* 2011; 67: 235-42.

[134] **McCoy AJ, Grosse-Kunstleve RW, Adams PD, *et al.*** Phaser crystallographic software. *J Appl Crystallogr.* 2007; 40: 658-74.

[135] **Adams PD, Afonine PV, Bunkoczi G, *et al.*** PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr.* 2010; 66: 213-21.

[136] **Emsley P, Cowtan K.** Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr.* 2004; 60: 2126-32.

[137] **Moriarty NW, Grosse-Kunstleve RW, Adams PD.** electronic Ligand Builder and Optimization Workbench (eLBOW): a tool for ligand coordinate and restraint generation. *Acta Crystallogr D Biol Crystallogr.* 2009; 65: 1074-80.

[138] **Dixon DP, McEwen AG, Lapthorn AJ, Edwards R.** Forced evolution of a herbicide detoxifying glutathione transferase. *J Biol Chem.* 2003; 278: 23930-5.

[139] **Morris RJ, Perrakis A, Lamzin VS.** ARP/wARP and automatic interpretation of protein electron density maps. *Methods Enzymol.* 2003; 374: 229-44.

[140] **Murshudov GN, Vagin AA, Dodson EJ.** Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr.* 1997; 53: 240-55.

[141] **Altschul SF, Gish W, Miller W, *et al.*** Basic local alignment search tool. *J Mol Biol.* 1990; 215: 403-10.

[142] **Salinas AE, Blundell TE.** Comparative protein modeling by satisfaction of spatial restraints. . *J Mol Biol.* 1993; 234: 779-815.

[143] **Chen R, Li L, Weng Z.** ZDOCK: an initial-stage protein-docking algorithm. *Proteins.* 2003; 52: 80-7.

[144] **Li L, Chen R, Weng Z.** RDOCK: refinement of rigid-body protein docking predictions. *Proteins.* 2003; 53: 693-707.

[145] **Brooks BR, Bruccoleri RE, Olafson BD, *et al.*** CHARMM: A program for macromolecular energy, minimization and dynamics calculations. *J Comput Chem.* 1983; 4: 187-217.

[146] **Pettersen EF, Goddard TD, Huang CC, *et al.*** UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem.* 2004; 25: 1605-12.

[147] **Lang PT, Brozell SR, Mukherjee S, *et al.*** DOCK 6: combining techniques to model RNA-small molecule complexes. *RNA.* 2009; 15: 1219-30.

[148] **Trott O, Olson AJ.** AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem.* 2010; 31: 455-61.

[149] **Lau CD, Levesque MJ, Chien S, *et al.*** ViewDock TDW: high-throughput visualization of virtual screening results. *Bioinformatics.* 2010; 26: 1915-7.

[150] **Hess B, Kutzner C, van der Spoel D, Lindahl E.** GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J Chem Theory Comput.* 2008; 4: 435-47.

[151] **Tsui V, Case DA.** Theory and applications of the generalized born solvation model in macromolecular simulations. *Biopolymers.* 2001; 56: 275-91.

[152] **Lovell SC, Davis IW, Arendall WB, *et al.*** Structure validation by Calpha geometry: phi,psi and Cbeta deviation. *Proteins.* 2003; 50: 437-50.

[153] **Laskowski RA, MacArthur MW, Moss DS, Thornton JM.** PROCHECK: A program to check the stereochemical quality of protein structures. *J Appl Cryst.* 1993; 26: 283-91.

[154] **Kleywegt GJ, Jones TA.** Model building and refinement practice. *Methods Enzymol.* 1997; 277: 208-30.

[155] **Shen MY, Sali A.** Statistical potential for assessment and prediction of protein structures. *Protein Sci.* 2006; 15: 2507-24.

[156] **Luthy R, Bowie JU, Eisenberg D.** Assessment of protein models with three-dimensional profiles. *Nature.* 1992; 356: 83-5.

[157] **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.

[158] **Krissinel EB, Winn MD, Ballard CC*, et al.*** The new CCP4 Coordinate Library as a toolkit for the design of coordinate-related applications in protein crystallography. *Acta Crystallogr D Biol Crystallogr.* 2004; 60: 2250-5.

[159] **Krissinel E, Henrick K.** Inference of macromolecular assemblies from crystalline state. *J Mol Biol.* 2007; 372: 774--797.

[160] **Binkowski TA, Naghibzadeh S, Liang J.** CASTp: Computed Atlas of Surface Topography of proteins. *Nucleic Acids Res.* 2003; 31: 3352-5.

[161] **Dundas J, Ouyang Z, Tseng J*, et al.*** CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues. *Nucleic Acids Res.* 2006; 34: W116-8.

[162] **Thompson JD, Higgins DJ, Gibson TJ.** CLUSTALW: 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.

[163] **Kabsch W, Sander C.** Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers.* 1983; 22: 2577-637.

[164] **Gouet P, Courcelle E, Stuart DI, Metoz F.** ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics.* 1999; 15: 305-8.

[165] **Laskowski RA, Swindells MB.** LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. *J Chem Inf Model.* 2011; 51: 2778-86.

[166] **Ramachandran GN, Ramakrishnan C, Sasisekharan V.** Stereochemistry of polypeptide chain configurations. *J Mol Biol.* 1963; 7: 95-9.

[167] **Read RJ.** Improved Fourier coefficients for maps using phases from partial structures with errors. *Acta Cryst.* 1986; A42: 140-9.

[168] **Liebau E, Eckelt VH, Wildenburg G*, et al.*** Structural and functional analysis of a glutathione S-transferase from *Ascaris suum*. *Biochem J.* 1997; 324: 659-66.

[169] **Kyte J, Doolittle RF.** A simple method for displaying the hydropathic character of a protein. *J Mol Biol.* 1982; 157: 105-32.

[170] **Skopelitou K, Dhavala P, Papageorgiou AC, Labrou NE.** A glutathione transferase from *Agrobacterium tumefaciens* reveals a novel class of bacterial GST superfamily. *PLoS One.* 2012; 7: e34263.

[171] **Kampranis SC, Damianova R, Atallah M*, et al.*** A novel plant glutathione S-transferase/peroxidase suppresses Bax lethality in yeast. *J Biol Chem.* 2000; 275: 29207-16.

[172] **Pemble SE, Wardle AF, Taylor JB.** Glutathione S-transferase class Kappa: characterization by the cloning of rat mitochondrial GST and identification of a human homologue. *Biochem J.* 1996; 319 (Pt 3): 749-54.

[173] **Thompson JD, Gibson TJ, Higgins DG.** Multiple sequence alignment using ClustalW and ClustalX. *Curr Protoc Bioinformatics.* 2002; Chapter 2: Unit 2.3.

[174] **Casalone E, Allocati N, Ceccarelli I, *et al*.** Site-directed mutagenesis of the *Proteus mirabilis* glutathione transferase B1-1 G-site. *FEBS Lett.* 1998; 423: 122-4.

[175] **Cocco R, Stenberg G, Dragani B, *et al*.** The folding and stability of human alpha class glutathione transferase A1-1 depend on distinct roles of a conserved N-capping box and hydrophobic staple motif. *J Biol Chem.* 2001; 276: 32177-83.

[176] **Stenberg G, Dragani B, Cocco R, *et al*.** A conserved "hydrophobic staple motif" plays a crucial role in the refolding of human glutathione transferase P1-1. *J Biol Chem.* 2000; 275: 10421-8.