



**HAL**  
open science

# Étude cristallographique de glutathion transférases de micro-organismes impliqués dans la dégradation de la lignine : le basidiomycète *Phanerochaete chrysosporium* et la bactérie *Sphingobium* sp. SYK-6

Pascalita Prosper

► **To cite this version:**

Pascalita Prosper. Étude cristallographique de glutathion transférases de micro-organismes impliqués dans la dégradation de la lignine : le basidiomycète *Phanerochaete chrysosporium* et la bactérie *Sphingobium* sp. SYK-6. Autre [cond-mat.other]. Université de Lorraine, 2013. Français. NNT : 2013LORR0155 . tel-01750222

**HAL Id: tel-01750222**

<https://hal.univ-lorraine.fr/tel-01750222v1>

Submitted on 29 Mar 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



## AVERTISSEMENT

Ce document est le fruit d'un long travail approuvé par le jury de soutenance et mis à disposition de l'ensemble de la communauté universitaire élargie.

Il est soumis à la propriété intellectuelle de l'auteur. Ceci implique une obligation de citation et de référencement lors de l'utilisation de ce document.

D'autre part, toute contrefaçon, plagiat, reproduction illicite encourt une poursuite pénale.

Contact : [ddoc-theses-contact@univ-lorraine.fr](mailto:ddoc-theses-contact@univ-lorraine.fr)

## LIENS

Code de la Propriété Intellectuelle. articles L 122. 4

Code de la Propriété Intellectuelle. articles L 335.2- L 335.10

[http://www.cfcopies.com/V2/leg/leg\\_droi.php](http://www.cfcopies.com/V2/leg/leg_droi.php)

<http://www.culture.gouv.fr/culture/infos-pratiques/droits/protection.htm>



Collegium Sciences & Technologies

Ecole Doctorale SESAMES

Thèse  
présentée pour l'obtention du titre de  
Docteur de l'Université de Lorraine  
en Sciences Physique  
par **Pascalita PROSPER**

**ETUDE CRISTALLOGRAPHIQUE DE GLUTATHION TRANSFERASES DE MICRO-  
ORGANISMES IMPLIQUES DANS LA DEGRADATION DE LA LIGNINE :**  
**LE BASIDIOMYCETE *Phanerochaete chrysosporium***  
**ET LA BACTERIE *Sphingobium* sp. SYK-6**

Soutenance publique prévue le 15 Novembre 2013

Membres du jury :

Rapporteurs :	M. Jean CAVARELLI	Professeur, Université de Strasbourg, France
	M. David COBESSI	Chargé de recherche CNRS, Grenoble
Examineurs :	M. Eric GELHAYE	Professeur, Université de Lorraine
	M. Claude DIDIERJEAN	Maître de Conférences, Université de Lorraine

Directeurs de Thèse :

M. Claude DIDIERJEAN	Directeur de thèse, Université de Lorraine
Mme Frédérique FAVIER	Co-directeur de thèse, Université de Lorraine



## Remerciements

Je remercie le Professeur Jean Cavarelli et le Docteur David Cobessi pour l'honneur qu'ils m'ont fait de juger ce travail. Je remercie également le Professeur Eric Gelhaye, le Docteur Claude Didierjean et le Docteur Frédérique Favier qui ont accepté de participer au jury de soutenance.

Je remercie les Professeur Claude Lecomte et Dominik Schaniel, successivement directeurs du CRM2, pour m'avoir accueillie au laboratoire et m'avoir permis de réaliser ces travaux.

Je remercie mes maîtres de stage Claude Didierjean et Frédérique Favier qui m'ont encadrée pour le travail présenté dans ce manuscrit.

Je remercie le Professeur Eric Gelhaye et toute son équipe au laboratoire IAM, en particulier le Docteur Mélanie Morel, ainsi qu'Edgar Meux et Yann Mathieu, avec qui notre équipe a fructueusement collaboré lors de ce projet.

Je remercie tout le personnel du CRM2, les chercheurs et enseignants chercheurs, le personnel technique et administratif pour l'aide qu'ils m'ont apportée ainsi que pour leur bonne humeur.

Je remercie l'ANR, le Ministère de la Recherche et de l'Enseignement Supérieur, l'Institut Jean Bariol, l'école doctorale SESAMES et l'Université de Lorraine qui m'ont permis de réaliser mes études de doctorat.

Je remercie les personnels techniques et d'encadrement des lignes synchrotron de l'ESRF et de SOLEIL, en particulier David Cobessi et Pierre Legrand pour leur aide précieuse lors des enregistrements de données de diffraction.

Je remercie mes anciens collègues de travail, en particulier Alda Navazza, Ahmed Haouz et les membres de la plateforme 6 de l'Institut Pasteur qui m'ont soutenue pendant toutes mes années de stage.

Je remercie enfin ma maman et toute ma famille, ainsi que mes amis, pour leur soutien.

Et enfin mon Dieu qui veille sur moi et qui m'a permis de finir ce travail.



## **Résumé :**

L'étude des champignons saprophytes tel que *Phanerochaete chrysosporium* est fondamentale car ces organismes, capables de dégrader la lignine, ont un fort potentiel en biotechnologie. Il a été démontré que la bactérie *Sphingobium* sp. SYK-6 possède des enzymes (*ligE*, *ligF* et *ligG*) appartenant à la famille des glutathion transférases (GST) qui successivement réduisent le lien éther (rôle éthérase) entre les unités de la lignine et déglutathionylent le produit dérivé conjugué (rôle lyase). Ce travail expose les relations entre la structure et la fonction des LigE, F et G de *Sphingobium* sp. SYK-6 et de deux classes de GST du champignon saprophyte *Phanerochaete chrysosporium* : une potentiellement liée aux éthérases (GSTFuA) et une potentiellement reliée aux lyases (GST Xi).

Les structures cristallographiques des GSTFuA1, 2 et 3 de *P. chrysosporium* ont été résolues. L'analyse des modèles a permis de révéler une nouvelle classe structurale de GST avec des propriétés uniques. Ces caractéristiques ont pu être reliées à la fonction de ligandine et au profil catalytique de cette nouvelle classe de GST. Parallèlement, les études structurales des LigE et F de *Sphingobium* sp. SYK-6 sont en voie d'achèvement.

La structure cristallographique de la GSTX1 de *P. chrysosporium* présente des caractéristiques structurales spécifiques qui nous ont conduit à la proposer comme leader d'une nouvelle classe structurale nommée Xi. L'enzyme ne présente pas d'activité lyase vis-à-vis des substrats des LigG, mais en revanche possède une fonction hydroquinone réductase (GHR). Parallèlement, la LigG a été étudiée, et s'apparente structurellement et fonctionnellement la classe Oméga des GST.

## **Abstract :**

*Phanerochaete chrysosporium* is a very interesting saprophytic fungus because it decays wood, by degrading lignin while leaving cellulose which is a renewable source of energy. The soil bacterium *Sphingobium* sp. SYK-6 possesses enzymes (LigE, LigF and LigG) that cleave the  $\beta$ -aryl ether linkage of lignin model compounds. LigE and LigF are glutathione dependent enzymes that reduce the ether bond (etherase activity) and LigG catalyzes the elimination of glutathione (lyase activity). This study presents the structure-function relationships of Lig enzymes and of two new classes of GST from *P. chrysosporium* : one potentially related to LigE (GSTFuA) and one potentially connected to LigG (GST Xi).

The crystallographic structures of GSTFuA1, 2 and 3 from *P. chrysosporium* were solved. The analysis of the models reveals a new structural class of GST with unique properties. These characteristics could be connected to the ligandin function and to the catalytic pattern of this new class of GST. In parallel, structural studies of LigE and LigF from *Sphingobium* sp. SYK-6 are nearly completed.

The crystallographic structure of the GSTX1 from *P. chrysosporium* exhibits specific structural properties which allowed us to define a new structural class (Xi) in the GST superfamily. The enzyme is a *S*-glutathionyl-(chloro)hydroquinone reductase (GHR) that does not present a lyase activity with LigG substrate. In parallel, high resolution structure of LigG was obtained; this enzyme can be related structurally and functionally to the GST Omega class.





## Table des matières

A Etat de l'art.....	5
A.I Les GST d'un point de vue général.....	6
A.I.1 Réactions catalysées par les GST.....	6
A.I.2 Rôles des GST.....	8
A.I.2.a La détoxification.....	8
A.I.2.b D'autres activités des GST.....	9
A.I.3 Classifications des GST.....	10
A.I.3.a Classification selon la localisation subcellulaire.....	10
A.I.3.b Classification courante à critères multiples.....	11
A.I.3.c Classification selon la nature du résidu catalytique.....	14
A.I.4 La structure tridimensionnelle des GST.....	15
A.I.4.a Les GST cytosoliques.....	15
i Le repliement tridimensionnel du monomère.....	15
ii Structures quaternaires.....	17
iii Le site de fixation du GSH : Site G.....	18
iv Le site de fixation des substrats hydrophobes : Site H.....	20
v Les autres sites de fixation de ligands : sites L.....	21
A.I.4.b Les GST mitochondriales.....	23
A.I.4.c Les GST MAPEG.....	25
A.I.5 Évolution des GST.....	26
A.II Les glutathion transférases à cystéine.....	29
A.II.1 Les GST Omega.....	29
A.II.2 Les CLIC.....	32
A.II.3 Les GST Bêta.....	33
A.II.4 Les GST Lambda et les DHAR.....	34
A.III Les GST et la dégradation du bois.....	35
A.III.1 La lignine.....	35
A.III.2 La dégradation de la lignine.....	38
A.III.3 La coupure de liens $\beta$ -aryl éther catalysée par des GST de <i>Sphingobium</i> sp. SYK-6.....	38
A.IV Les GST de <i>Phanerochaete chrysosporium</i> .....	40
A.IV.1 Présentation du champignon.....	40
A.IV.2 Les GST fongiques.....	42
A.V Objectifs de la thèse.....	44
B Résultats.....	47
B.I Etude structurale de la PcGHR1 (ou PcGSTX1).....	48
B.I.1 Introduction.....	48
B.I.2 Résultats et discussion.....	49
B.I.3 Publication.....	51
B.II Etude structurale de la PcGSTX1-C86S en complexe avec des ligands.....	67
B.II.1 Introduction.....	67
B.II.2 Matériel et Méthodes.....	67
B.II.2.a Cristallisations.....	67
B.II.2.b Collecte et traitement des données de diffraction.....	68

B.II.3 Résultats.....	70
B.II.3.a Complexes avec le GS-hexyl et le GSDNB.....	70
B.II.3.b Complexe avec la GS-ménadione.....	70
B.II.3.c Conclusion.....	72
B.III Etudes structurales des PcGSTFuA.....	74
B.III.1 PcGSTFuA1.....	74
B.III.1.a Introduction.....	74
B.III.1.b Résultats et discussion.....	74
B.III.1.c Publication PcGSTFuA1.....	75
B.III.2 PcGSTFuA2 et PcGSTFuA3.....	100
B.III.2.a Introduction.....	100
B.III.2.b Résultats et discussion.....	100
B.III.2.c Publication GSTFuA2-3.....	101
B.IV Etudes structurales des Lig.....	130
B.IV.1 LigE.....	131
B.IV.1.a Introduction.....	131
B.IV.1.b Matériel et Méthodes.....	131
i Cristallisation.....	131
ii Collecte des données de diffraction, phasage et construction du modèle.....	132
B.IV.2 LigF.....	134
B.IV.2.a Introduction.....	134
B.IV.2.b Matériel et Méthodes.....	135
i Cristallisation de la protéine.....	135
ii Collecte des données de diffraction, phasage et construction du modèle.....	136
B.IV.2.c Résultats.....	136
i Analyse globale du modèle préliminaire.....	136
ii Analyse de la structure quaternaire.....	138
iii Analyse préliminaire de la structure.....	139
B.IV.3 LigG.....	140
B.IV.3.a Introduction.....	140
B.IV.3.b Résultats et discussion.....	141
B.IV.3.c Publication.....	142
C Discussion.....	153
C.I Quelles sont les caractéristiques structurales communes aux GST ? .....	156
C.II Quelles sont les caractéristiques structurales distinguant les GST ?.....	159
C.II.1 La structure du monomère.....	159
C.II.2 La structure quaternaire.....	162
C.II.3 Les résidus permettant la catalyse.....	164
D Conclusion et perspectives:.....	167
E Annexe : l'arbre de similarité des structures de GST.....	171
E.I Recherche des structures de GST dans la Protein Data Bank.....	171
E.II Constitution d'une base de données non redondante de structures de GST.....	171
E.III Alignement structural des entrées de la base de données non redondante des GST.....	172
E.IV Construction de l'arbre de similarité des séquences de GST alignées selon la superposition structurale.....	172
F Références bibliographiques.....	185

## Abréviations utilisées

3MGA:	3-O-methylgallate
ABC :	'ATP Binding Cassette'
ADN :	Acide DesoxyriboNucleique
ARNt :	Acide RiboNucleique de tranfert
CDNB :	1-chloro-2,4-dinitrobenzène
CLIC :	'Intacellular Chloride Ions Channel'
FLAP :	'Arachidonate 5-lipoxygenase-activating protein'
DHAR:	DésHydroAscorbate Réductase
DsbA :	Disulfide Binding Protein A
eEF1B $\gamma$ :	'Eukaryotic elongation factor 1B', sous-unité gamma
GDP/GTP :	Guanosine DiPhosphate/ Guanosine TriPhosphate
GGE:	Guaiacylglycérol- $\beta$ -guaiacyl éther
GHR :	S-glutathionyl-(chloro)hydroquinone réductase
GRX2 :	Glutarédoxine 2
GS-HPV :	$\beta$ -hydroxypropiovanillone glutathionylée
GSDNB :	1-(S-glutathionyl)-2,4-dinitrobenzène
GSH :	Glutathion
GST :	Glutathion transférases
HEPES :	Acide 4-(2-hydroxyéthyl)-1-piperazine éthane sulfonique
H <sub>2</sub> O <sub>2</sub> :	Hydroperoxyde d'hydrogène
HO $\cdot$ :	Radical hydroxyle
LiP :	Lignine peroxydases
LTC4S :	LeukoTriene C <sub>4</sub> Synthase
MAK16:	'Maintenance of killer protein 16'
MAPEG :	'Membrane-Associated Proteins in Eicosanoid and Glutathione metabolism'
MES :	Acide 2-(N-morpholino)-éthanesulfonique
MnP :	Manganèse peroxydases
MPPHV :	$\alpha$ -(2-méthoxyphénoxy)- $\beta$ -hydroxypropiovanillone
MRP :	'Multidrug Resistance-related Protein'
mPGES-2:	Prostaglandine E synthases de type microsomales
NO $\cdot$ :	Radical nitrosyle

$O_2^{\cdot-}$ :	Radical superoxyde
PDB :	'Protein Data Bank'
PCA :	Protocatechuate
RNS :	'Reactive Nitrogen species'
ROS :	'Reactive Oxygen species'
-SH :	Thiol
SN :	Substitutions nucléophiles (de type SN1, SN2)
SNAr :	Substitutions nucléophiles aromatiques
SspA :	'Stringent starvation protein A'

## A Etat de l'art

Cela fait plusieurs décennies que les scientifiques voient dans les composés de la biomasse une nouvelle source d'énergie renouvelable. Des techniques industrielles sont en cours de développement afin d'extraire et de transformer la cellulose du bois en biocarburant appelé éthanol de deuxième génération. Ces techniques tentent de copier différentes étapes du processus de biotransformation naturelle du bois qui font intervenir principalement des microorganismes (champignons et bactéries).

Les champignons ligninolytiques sont des acteurs majeurs intervenant dans la dégradation du bois. Ils sont capables de dégrader la totalité des polymères constitutifs du bois en rompant les liaisons notamment par oxydation radicalaire. L'étape de dégradation de la lignine, qui est un polymère phénolique, est très importante dans le processus de production de bioéthanol car elle permet d'avoir accès à la cellulose, un polymère de glucose transformable en éthanol par fermentation sous l'action de levures selon le même processus que pour l'éthanol dit de première génération.

Le basidiomycète *Phanerochaete chrysosporium* est un champignon très étudié et très intéressant car il dégrade entièrement la lignine. Dans le cadre de cette étude une analyse génomique a permis de montrer que *P. chrysosporium* et les saprophytes apparentés présentent la particularité de posséder un nombre élevé de gènes codant pour des protéines appartenant à la famille structurale des glutathion transférases. Ces enzymes sont particulièrement bien connues pour leur implication dans la phase II de la détoxification cellulaire des composés xénobiotiques. Des travaux menés par Masai et son équipe (Masai *et al.*, 2003) ont mis en évidence, dans la bactérie *Sphingobium* sp. SYK-6, un cluster de gènes (*ligDFEG*) codant majoritairement pour des glutathion transférases impliquées dans la dégradation de fragments de lignine en rompant ses liaisons éther en présence de glutathion. Une analyse de séquence a permis de relier la protéine LigE à la famille de protéines GSTFuA spécifique du règne des champignons et plus précisément des basidiomycètes saprophytes. Ce travail de thèse fait partie d'une étude plus globale où l'un des buts est de caractériser la famille de protéines GSTFu chez *P. chrysosporium* pour tenter de découvrir si ce champignon possède en plus de sa voie classique impliquant en particulier les ligninases, une voie similaire à celle de la bactérie du sol *Sphingobium* sp. SYK-6 pour dégrader la lignine.

Ce manuscrit présente la contribution cristallographique de cette étude, au cours de laquelle les structures tridimensionnelles de trois GSTFuA ont été déterminées et analysées. Par

ailleurs, la structure cristallographique d'une GST de *P. chrysosporium* (PcGSTO1) ayant potentiellement la même activité lyase que la protéine LigG de *Sphingobium* sp. SYK-6 a été étudiée. Il a été également nécessaire d'entreprendre les études cristallographiques des LigF, E et G de *Sphingobium* sp. SYK-6. Malheureusement seule la structure de LigG a pu être obtenue à haute résolution. Après une présentation générale de la superfamille des glutathion transférases et de son potentiel dans la dégradation du bois, l'ensemble des résultats est décrit et discuté.

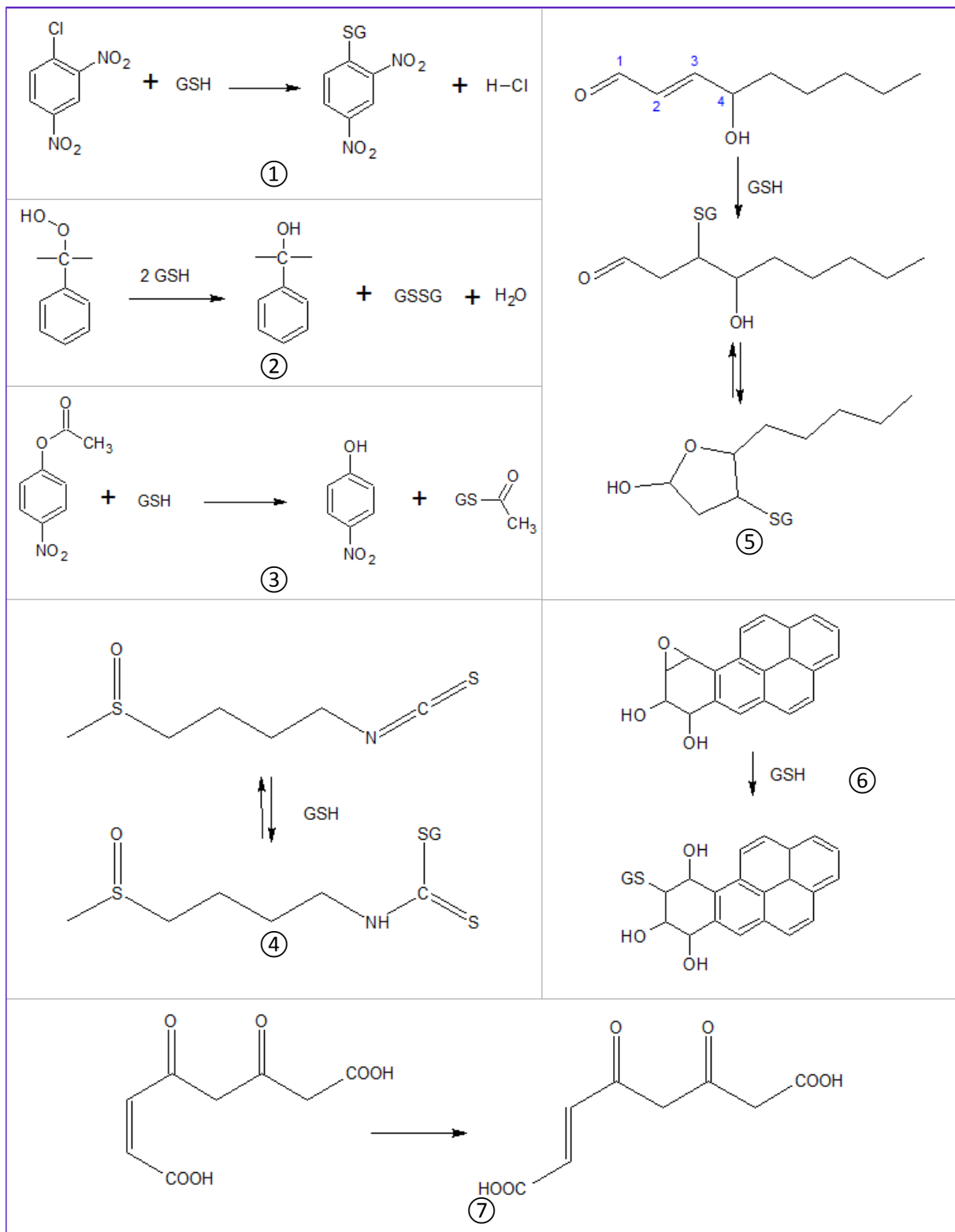
## **A.I Les GST d'un point de vue général**

### **A.I.1 Réactions catalysées par les GST**

Les GST forment une superfamille d'enzymes présentes a priori dans tous les organismes. Des GST ont été caractérisées expérimentalement dans tous les règnes du vivant à l'exception des archaées. Concernant ces dernières, une étude génomique *in silico* récente a identifié des gènes d'archaées de GST proches de ceux des bactéries (Allocati *et al.*, 2012). Les GST catalysent de nombreuses réactions. La plus courante, et qui a donné son nom aux GST, est la conjugaison du glutathion (activité transférase) aux xénobiotiques, avec la formation d'un thioether (Hayes *et al.*, 2005). Parmi les réactions de conjugaison, on distingue les substitutions nucléophiles (de type SN1, SN2), les substitutions nucléophiles aromatiques (SNAr), les réactions d'addition de Michaël d' $\alpha$ - $\beta$ -cétones insaturées ou de quinones et les réactions d'ouverture de cycles époxydes (Illustration 1). Comme la majorité des GST transfère le glutathion sur le 1-chloro-2,4-dinitrobenzène (CDNB) pour donner le 1-(S-glutathionyl)-2,4-dinitrobenzène (GSDNB), ce substrat électrophile est considéré comme le substrat classique de cette superfamille d'enzyme (Clark *et al.*, 1973).

Notons que certaines GST (qui seront décrites dans la suite du manuscrit) catalysent préférentiellement la réaction inverse de la conjugaison, autrement dit la déglutathionylation de composés initialement conjugués au GSH. Cette activité est également appelée thiol transférase et nécessite le plus souvent la présence de glutathion pour régénérer l'enzyme. Par exemple, une activité S-glutathionyl-(chloro)hydroquinone réductase (GHR) a été découverte récemment dans une GST de la bactérie *Sphingobium chlorophenicum* qui a la capacité de minéraliser les pentachlorophénols (Xun *et al.*, 2010).

Les GST catalysent également la réduction de peroxydes, du déshydroascorbate, de la nitroglycérine, ou de l'acide monométhylarsonique (Illustration 1). Elles peuvent aussi conduire à des thiolyses d'esters aromatiques en alcools correspondants (Hayes *et al.*, 2005).



**Illustration 1 - Quelques exemples de réactions catalysées par les GST**

① : glutathionylation du 1-chloro-2,4-dinitrobenzène (CDNB); ② : coupure peroxydatique du cumène hydroperoxyde; ③ : thiolyse de 4-nitrophénylacétate; ④ : glutathionylation de l'isothiocyanate sulfuraphane; ⑤ : glutathionylation du 4-hydroxynonéal (4-HNE); ⑥ : glutathionylation du 7,8-dihydroxy-9,10-oxo-7,8,9,10-tetrahydrobenzo[a]pyrène (BPDE); ⑦ : isomérisation du maleylacétoacétate.

Sans pouvoir être exhaustif, citons finalement l'isomérisation des alcènes comme le maleylacétoacétate et le  $\Delta^5$ -3-cétostéroïde catalysée par les GST (Hayes *et al.*, 2005).

## **A.I.2 Rôles des GST**

### **A.I.2.a La détoxification**

On appelle détoxification la transformation biologique que subissent certains composés exogènes (médicaments par exemple) ou endogènes (hormones stéroïdes, bilirubine) en vue de leur élimination de l'organisme (Jean-Paul Leroux, Encyclopédie Universalis). La compréhension fine de ce processus complexe dans les organismes vivants est aujourd'hui plus que nécessaire. En effet ceux-ci sont en permanence confrontés à de nouvelles espèces chimiques toxiques provenant de l'industrie et qui polluent l'environnement. La superfamille des glutathion transférases joue un rôle clé dans la détoxification des composés endogènes et exogènes à travers de multiples réactions.

La voie d'élimination dans laquelle la plupart des GST sont impliquées est décrite en trois phases (Sheehan *et al.*, 2001):

Les phases I et II sont responsables de la conversion des xénobiotiques lipophiles non polaires en métabolites plus solubles et donc moins toxiques qui peuvent être alors éliminés plus facilement de la cellule au cours de la phase III.

La phase I est prise en charge principalement par le système du cytochrome P450 qui est une famille de protéines microsomales responsables d'un grand nombre de réactions, telles que l'oxydation qui semble être la fonction la plus importante.

Les enzymes de la phase II catalysent la conjugaison de xénobiotiques avec un substrat endogène soluble dans l'eau, comme le glutathion réduit, l'acide UDP-glucorique ou la glycine. La phase II fait intervenir principalement des enzymes appelées les glutathion transférases (GST) dont la principale fonction est de transférer un glutathion sur une molécule contenant un centre électrophile. Les composés conjugués en résultant sont plus solubles et moins toxiques que les composés dont ils dérivent, et donc plus faciles à éliminer de la cellule.

Lors de la phase III les composés conjugués sont excrétés de la cellule pour être soit éliminés soit séquestrés selon l'organisme. Lors de ce processus plusieurs systèmes de transport ont été décrits tels que la protéine MRP1 pour « *Multidrug Resistance-related Protein* » qui fait partie de la famille des transporteurs ABC (ATP Binding Cassette) (Sheehan *et al.*, 2001; Oakley, 2011).

Dans le cadre de ce processus de détoxification des xénobiotiques, les GST interviennent dans l'élimination de nombreux composés tels que des carcinogènes, des drogues thérapeutiques. Ainsi, les GST peuvent empêcher l'action d'agents thérapeutiques comme cela a été montré dans le cancer. Par exemple, l'augmentation de l'expression de GST dans des cellules tumorales semble être corrélée à la résistance de ces cellules à certains médicaments tels que le chlorambucil (Laborde,



2010). Dans les plantes et les insectes, les GST confèrent respectivement une résistance aux herbicides (Andrews *et al.*, 1997) et aux insecticides (Tang & Tu, 1994) et dans les bactéries une résistance aux antibiotiques (Arca *et al.* 1997). Les GST conjuguent également le glutathion à des composés endogènes tels que des hormones stéroïdiennes (Johansson & Mannervik, 2001), des intermédiaires métaboliques de la dégradation de la tyrosine (Lim *et al.*, 2004) et des produits du stress oxydant.

Le stress oxydant est un processus de régulation cellulaire faisant intervenir des espèces oxygénée et azotées très réactives appelées ROS ( $O_2^{\cdot-}$ ,  $HO^{\cdot-}$  et  $H_2O_2$ ) et RNS ( $NO^{\cdot}$ ) respectivement. Ces espèces oxygénées ou azotées proviennent du métabolisme de l'oxygène de l'air dans la voie de respiration aérobie et du métabolisme de l'azote. Ce sont des radicaux libres capables de causer de sérieux dommages aux lipides membranaires, à l'ADN, aux protéines et aux carbohydrates. Le radical superoxyde  $O_2^{\cdot-}$  peut attaquer l'ADN soit directement, soit indirectement via les sous-produits issus de l'attaque des acides gras de la membrane. Ces métabolites ont une durée de vie très courte et se dégradent en des électrophiles secondaires tels que les époxyaldéhydes, les 2-alcénals, les 4-hydroxy-2-alcénals, les cétoaldéhydes et les hydroperoxydes de phospholipides. En plus de participer à l'élimination d'aldéhydes et d'époxydes par conjugaison, les GST participent également à la réduction des hydroperoxydes (Hayes *et al.*, 2005).

### **A.I.2.b D'autres activités des GST**

Plusieurs GST utilisent le GSH non comme substrat mais comme co-facteur dans la biosynthèse d'eicosanoïdes tels que les leucotriènes et les prostaglandines (Hayes *et al.*, 2005). Des fonctions particulières ont également été découvertes comme l'implication de certaines GST dans l'inhibition de l'activité de signalisation des c-Jun N-terminales kinases (Adler *et al.*, 1999).

L'Ure2p de la levure *Saccharomyces cerevisiae* est un cas particulier dans la superfamille des GST. En effet, elle présente une élongation N-terminale capable de former des fibrilles amyloïdes. Cette Ure2p est un facteur de transcription impliqué dans la répression de la voie d'utilisation de l'urédosuccinate en inhibant le facteur de transcription Gln3p. Dans ce processus dépendant de l'azote, l'Ure2p empêche l'entrée de Gln3p dans le noyau (Bousset *et al.*, 2001). Les études *in vitro* ont montré que la partie soluble de l'Ure2p présentait une activité peroxydase et une activité disulfure oxydoréductase similaire à celle des glutarédoxines (Bai *et al.*, 2004 ; Zhang & Perret, 2009).

Il est également intéressant de citer le facteur d'élongation eEF1B de levure (pour « eukaryotic elongation factor 1B ») qui est constitué de deux sous-unités, un domaine d'échange de

GDP/GTP nommé eEF1B $\alpha$  et un domaine de structure nommé eEF1B $\gamma$ . Ce dernier appartient à la famille des GST. Le facteur eEF1B a pour rôle de recharger en GTP le facteur eEF1A, qui catalyse le transfert de l'aminoacyl-ARNt sur le ribosome lors de phase d'élongation de la synthèse des protéines. Le domaine eEF1B $\gamma$  semble être important dans l'activité du domaine eEF1B $\alpha$  du facteur eEF1B (voir références dans l'article de Jeppesen *et al.*, 2003). De plus, il a été montré que la sous-unité eEF1B $\gamma$  pouvait présenter une activité avec le CDNB.

Par ailleurs, les GST peuvent fixer un grand nombre de molécules lipophiles (masse moléculaire > 400 daltons), et sont de ce fait impliquées dans le stockage et le transport rapide de molécules dans la phase aqueuse de la cellule (Litwack *et al.*, 1971 ; Habig *et al.*, 1974 ; Tipping & Ketterer, 1981 ; Caccuri *et al.*, 1990). Cette propriété qualifiée de « ligandine » a été observée dès le début des années 70 dans plusieurs études sur une protéine abondante dans le rein des rats, et plus tard identifiée comme une GST. Cette GST montre une forte activité de fixation (avec des constantes d'affinité de  $10^5 - 10^7 \text{ M}^{-1}$ ) envers de nombreux ligands anioniques lipophiles (Oakley *et al.*, 1999). Ces ligands incluent des molécules comme l'hème, la bilirubine, les sels biliaires, les stéroïdes, les hormones thyroïdiennes, les acides gras et les médicaments.

Ce caractère de « ligandine » a été montré plus tard être une propriété de beaucoup d'isoformes de GST. La fixation de type « ligandine » est caractérisée par une inhibition non-compétitive envers le substrat classique des GST, le CDNB (Bico *et al.*, 1995). Plus récemment, des études biochimiques ont révélé que des GST de maïs avaient la capacité de fixer des porphyrines sans le conjuguer au glutathion (Lederer & Böger, 2003 ; Dixon *et al.*, 2008).

### **A.I.3 Classifications des GST**

#### **A.I.3.a Classification selon la localisation subcellulaire**

Les glutathion transférases sont tout d'abord divisées en trois grandes familles<sup>1</sup> selon leur localisation subcellulaire : les GST cytosoliques, les GST mitochondriales et les GST microsomaux appelées MAPEG pour « *Membrane-Associated Proteins in Eicosanoid and Glutathione metabolism* ». Chaque famille est subdivisée en plusieurs sous-groupes différents d'isoenzymes appelées classes. Par exemple, les GST MAPEG ont été subdivisées en 4 classes MAPEG I à IV sur la base d'une analyse phylogénétique (Bresel *et al.*, 2005). Les trois grandes familles de GST ont en commun les activités glutathion transférase et peroxydase. Néanmoins, des différences structurales significatives sont présentes (voir paragraphe A.I.4).

---

1 Il existe aussi des GST localisées dans le chloroplaste chez les plantes (voir par exemple paragraphe A.II.4).

### A.I.3.b Classification courante à critères multiples

Les GST cytosoliques de mammifères ont été les premières à être particulièrement bien caractérisées et classifiées selon une combinaison de critères tels que la similitude des séquences primaires, la similitude des structures tridimensionnelles, les spécificités d'inhibiteurs et de substrats et les propriétés immunologiques (Hayes & Mantle, 1986). A l'origine, les GST humaines ont été regroupées dans des classes spécifiques, nommées Alpha, Mu, Pi et Thêta d'après une nomenclature basée sur des lettres grecques (Mannervik *et al.*, 1992). Dans une classe, les GST sont numérotées suivant l'ordre de découverte des isoformes. Ces GST étant dimériques, la nomenclature a également pour objectif de distinguer les GST homo et hétérodimériques. Par exemple, une GSTP1-1 signifie que l'enzyme est une GST homodimérique de la classe Pi et GSTP1-2 veut dire que l'on a un hétérodimère entre l'isoforme P1 et l'isoforme P2. Plus récemment, cette nomenclature a été étendue aux GST de mammifères, avec l'apparition des classes Oméga, Sigma et Zêta (Mannervik *et al.*, 2005). La classe Thêta est très distincte des classes Alpha, Pi et Mu car elle ne répond pas à certains critères classiques des GST établis auparavant tels que l'affinité pour le substrat classique (le CDNB) et aussi le fait que les enzymes de cette classe ne se fixent pas sur des colonnes de purification d'agarose (Meyer *et al.*, 1991).





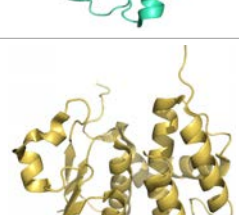

Dans un premier temps, les GST des organismes autres que les mammifères, moins bien caractérisées, ont été rangées dans la classe Thêta. plus tard, les études cristallographiques et biochimiques sur ces GST ont révélé de nouvelles caractéristiques impliquant l'établissement de nouvelles classes de GST. Sans parler des classifications annexes, les GST non mammifères ont été aussi classées avec des lettres grecques. Certaines classes sont spécifiques à certains organismes et d'autres sont retrouvées dans plusieurs règnes. Par exemples les classes Phi et Delta sont spécifiques respectivement aux plantes et aux insectes. La classe Oméga est présente à la fois chez les animaux et les champignons. Les classes répertoriées à ce jour utilisant des lettres grecques sont listées dans le Tableau 1.


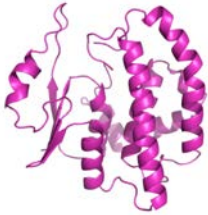
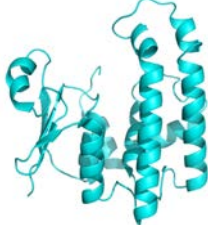
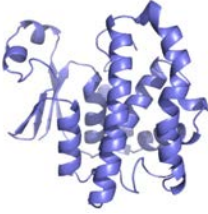



Aujourd'hui, la classification n'utilise quasiment que le critère d'identité de séquence. Ainsi, deux GST appartiennent sans ambiguïté à la même classe si elles partagent plus de 60 % d'identité et n'appartiennent pas à la même classe si l'identité est inférieure à 30 % (Sheehan *et al.*, 2001). Ces bornes varient toutefois selon les auteurs (respectivement 40 et 25 % selon Oakley (2011)).



Par ailleurs, plusieurs protéines possédant les caractéristiques topologiques des GST cytosoliques mais ne possédant pas d'activité transférase n'ont pas été classées avec des lettres grecques. Ce sont les CLIC (pour « *Intacellular Chloride Ions Channel* ») (Harrop *et al.*, 2001), les prostaglandine E synthases de type microsomales (mPGES-2) (Yamada *et al.*, 2005), le facteur

d'élongation de la levure EF1B $\gamma$  (Jeppesen *et al.*, 2003), la protéine Ure2p de la levure (Bousset *et al.*, 2001), la protéine SspA (pour « Stringent starvation protein A ») (Hansen *et al.*, 2005) et la déshydroascorbate réductase (DHAR) (Dixon *et al.*, 2002).

Pour finir, les GST mitochondriales ont été classées dans une seule et unique classe nommée Kappa.

Classe	Organismes	Résid.	Activité catalytique	Activité non catalytique	Structure
Alpha	mammifères oiseaux	Y	Conjugaison du GSH, peroxydase, isomérase	Modulation du signal	
Bêta	bactéries	C/S	Conjugaison du GSH, thiol transférase	Cys pas toujours nécessaire à la fonction transférase	
Delta	insectes	S	conjugaison du GSH		
Epsilon	insectes	S	Peroxydase,		
Zêta	mammifères insectes plantes bactéries	S	Conjugaison du GSH, peroxydase, isomérase	Isomérisation glutathion- dépendante du maleylacétoacétate en fumarylacétoacétate : catabolisme de Phe et Tyr	
Êta	bactéries	R	Conjugaison du GSH, peroxydase		

Classe	Organismes	Résid.	Activité catalytique	Activité non catalytique	Structure
Thêta	mammifères insectes plantes bactéries	S	Conjugaison du GSH, peroxydase	Glutathion peroxydases intervenant dans la détoxification des hydroperoxydes organiques formés sous stress oxydant	
Lambda	plantes	C	Thiol transférase		?
Mu	mammifères oiseaux nématodes	Y	Conjugaison du GSH, peroxydase, isomérase	Modulation du signal	
Nu	nématodes	Y		Ligandine	
Pi	mammifères	Y	Conjugaison du GSH, peroxydase, isomérase	Modulation du signal	
Rhô	mollusques	S	Conjugaison du GSH		
Sigma	mammifères insectes	Y	Conjugaison du GSH, peroxydase, isomérase	Glutathion-dépendant prostaglandine D2 synthase	
Tau	plantes	S	Conjugaison du GSH, peroxydase,	Ligandine et modulation du signal	

Classe	Organismes	Résid.	Activité catalytique	Activité non catalytique	Structure
Phi	plantes	S	Conjugaison du GSH, peroxydase, isomérase	Ligandine	
Chi	cyano-bactéries	?		Comparable aux GST Bêta mais pas de résidu Cys catalytique	?
Oméga	mammifères insectes nématodes champignons	C	Conjugaison du GSH, thiol transférase, déshydroascorbate réductase	Modulation des canaux ioniques	

**Tableau 1** - Principales propriétés des GST cytosoliques présentées selon la nomenclature utilisant des lettres grecques.

Pour chaque classe, une référence significative est donnée et lorsque la structure est connue le code pdb qui a été utilisé pour réaliser la figure d'un monomère représenté en flèches et rubans : alpha (Sinning *et al.*, 1993 ; code pdb : 1GUK), Bêta (Rossjohn *et al.*, 1998b ; code pdb : 2V6K), Delta (Wilce *et al.*, 1995; Oakley *et al.*, 2001 ; code pdb : 1V2A), Epsilon (Sawicki *et al.*, 2003 ; code pdb : 2IL3), Zêta (Polekhina *et al.*, 2001 ; code pdb : 1E6B), Êta (Skopelitou *et al.*, 2012 ; code pdb : 2YCD), Thêta (Rossjohn *et al.*, 1998a ;code pdb : 2C3N), Lambda (Dixon *et al.*, 2002), Mu (Ji *et al.*, 1994 ; code pdb : 3GST), Nu (Schuller *et al.*, 2005 ; code pdb : 2WS2), Pi (Reinemer *et al.*, 1992 ; code pdb : 3GSS), Sigma (Ji *et al.*, 1995; Kanaoka *et al.*, 1997 ; code pdb : 1MOU), Tau (Thom *et al.*, 2002 ;code pdb : 2VO4), Phi (Reinemer *et al.*, 1996 ; code pdb : 1AW9), Khi (Wiktelius & Stenberg, 2007), Rhô (Park *et al.*, 2013 ;code pdb : 3QAV) et Oméga (Board *et al.*, 2000 ; code pdb : 1EEM).

### A.I.3.c Classification selon la nature du résidu catalytique

Une analyse récente basée sur l'alignement des séquences d'un grand nombre de GST cytosoliques propose une classification en deux grands groupes selon la nature du résidu catalytique (Atkinson & Babbitt, 2009). Le premier groupe concerne les Y-GST qui utilisent une tyrosine pour activer le glutathion, le second groupe correspond aux S/C-GST qui utilisent une sérine ou cystéine dans leur mécanisme catalytique. Suivant le groupe, la position du résidu catalytique dans les structures primaires est légèrement différente. La position de la tyrosine précède de quelques résidus la position de la sérine ou la cystéine. Les Y-GST sont des enzymes appartenant principalement aux classes spécifiques des GST de mammifères (Alpha, Mu et Pi) et à la classe Sigma commune à de nombreux organismes. Par contre les S/C-GST sont plus largement distribuées.

Les GST possédant soit une sérine soit une tyrosine comme résidu catalytique présentent majoritairement une activité GSH transférase (Tableau 1). Les structures cristallographiques montrent que le GSH est stabilisé par le groupe hydroxyle en formant une liaison hydrogène avec

son groupe sulfhydryl (-SH). Par ailleurs, il a été montré que la tyrosine abaisse le pKa du groupe sulfure jusqu'à 2,4 unité de pH (Kong *et al.*, 1992). Ceci permet la formation de l'anion thiolate (GS<sup>-</sup>) qui est 10<sup>9</sup> fois plus réactif que le GSH (Armstrong, 1991). Ce mécanisme catalytique a été bien décrit dans la revue du professeur R. Armstrong (1997)

Les GST possédant une cystéine comme résidu catalytique présentent à l'inverse une activité thiol transférase en formant un pont disulfure avec le glutathion. Ces GST se comportent comme la glutarédoxine dont elles réduisent les substrats classiques en présence de glutathion (Oakley, 2011).

#### **A.I.4 La structure tridimensionnelle des GST**

##### **A.I.4.a Les GST cytosoliques**

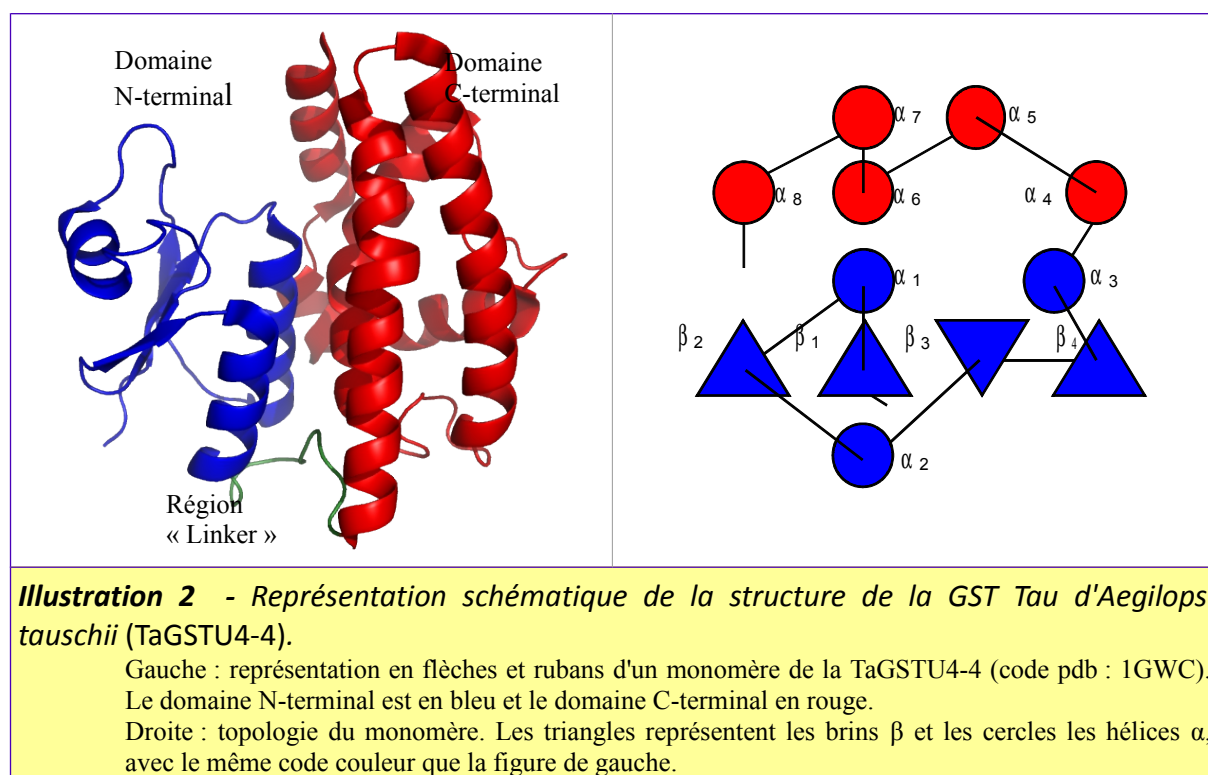
###### ***i* Le repliement tridimensionnel du monomère**

Malgré le faible taux d'identité de séquence entre les classes, tous les monomères des GST cytosoliques dont la structure est connue partagent la même structure tertiaire (Oakley, 2011), présentée dans l'illustration 2, ici avec la 4<sup>ème</sup> isoforme de la GST Tau d'*Aegilops tauschii*, notée TaGSTU4-4 (code pdb 1GWC, Thom *et al.*, 2002). Le monomère se compose de deux domaines distincts. Le domaine N-terminal (comprenant approximativement les résidus de 1 à 80) adopte une topologie ( $\beta_1\alpha_1\beta_2\alpha_2\beta_3\beta_4\alpha_3$ ) semblable à celle des thiorédoxines (Murzin *et al.*, 1995). Le domaine C-terminal (approximativement du résidu 90 au dernier) est composé d'un paquet minimal de 4 hélices ( $\alpha_4\alpha_5\alpha_6\alpha_7$ ) additionné d'éventuelles hélices supplémentaires de position variable. Les deux domaines sont reliés par une courte région d'une dizaine de résidus sans structure secondaire caractéristique.

L'observation d'un repliement thiorédoxine dans le domaine N-terminal concorde avec sa présence dans plusieurs autres familles d'enzymes (avec une identité de séquence limitée) qui partagent la capacité de réagir sur une cystéine ou le GSH (Sheehan *et al.*, 2001). Les exemples sont la DsbA (l'enzyme bactérienne équivalente à la protéine isomérase disulfure), la glutarédoxine et la glutathion peroxydase.

Le repliement thiorédoxine du domaine N-terminal des GST se compose de deux sous-motifs distincts suivant la topologie  $\beta_1\alpha_1\beta_2$  et  $\beta_3\beta_4\alpha_3$ . Ils sont reliés par l'hélice  $\alpha_2$ . Les quatre brins bêta forment un feuillet avec le brin  $\beta_3$  antiparallèle aux trois autres brins. Les deux hélices  $\alpha_1$  et  $\alpha_3$  sont parallèles et se situent d'un côté du feuillet et l'hélice  $\alpha_2$  de l'autre côté, faisant face au solvant. La boucle qui relie  $\alpha_2$  à  $\beta_3$  comporte un résidu proline caractéristique qui est dans une conformation *cis* et qui est invariable dans les GST cytosoliques (correspondant à Pro57 chez TaGSTU4-4). Cette boucle, connue sous le nom de boucle *cis*-Proline, n'a pas de rôle direct dans la catalyse. Néanmoins, elle semble être importante en maintenant la protéine dans une structure

catalytiquement compétente (Allocati *et al.*, 1999). Le domaine N-terminal fournit la plupart des résidus du site de fixation du GSH et il contient également le résidu catalytique. Dans les Y-GST la tyrosine catalytique se trouve au niveau C-terminal du brin  $\beta_1$  alors que la sérine ou la cystéine des S/C-GST occupe la dernière position de la boucle  $\beta_1\alpha_1$ . Toutes ces propriétés structurales, communes à quasiment toutes les GST, expliquent probablement pourquoi le domaine N-terminal est mieux conservé dans les GST que le domaine tout hélice.



Le domaine C-terminal, plus variable, est formé d'un paquet de quatre à huit hélices  $\alpha$ , ce nombre variant d'une classe à l'autre. Le domaine C-terminal a un repliement unique dont le cœur est composé des quatre hélices consécutives  $\alpha_4$ ,  $\alpha_5$ ,  $\alpha_6$  et  $\alpha_7$ . Les hélices  $\alpha_4$ ,  $\alpha_5$ ,  $\alpha_6$  sont orientées approximativement suivant la même direction que les hélices  $\alpha_1$  et  $\alpha_3$  alors que la dernière hélice  $\alpha_7$  est quasiment perpendiculaire aux trois autres. Quand des hélices supplémentaires sont présentes, elles se trouvent le plus souvent à la suite de l'hélice  $\alpha_7$ . Le domaine C-terminal contribue à fournir la plupart des résidus formant le site de fixation du substrat électrophile qui est le plus souvent de nature hydrophobe. Ce site de fixation a ainsi été appelé site H. Les différences observées dans le domaine C-terminal entre les différentes classes de GST sont supposées être responsables de la spécificité de substrat de chacune d'elles (Armstrong, 1997). Un résidu, néanmoins, est très bien conservé dans ce domaine ; il s'agit d'un aspartate (Asp162 dans la TaGSTU4-4) situé dans le premier tour de l'hélice  $\alpha_6$  qui participe à un motif « *N-capping Box* » du type S/TXXD. Ce motif permet, via des chaînes latérales, de stabiliser dans une hélice  $\alpha$  deux des quatre premiers



groupements NH qui sont sans accepteur de liaison hydrogène (Richardson & Richardson, 1988). Ainsi le carboxylate de cet acide aspartique stabilise le groupement NH du premier résidu du motif. Ce motif semble crucial pour le repliement et la stabilité des GST (Dragani *et al.*, 1997).

Les contacts à l'interface des domaines font intervenir dans la plupart des cas, des résidus des hélices  $\alpha_1$  et  $\alpha_3$  du domaine N-terminal et des résidus des hélices  $\alpha_6$  et  $\alpha_4$  du domaine C-terminal.

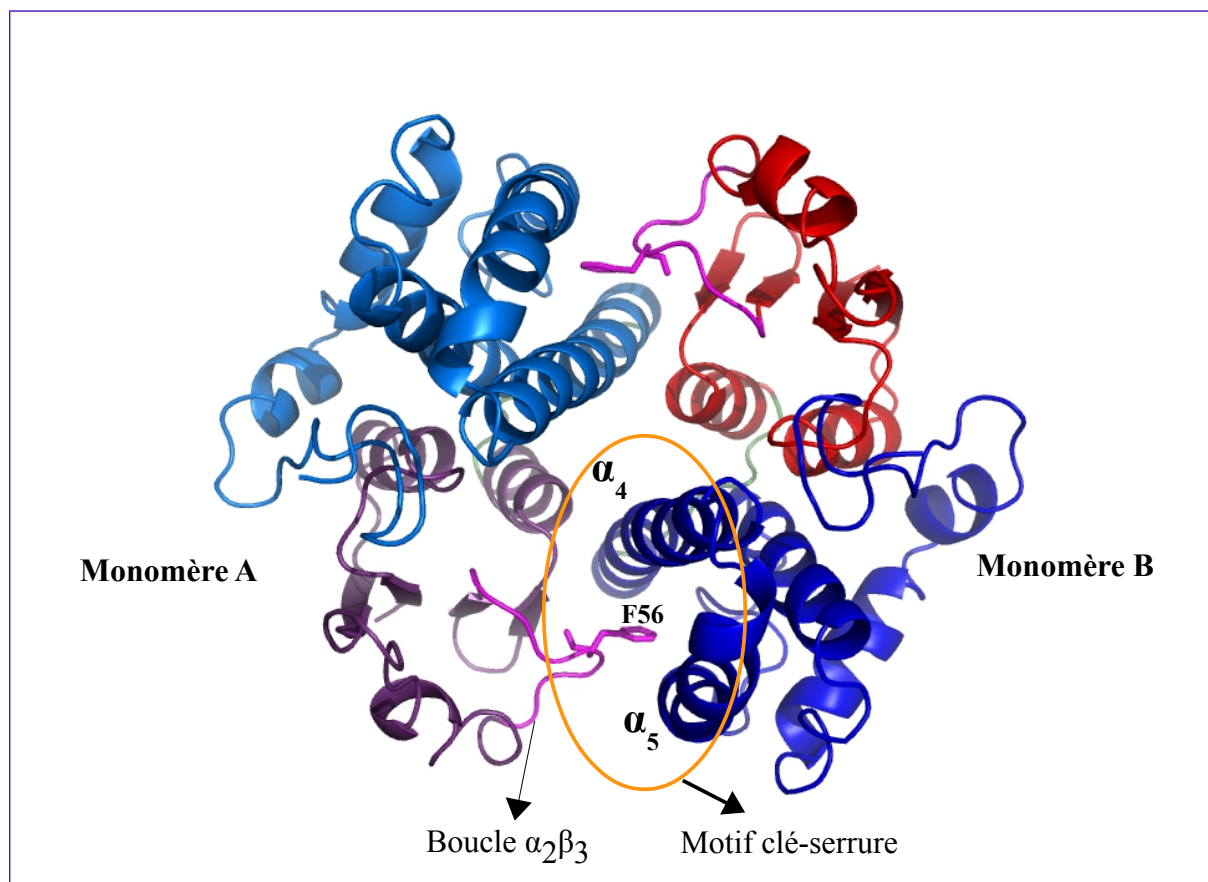
En dépit de leurs topologies globales similaires, les structures des GST présentent des différences significatives localement en particulier au niveau du site de fixation du substrat électrophile (site H). De plus, les différences topologiques les plus remarquables sont : la boucle dite « Mu » qui est une longue boucle caractéristique des GST Mu située entre le brin  $\beta_2$  et l'hélice  $\alpha_2$ ; l'hélice  $\alpha_9$  des classes Alpha et Thêta qui ferment le site H en le rendant plus hydrophobe (Sinning *et al.*, 1993); l'extension N-terminale et l'hélice  $\alpha_{10}$  se rejoignant dans la classe Oméga (Board *et al.*, 2000).

## ***ii Structures quaternaires***

Les GST sont pour la plupart des homo-dimères globulaires d'environ 50 kDa et de rayon approximatif 50 Å. La surface enfouie à l'interface du dimère varie de façon significative d'une classe à l'autre, d'environ 2000 Å<sup>2</sup> à 3400 Å<sup>2</sup> (Board *et al.*, 2000). La forme dimérique semble conférer à l'enzyme une certaine stabilité. Concernant la GST Pi, il a été montré que les deux sites actifs fonctionnent en synergie (Hegazy *et al.*, 2004) et la forme monomérique est active (Huang *et al.*, 2008). Il existe aussi des hétérodimères de GST mais uniquement au sein d'une même classe (Dirr *et al.*, 1994). Les deux sous-unités (ou monomères) sont reliées par une symétrie axiale binaire. Le domaine N-terminal d'une sous-unité interagit avec le domaine C-terminal de l'autre sous-unité. Ainsi la boucle  $\alpha_2\beta_3$ , le brin  $\beta_4$  et l'hélice  $\alpha_3$  d'une sous-unité interagissent avec les hélices  $\alpha_4$  et  $\alpha_5$  de l'autre sous-unité (voir Illustration 3). Ce dimère est considéré comme le dimère classique des GST. La suite de ce manuscrit montrera toutefois qu'il peut aussi exister d'autres types de dimères de GST. Par ailleurs, il existe également des GST cytosoliques monomériques, ce sont les GST Lamda, les DHAR et les CLIC (Frova, 2006).

Dans le dimère canonique, l'analyse des interfaces entre les sous-unités des enzymes de mammifères Alpha, Mu et Pi montre qu'elles partagent une interaction hydrophobe spécifique appelée « *ball-and-socket* » ou « *lock-and-key* » (« clé-serrure ») (Dirr *et al.*, 1994). Dans cette interaction la chaîne latérale d'un résidu phénylalanine qui se trouve sur la boucle  $\alpha_2\beta_3$  joue le rôle de la clé (F56 dans la première isoforme découverte de GST Mu de rat, rGSTM1-1, code PDB 5GST) (Illustration 3). Celle-ci s'insère presque parfaitement dans une crevasse (serrure) située entre les hélices  $\alpha_4$  et  $\alpha_5$  du domaine C-terminal de l'autre monomère. Cette interaction particulière n'est pas présente dans les GST des classes Thêta et Sigma dont l'interface des sous-unités est

caractérisée majoritairement par des interactions hydrophiles (Armstrong, 1997). En effet, le résidu phénylalanine en question et la poche hydrophobe sont absents dans les structures des classes Thêta et Sigma (Armstrong, 1997).



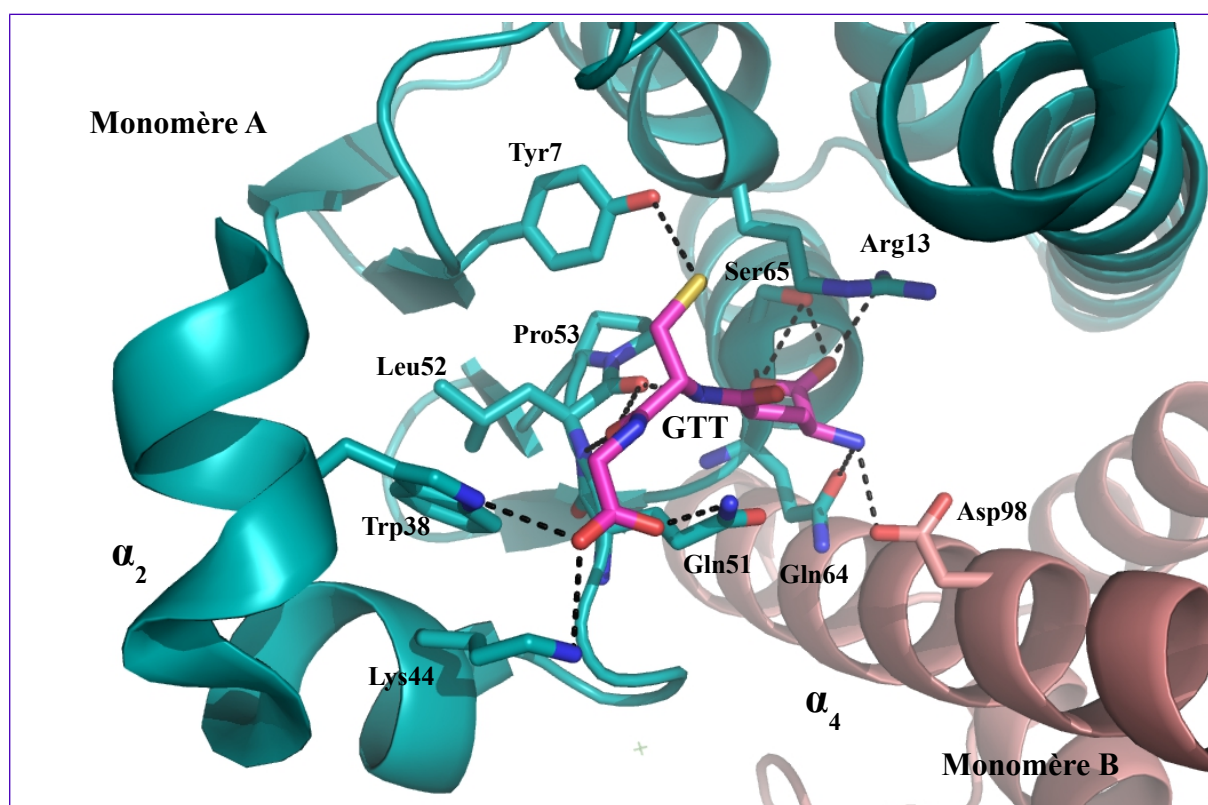
**Illustration 3** - Représentation schématique du dimère classique des GST cytosoliques.

Représentation en flèches et rubans d'un dimère de la GST mu de rat (rGSTM1-1, code pdb 5GST). Dans le monomère A, le domaine N-terminal est en violet et le domaine C-terminal en bleu clair. Dans le monomère B, le domaine N-terminal est en rouge et le domaine C-terminal en bleu foncé. Un des deux motifs « clé-serrure » permettant la stabilisation du dimère est mis en valeur dans l'ellipse jaune. La clé, à savoir la phénylalanine 56 provenant de la boucle  $\alpha_2\beta_3$ , représentée en rose (monomère A), s'insère dans une poche hydrophobe située entre les hélices  $\alpha_4$  et  $\alpha_5$  du domaine C-terminal de l'autre sous-unité (monomère B).

### iii Le site de fixation du GSH : Site G

Le site de fixation du glutathion appelé site G est situé en majeure partie dans le domaine N-terminal où le glutathion est fixé dans une conformation étendue. Le résidu  $\gamma$ -glutamate pointe vers l'interface du dimère, la chaîne latérale de la cystéine est localisée à proximité du résidu catalytique et le groupement carboxylate de la glycine se trouve à la surface de la protéine. Six des huit groupements donneurs ou accepteurs de liaisons hydrogène du glutathion forment quasiment toujours des interactions directes avec des résidus du site G. Seuls les groupements carboxylate du  $\gamma$ -glutamate et amine de la glycine ne forment que très rarement des liaisons hydrogène directes.

Les groupements  $\text{NH}_3^+$  et  $\text{COO}^-$  du  $\gamma$ -Glu sont le plus souvent stabilisés par un motif Q/ES/T (par exemple  $^{64}\text{QS}^{65}$  dans la GST Pi humaine, notée hGSTP1-1, représentée dans l'illustration 4) situé au niveau de la courte boucle  $\beta_4\alpha_3$ . Le résidu Q/E présente des angles de torsion anormaux dans le squelette de la chaîne principale indiquant une contrainte conformationnelle liée à son interaction avec le glutathion (Dirr *et al.*, 1994). Des interactions quaternaires avec le groupement amine N-terminal du glutathion sont parfois observées. Par exemple dans les GST de mammifères Pi, Mu et Alpha, ce groupement est stabilisé par la chaîne latérale d'un résidu aspartate provenant de l'hélice  $\alpha_4$  de l'autre sous-unité (D105 dans la rGSTM1-1 ; Dirr *et al.*, 1994).



**Illustration 4** - Représentation schématique du site G de la GST Pi humaine (hGSTP1-1).

Représentation en flèches et rubans du site G de la hGSTP1-1 (code pdb 1PX7). Les monomères A et B sont respectivement en vert et en saumon. Le glutathion est en rose. Le glutathion et des résidus impliqués dans sa stabilisation sont représentés en bâtons. Les liaisons hydrogène sont représentées par des pointillés.

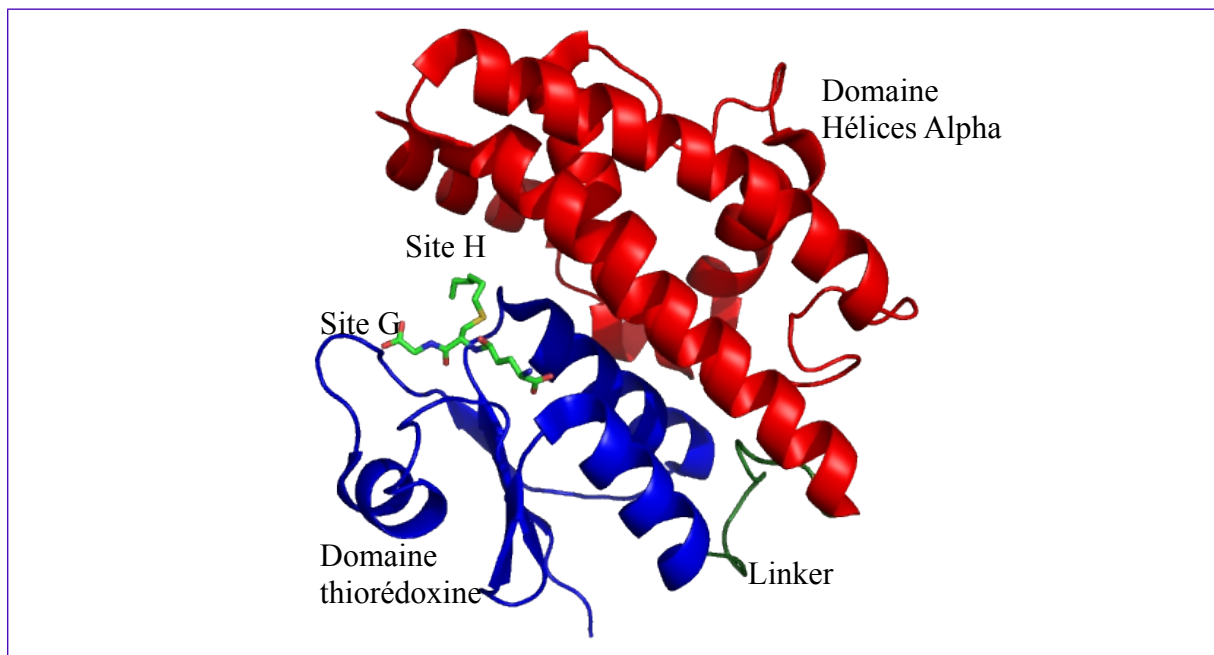
Les chaînes principales de la cystéine du glutathion et du résidu précédant la *cis*-proline forment des liaisons hydrogène complémentaires ( $^{52}\text{LP}^{53}$  dans la hGSTP1-1). Le groupement thiol du glutathion forme une liaison hydrogène avec le résidu catalytique s'il s'agit d'une sérine ou d'une tyrosine (Y7 dans la hGSTP1-1). Sinon le glutathion est lié de manière covalente à la cystéine catalytique.

Le groupement carboxylate terminal du glutathion est le plus souvent stabilisé par la chaîne latérale d'un résidu basique provenant de la région entre les brins  $\beta_2$  et  $\beta_3$ . Par exemple dans la TaGSTU4-4, il s'agit d'une lysine (K42) situé dans la boucle  $\beta_2\alpha_2$  (Thom *et al.*, 2002). Dans la GST Phi d'*Arabidopsis thaliana* (AtGSTF1-1), c'est également une lysine, mais celle-ci est au cœur de l'hélice  $\alpha_2$ . (K41, Reinemer *et al.*, 1996). Même si ces résidus basiques ne sont pas alignés dans les séquences, ils occupent souvent la même position dans les structures (Dirr *et al.*, 1994).

L'analyse fine des interactions entre le glutathion et la GST Delta d'*Anopheles dirus* (AdGSTD3-3) a permis à Winayanuwattikun et Kitterman (2004) de proposer que le carboxylate du  $\gamma$ -Glu joue le rôle d'une base pendant la catalyse. Assisté des résidus chargés environnants, il est supposé déprotoner le groupement thiol de la cystéine afin d'obtenir la forme la plus réactive du glutathion. Les auteurs parlent alors de la formation d'un réseau partagé d'électrons.

#### ***iv Le site de fixation des substrats hydrophobes : Site H***

La localisation du site de fixation des substrats hydrophobes (site H) a été décrite pour la première fois dans la structure cristallographique de la GST Pi humaine (hGSTP1-1), qui a été co-cristallisée avec le *S*-hexylglutathion (Reinemer *et al.*, 1992). Le site H, qui est adjacent au site G, est localisé dans une crevasse entre les domaines N-terminal et C-terminal. Les résidus des deux domaines contribuant à la formation de ce site H varient largement entre les classes, de même que la position précise du site lui-même. Ces différences semblent être liées au fait que les enzymes de la superfamille des GST catalysent plusieurs réactions distinctes avec des substrats très variables. Ainsi chaque isoenzyme dispose d'un site H spécifique afin d'optimiser son efficacité catalytique et en particulier son ou ses états de transition (Armstrong, 1997) (Illustration 5). Malgré tout, les structures tridimensionnelles d'un certain nombre d'enzymes en complexe avec leurs substrats et leurs analogues d'état de transition ont mené à l'identification de trois modes différents de fixation pour les peptides conjugués (GS-R), qui ont été appelés « *in* », « *out* », ou « *far-out* » (Armstrong, 1997). Dans les modes « *in* » ou « *out* », la partie peptidyl de la molécule GS-R occupe le site G. Seule diffère la position du groupement R. Dans le mode de fixation « *in* », il occupe une cavité dont le fond est défini par les résidus du domaine N-terminal appartenant à la boucle  $\alpha_1\beta_1$ . Les parois de cette cavité sont formées par une face de l'hélice  $\alpha_4$  et la queue C-terminale (Illustration 6). Cette cavité est supposée être le site du substrat xénobiotique catalytiquement productif (Armstrong, 1997). Dans le mode de fixation « *out* », le groupement R est dans une conformation s'éloignant du site actif en s'orientant vers le solvant entre les deux sous-unités du dimère. Ce mode de fixation est communément observé avec le *S*-(2,4-dinitrophenyl)-SG (Illustration 6).



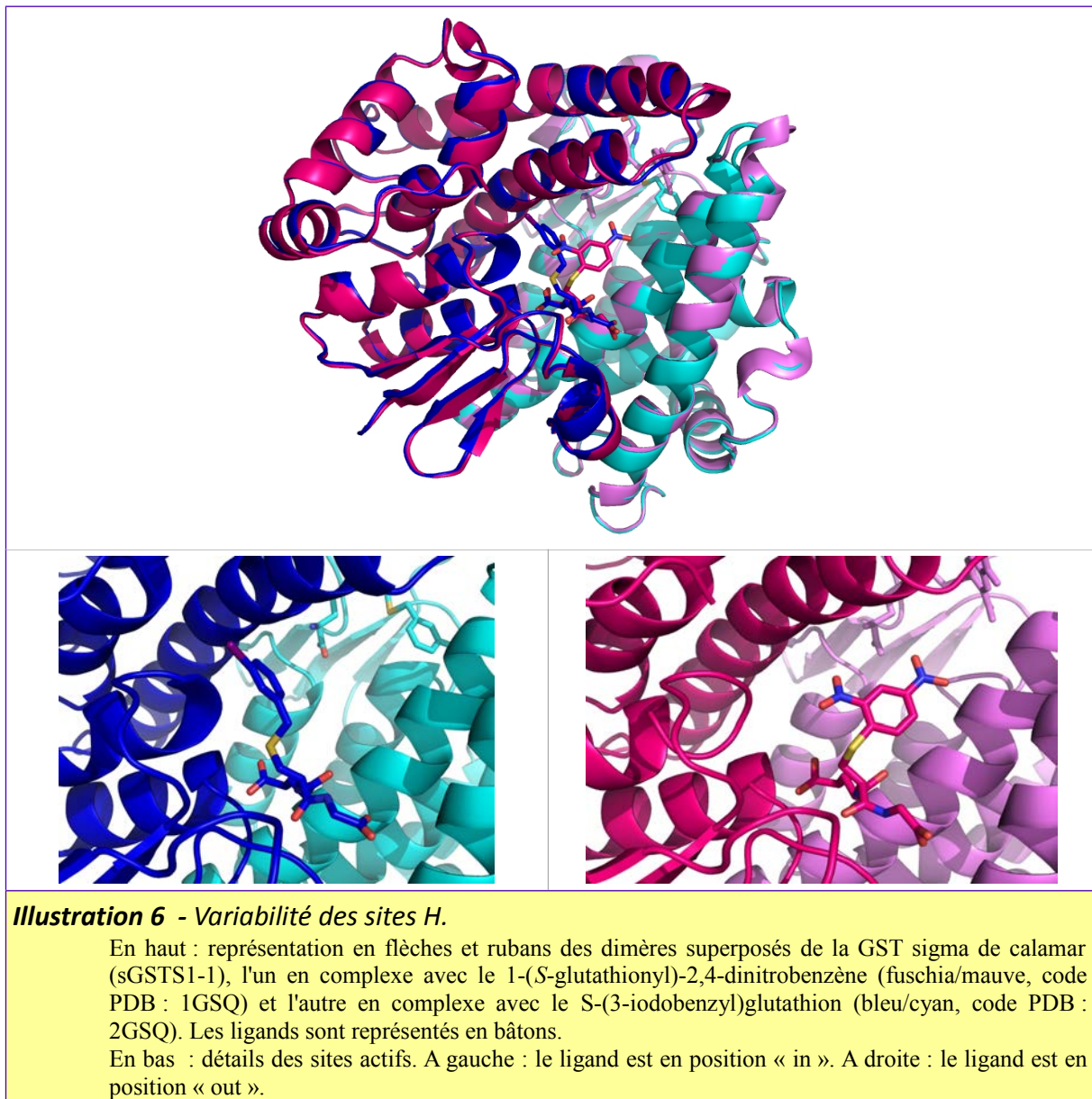
**Illustration 5** - Représentation schématique d'un monomère de GST en complexe avec un substrat glutathionylé.

Représentation en flèches et rubans d'un monomère de la GST Tau d'*Aegilops tauschii* (TaGSTU4-4, code pdb : 1GWC). Le domaine thiorédoxine est en bleu, le domaine tout hélice en rouge. Le ligand est le S-hexylglutathion, représenté en bâtons verts. Le groupement hexyl se trouve dans le site H alors que le tripeptide se trouve dans le site G.

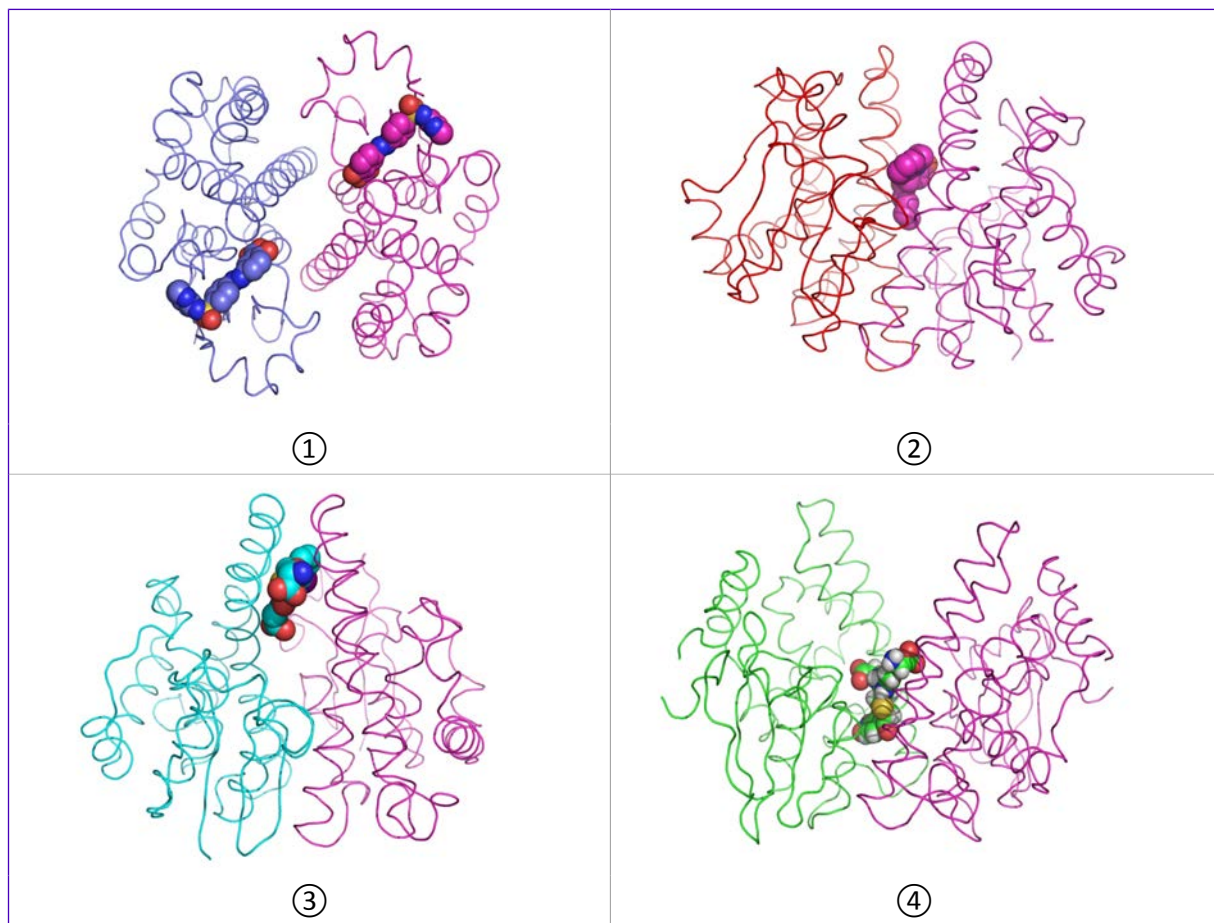
Dans le mode de fixation « *far-out* » la partie peptidyl n'occupe pas le site G et le groupe R n'occupe également pas le site de fixation du substrat xénobiotique. Cela a été observé notamment dans la structure de la GST Sigma de calamar où trois molécules de S-(3-iodobenzyl)-SG sont fixées : une dans chaque site actif du dimère et la troisième à l'interface du dimère (Ji *et al.*, 1995). Ce dernier site très éloigné du site catalytique de l'enzyme a plutôt été classé par la suite comme site L pour site de fixation de ligands, en lien avec des activités de transport des GST (voir ci-dessous).

#### v Les autres sites de fixation de ligands : sites L

Dès le début des années 70, les GST ont été qualifiées de « ligandine » car il a été montré qu'elles ont la capacité d'être impliquées dans le stockage et le transport rapide de molécules dans la phase aqueuse de la cellule (Litwack *et al.*, 1971 ; Habig *et al.*, 1974 ; Tipping & Ketterer, 1981 ; Caccuri *et al.*, 1990). Les caractérisations cristallographiques des GST dans les années 90 ont mis en évidence des sites de fixation de ligands différents des sites G et H appelés sites L. Cela dit, il n'est pas rare de retrouver un ligand se fixant en partie dans le site H, tel que la sulfasalazine qui se fixe dans le site H de la GST Pi humaine (hGSTP1-1, Oakley *et al.*, 1999) et qui est connue pour être un inhibiteur des GST des classes Alpha, Mu et Pi (Illustration 7).



Plusieurs études ont mis en évidence un site de fixation de molécules de tampon dans la hGSTP1-1 (Ji *et al.*, 1997 ; Oakley *et al.*, 1997, Prade *et al.*, 1997). Ce site, adjacent au tryptophane 28 et localisé au niveau N-terminal du brin  $\beta_1$ , a la capacité de fixer les molécules des tampons MES (acide 2-(N-morpholino)-éthanesulfonique) et HEPES (acide 4-(2-hydroxyéthyl)-1-piperazine éthane sulfonique). Par ailleurs, un site à l'interface des dimères a été caractérisé dans plusieurs classes (Mu : McTigue *et al.*, 1995 ; Sigma : Ji *et al.*, 1996 ; Oméga : Brock *et al.*, 2013). Dans ce cas il n'y a qu'un site par dimère, dont la position exacte diffère d'une classe à l'autre (Illustration 7). Il a été supposé que ce site L à l'interface des dimères pouvait être le site de fixation des inhibiteurs des GST Oméga humaines (Brock *et al.*, 2013). Dans sa revue récente, A. Oakley (2011) souligne le fait qu'il n'est pas possible de généraliser la position d'un site L observé dans une classe à toutes les classes.



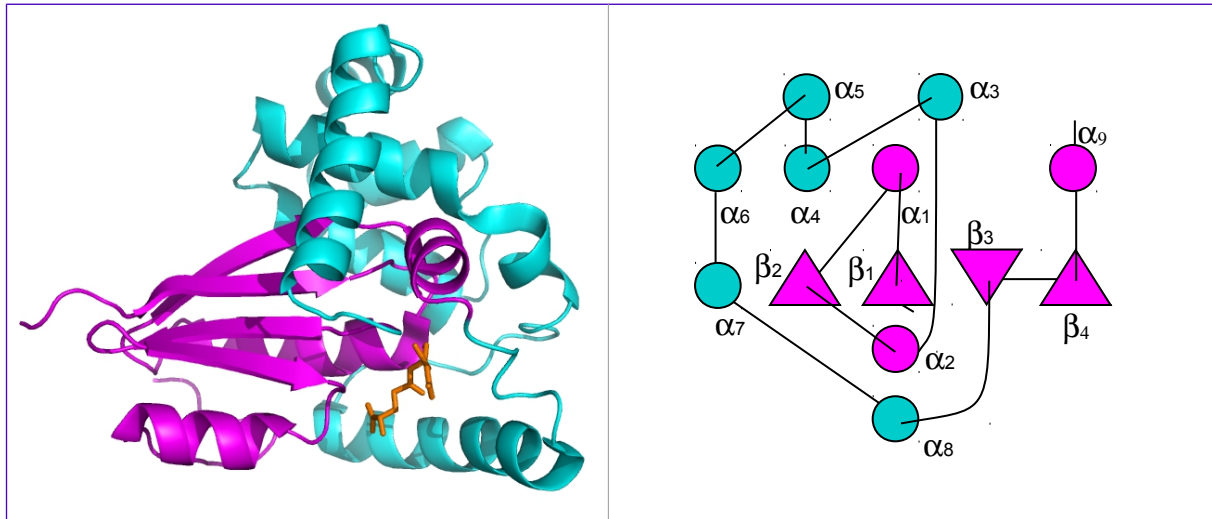
**Illustration 7** - Représentation schématique des sites L de quelques GST.

① : dimère de la hGSTP1-1 en complexe avec la sulfasalazine représentée en sphères (code pdb : 13GS). ② : dimère de la GST de classe mu de *S. japonica* en complexe avec le praziquantel représenté en sphères (code pdb : 1GTB); ③ : dimère de la GST de classe sigma de calamar en complexe avec le S-(3-iodobenzyl)glutathion représenté en sphères (code PDB 2GSQ); ④ : dimère de la hGSTO1-1 en complexe avec le S-(4-nitrophénacyl)glutathion représenté en sphères (code PDB : 4IS0).

#### A.I.4.b Les GST mitochondriales

Les GST mitochondriales sont rassemblées dans une classe unique appelée Kappa et ont été détectées dans les mitochondries et les peroxisomes (Pemble *et al.*, 1996). Les structures de cette classe montrent à la fois des différences et des similarités frappantes avec les GST cytosoliques (Ladner *et al.*, 2004). A ce jour, cinq structures de GST Kappa ont été résolues. Les enzymes de la classe Kappa possèdent un repliement particulier similaire à celui des DsbA. Celles-ci appartiennent à une famille d'enzymes impliquées dans la formation des ponts disulfures des protéines (DSB). Dans ce repliement particulier, le domaine tout hélice est inséré entre l'hélice  $\alpha_2$  et le brin  $\beta_3$  du domaine thiorédoxine (Illustration 8). Malgré ces différences, d'une part le mode de fixation du glutathion des GST Kappa est similaire à celui des GST cytosoliques, d'autre part la

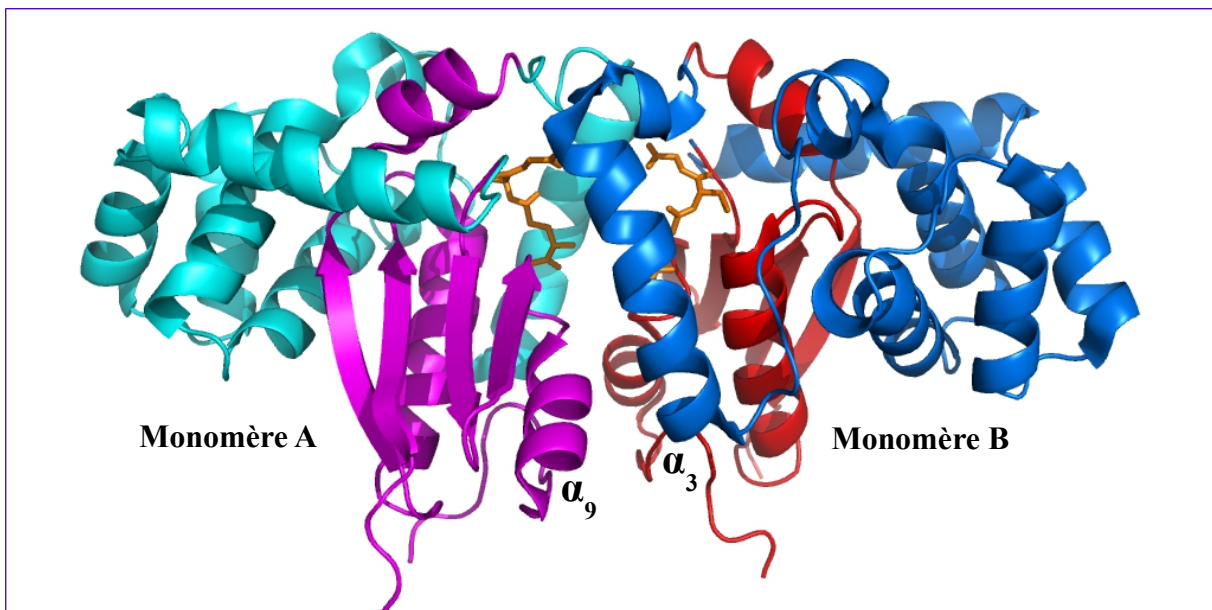
sérine catalytique des GST mitochondriales est localisée à la même position que le résidu catalytique des S/C-GST. La structure tridimensionnelle des GST Kappa est également proche de celle d'autres enzymes telles que la 2-hydroxychromène-2-carboxylate isomérase bactérienne qui est une oxydoréductase GSH-dépendante (Hayes *et al.*, 2005).



**Illustration 8** - Représentation schématique de la structure de la GST mitochondriale Kappa de rat (rGSTK1-1).

Gauche : représentation en flèches et rubans du monomère de la GST Kappa de rat (rGSTK1-1, code pdb 1R4W). Le domaine thiorédoxine est en magenta, le domaine tout hélice en cyan. Le glutathion est représenté en bâtons orange.

Droite : topologie d'un monomère. Les triangles représentent les brins  $\beta$  et les cercles représentent les hélices  $\alpha$ , avec le même code couleur.



**Illustration 9** - Représentation schématique d'un dimère de GST Kappa mitochondriale.

Le monomère de gauche est représenté avec le même code couleur que dans l'illustration 8, dans le monomère de droite, le domaine thiorédoxine est en rouge et le domaine tout hélice en bleu. Le glutathion est représenté en bâtons orange.

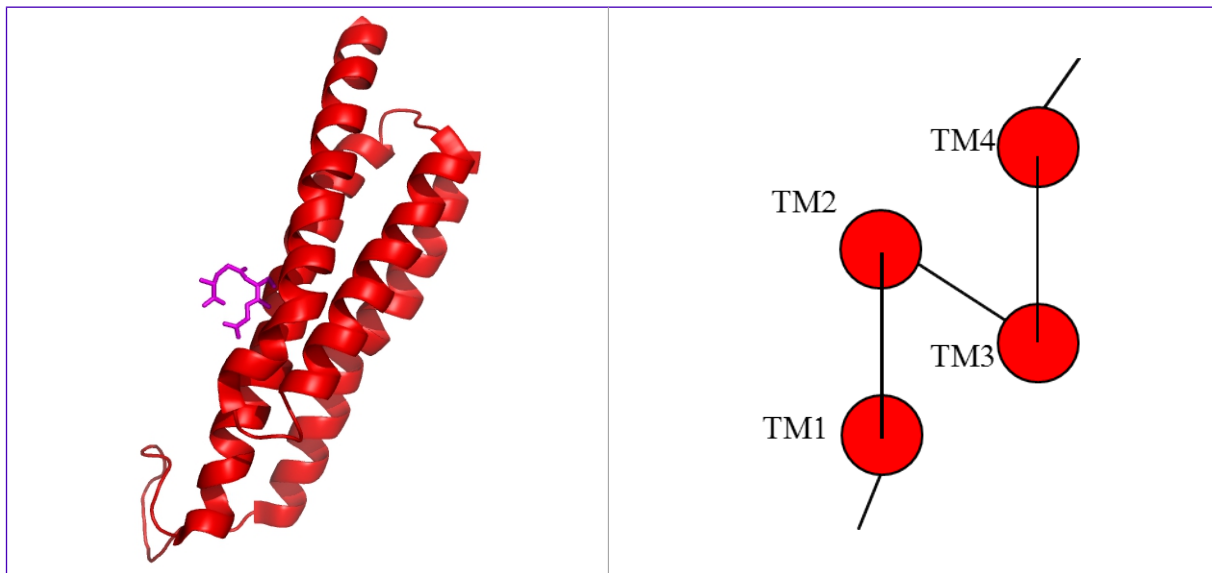


Le dimère des GST Kappa, de forme allongée, est très différent du dimère canonique des GST cytosoliques qui a une forme globulaire (Illustration 9). Les monomères sont en contact par l'intermédiaire de l'hélice  $\alpha_3$  (première hélice du domaine tout hélice des GST Kappa) d'un monomère qui s'assemble avec l'hélice  $\alpha_9$  de l'autre monomère. L'hélice  $\alpha_9$  correspond à la dernière hélice du domaine thiorédoxine. L'étendue de l'interface du dimère des GST Kappa est du même ordre de grandeur que celle des GST cytosoliques (env. 3000 Å<sup>2</sup>) (Ladner *et al.*, 2004).

#### **A.I.4.c Les GST MAPEG**

Les MAPEG (pour « Membrane-Associated Proteins in Eicosanoid and Glutathione metabolism ») constituent une branche unique de la superfamille des GST où la plupart sont impliquées dans la synthèse des eicosanoïdes. Elles sont divisées en 4 sous-groupes (I à IV) où les membres d'un sous-groupe partagent entre eux une identité de séquence supérieure à 20 % (Hayes *et al.*, 2005).

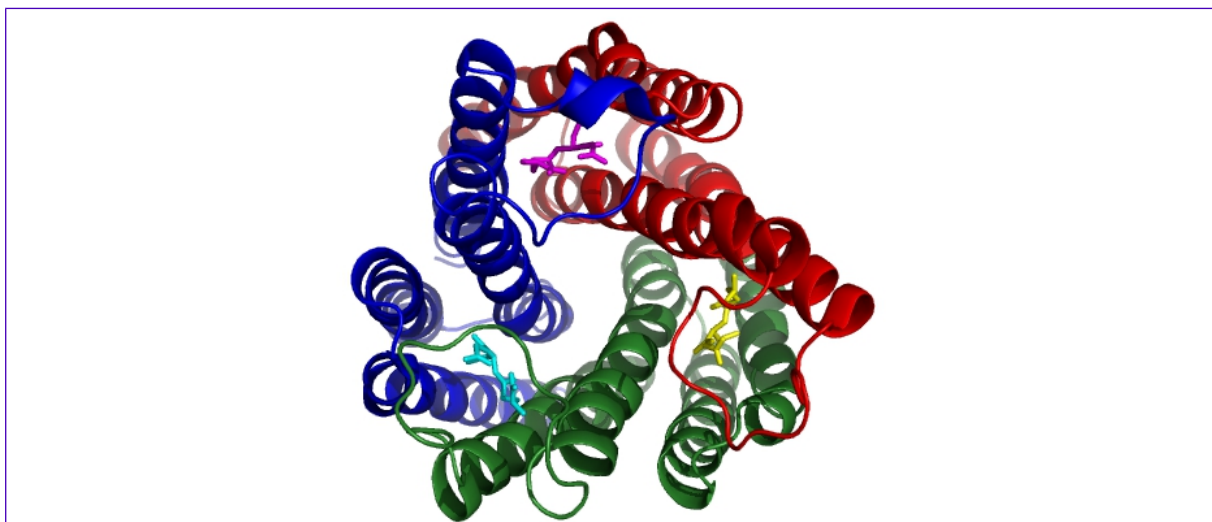
Malgré les faibles identités de séquence entre les sous-groupes, les MAPEG partagent une structure tridimensionnelle unique qui est totalement différente de la structure canonique des GST cytosoliques. Il est important de rappeler que les MAPEG ont été classées dans la superfamille des GST car la première d'entre elles à avoir été caractérisée présente une activité GSH transférase avec divers substrats tels que le CDNB (Morgenstern *et al.*, 1982). Les études cristallographiques ayant permis d'apporter des informations à haute résolution sur les MAPEG sont celles réalisées sur la MGST1 de rat (Holm *et al.*, 2006) et sur la LTC4S (LeukoTriene C<sub>4</sub> Synthase) humaine (Martinez Molina *et al.*, 2007), résolues par la diffraction des rayons X. La structure des MAPEG est formée d'un paquet de quatre hélices gauches (TM1-TM4) assemblées en un trimère (Illustration 10 et Illustration 11) où le paquet d'hélices est orienté orthogonalement au plan de la membrane. La topologie globale des structures tridimensionnelles des MGST1 et LTC4S est étroitement similaire à celle de la sous-unité 1 de la protéine ba<sub>3</sub>-cytochrome *c* oxydase (Oakley, 2011). Il y a trois sites de fixation du glutathion par trimère, situés du côté cytosolique. Contrairement aux GST cytosoliques et mitochondriales, les trois monomères des GST MAPEG fonctionnent ensemble et les sites G sont formés par la participation des résidus de deux monomères. Dans la structure de la MGST1 du rat, le glutathion se fixe dans une conformation étendue de la même manière que dans les GST cytosoliques et mitochondriales. Par contre, dans la LTC4S, le glutathion adopte une conformation repliée en V avec son groupement thiol dirigé vers la membrane pour interagir avec son substrat lié à la membrane. Pour finir, la structure de MAPEG FLAP (qui est protéine activatrice de la 5-lipoxygénase) a été obtenue en présence d'un inhibiteur qui se fixe dans une poche adjacente au site de fixation du GSH (Ferguson *et al.*, 2007).



**Illustration 10** - Représentation schématique de la structure d'un monomère de la prostaglandine E synthase 1 microsomale humaine (hMPGES1).

Gauche : représentation en rubans du monomère observé (en rouge), en complexe avec le glutathion (en bâtons magenta) (code PDB 3DWW).

Droite : topologie d'un monomère. Les cercles représentent les hélices  $\alpha$ .



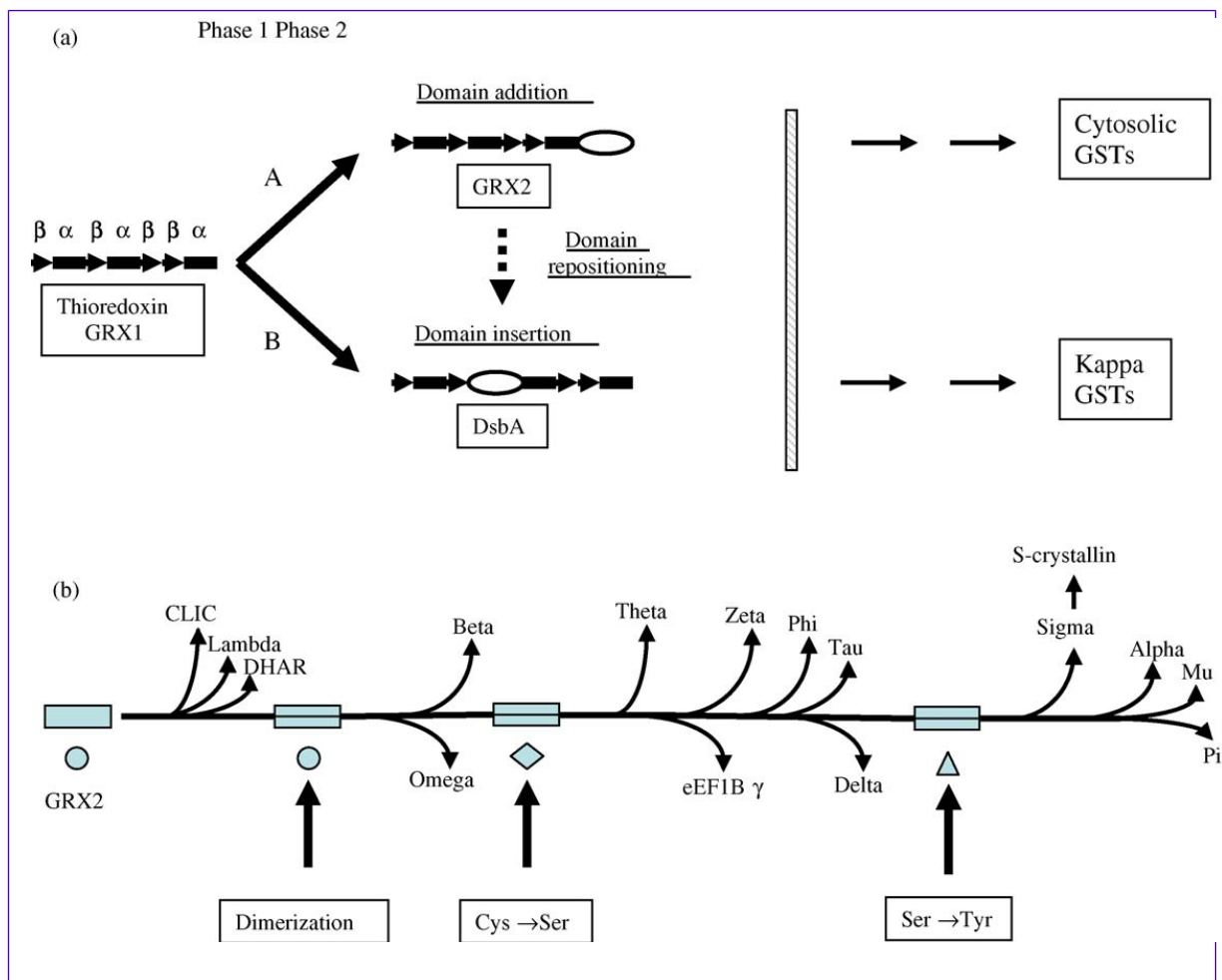
**Illustration 11** - Représentation schématique du trimère de la hMPGES1.

Représentation en rubans des trois monomères (en rouge, vert et bleu), en complexe avec leur glutathion (en bâtons magenta, jaune et cyan) (code pdb 3DWW).

### A.I.5 Évolution des GST

A la fin des années 1990 un modèle d'évolution des GST a été proposé où il a été supposé que l'apparition des diverses classes de GST provient d'un processus d'amplification de gènes suivie d'une évolution divergente conduisant à la création de nouvelles activités catalytiques (Armstrong, 1998 ; Hansson *et al.*, 1999). Ce modèle a été affiné dans les années 2000 par divers auteurs (Sheehan *et al.*, 2001 ; Hayes *et al.*, 2005 ; Frova, 2006 ; Öztetik, 2008). Avant tout, ce modèle

(Illustration 12) est basé sur la similarité du domaine N-terminal des GST cytosoliques, dont le repliement canonique semble avoir évolué d'un ancêtre commun thiorédoxine/glutarédoxine. L'évolution des enzymes cytosoliques commence par l'addition d'un domaine tout hélice au domaine thiorédoxine  $\beta\alpha\beta\alpha\beta\alpha$ . Par contre, la résolution de la première structure cristalline d'une GST Kappa ( $\tau$ GSTK1-1, Ladner *et al.*, 2004) indique clairement que les GST mitochondriales ont évolué suivant une voie parallèle, voire différente. En effet, le domaine tout hélice des GST Kappa, responsable de la fixation du substrat électrophile, est inséré dans le cœur du repliement  $\beta\alpha\beta\alpha\beta\alpha$  après l'hélice  $\alpha_2$ . Il est intéressant de noter que l'organisation des structures secondaires des GST Kappa ressemble à celle la DsbA qui est considérée comme un parent éloigné des thiorédoxines d'un point de vue de l'évolution (Shouldice *et al.*, 2011)



**Illustration 12** - Modèle de l'évolution/diversification des GST tiré de Frova *et al.* (2006).

(a) : les premières étapes à partir du repliement thiorédoxine.

(b) : séquence probable des étapes majeures de la différenciation des GST cytosoliques. Les évènements majeurs (dimérisation, changement de résidu catalytique) sont pointés par des flèches verticales.

La glutarédoxine 2 (GRX2) d'*E. coli* dont la première structure a été déterminée en 2000 semble être l'ancêtre commun des GST cytosoliques car elle possède en plus du domaine glutarédoxine classique (repliement du type thiorédoxine) un domaine C-terminal tout hélice similaire à celui des GST cytosoliques (Xia *et al.*, 2001). Les premières classes de GST à apparaître dans ce modèle d'évolution sont les CLIC, les DHAR et les GST Lambda car elles conservent la structure quaternaire monomérique de la GRX2. Ensuite arrive une étape de dimérisation avec l'apparition des classes Oméga et Bêta. Toutes ces classes conservent la cystéine catalytique de la GRX2. De plus, à l'exception des CLIC qui sont spécialisées dans la formation des canaux ioniques, ces enzymes possèdent une activité thiol transférase et/ou déshydroascorbate réductase comme les glutarédoxines. Ainsi, elles forment un groupe fonctionnel distinct des GST canoniques de conjugaison. Toutefois les GST Bêta agissent en tant que thiol-disulfure oxydoréductases et aussi en tant que glutathion transférases (Caccuri *et al.*, 2002). Il est supposé que leur évolution est postérieure à celle des GST Oméga, mais précède le second événement crucial dans l'évolution des GST qui est la transformation chimique du résidu catalytique en sérine.

Parmi les GST à sérine, les classes représentatives sont les GST Thêta et Zêta et le facteur d'élongation eEF1B $\gamma$  qui sont présents chez tous les eucaryotes, incluant les champignons et les nématodes. Dans ce groupe, la classe Thêta paraît être la plus ancienne car elle est également présente dans les bactéries. Malgré leur diversification fonctionnelle, les GST Zêta et eEF1B $\gamma$  sont similaires aux GST Thêta en séquence et en propriétés biochimiques, par exemple elles possèdent toutes l'activité de conjugaison avec le CDNB. Les GST Zêta et eEF1B $\gamma$  semblent ensuite avoir pris des chemins divergents, chez les eucaryotes (Frova, 2006).

Plusieurs indices indiquent que les autres GST à sérine sont plus récentes dans l'évolution. Premièrement, ces classes sont spécifiques à des groupes phylogénétiques. Par exemple, les classes Phi et Tau sont spécifiques des plantes alors que les GST Delta sont spécifiques des insectes. Deuxièmement, dans chaque organisme, il existe un grand nombre de membres de ces classes et leurs clusters de gènes sont fréquemment trouvés dans des courtes régions des génomes. De plus ces gènes présentent souvent une grande similarité de séquences. Toutes ces observations semblent indiquer des événements récents de duplication.

Le changement du résidu catalytique de sérine en tyrosine marque une autre séparation évolutive des GST cytosoliques. Il y a en plus de la modification de la nature du résidu catalytique une modification de position. Alors que dans les GST à sérine le résidu catalytique se trouve à la fin de la boucle  $\beta_1\alpha_1$ , la tyrosine se trouve à la fin du brin  $\beta_1$ . Les GST de mammifères des classes Alpha, Mu, Pi sont des GST à tyrosine, ainsi que les GST Sigma. Cette séparation a été proposée sur la base d'alignements de séquences (Board *et al.*, 1997). La présence de GST Sigma à la fois chez les vertébrés et chez les invertébrés, suggère qu'elles ont divergé avant le groupe des GST

Alpha/Mu/Pi, en évoluant avec des propriétés fonctionnelles distinctes dans diverses lignées. Chez les céphalopodes, il existe deux GST Sigma fonctionnellement distinctes: une GST Sigma présente une bonne activité de conjugaison du glutathion, tandis que l'autre est une protéine abondante dans le cristallin et ne présente pas d'activité catalytique (Tomarev *et al.*, 1993). Cette GST S-cristalline semble avoir évolué d'une GST Sigma dupliquée et incorporant une insertion de 11 résidus entre les hélices  $\alpha_4$  et  $\alpha_5$  qui serait responsable du manque d'activité enzymatique (Sheehan *et al.*, 2001). Par contre les GST Sigma des vertébrés et des parasites multicellulaires ont acquis des fonctions spécialisées dans la synthèse de prostaglandines (Sheehan *et al.*, 2001).

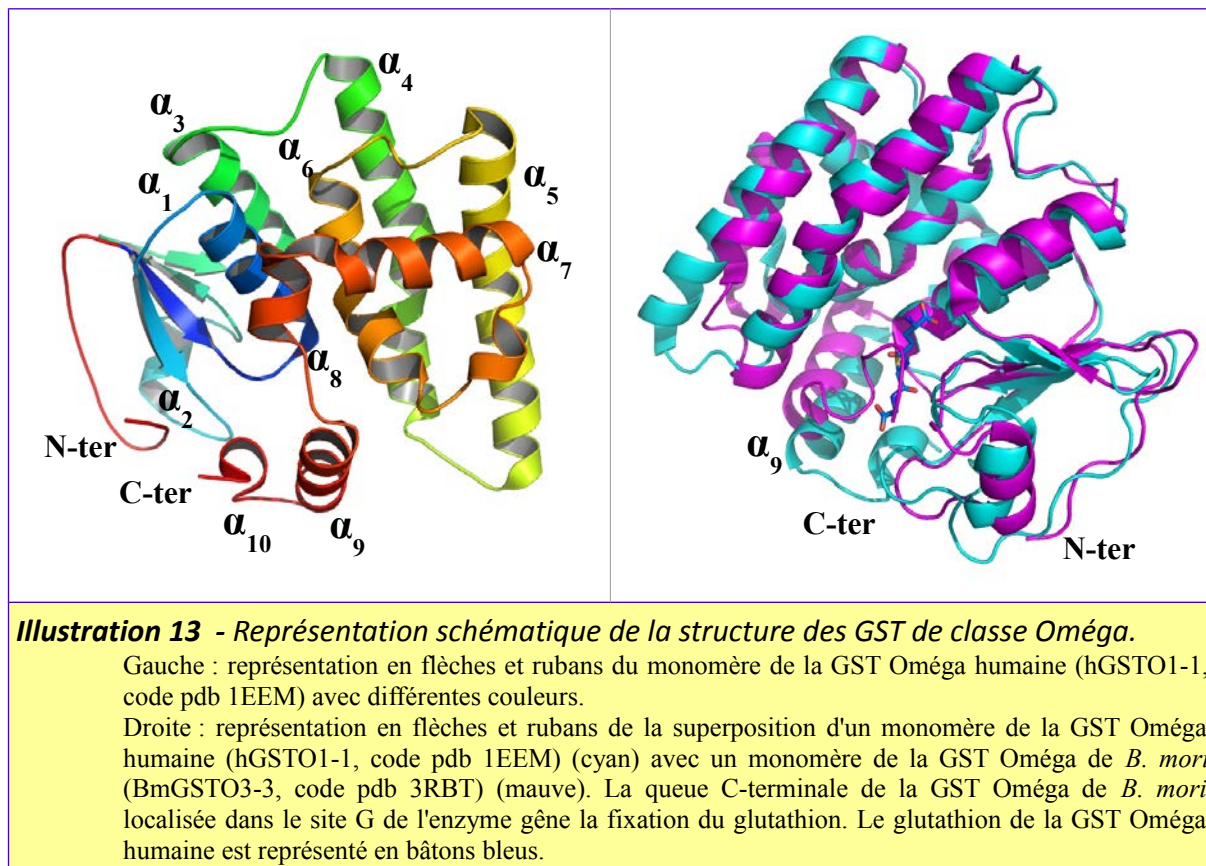
Ce modèle d'évolution devrait probablement devenir beaucoup complexe pour prendre en compte de nouvelles GST caractérisées ces dernières années, dans lesquels le résidu catalytique n'est ni la sérine ni la tyrosine classiques. Dans la GST Gtt2 de *S. cerevisiae*, Ma *et al.* (2009) ont découvert que la déprotonation du groupement thiol du glutathion d'un monomère est assisté, via une molécule d'eau, par une sérine et/ou histidine de l'hélice  $\alpha_4$  de l'autre monomère. Plus récemment, il a été montré que le résidu catalytique de la GST d'*Agrobacterium tumefaciens* est une arginine localisée dans la boucle  $\beta_1\alpha_1$  (Skopelitou *et al.*, 2012). Cette dernière a été définie dans une nouvelle classe appelée Êta.

## **A.II      Les glutathion transférases à cystéine**

### **A.II.1      Les GST Omega**

La classe Oméga est la première classe de GST à cystéine a avoir été caractérisée de façon approfondie (Board *et al.*, 2000). La cystéine catalytique est située à la fin de la boucle  $\beta_1\alpha_1$ , qui est la position de la cystéine catalytique des glutarédoxines et thiorédoxines. La classe des GST Oméga (GSTO) a été initialement identifiée et caractérisée chez l'homme (Board *et al.*, 2000), plus tard dans tout le règne animal (Whitbread *et al.*, 2005), et enfin plus récemment dans le règne fongique (Morel *et al.*, 2009). Chez l'homme et la souris, deux gènes de GSTO ont été découverts codant pour les protéines GSTO1-1 et GSTO2-2. Dans les deux organismes, l'ARNm de la GSTO1-1 est exprimé dans le foie, le cœur et d'autres tissus. Alors que pour la GSTO2-2, l'expression est prédominante dans les testicules (Whitbread *et al.*, 2005). La stabilité et la structure d'un variant de la GSTO1-1 humaine contenant une délétion polymorphique de la Glu155 ont été étudiées récemment (Zhou *et al.*, 2011). Ce polymorphisme a été identifié dans la plupart des groupes ethniques et sa fréquence peut atteindre une proportion supérieure à 10 %. Dans le ver à soie *Bombyx mori*, quatre gènes de GSTO ont été caractérisés. Parmi les quatre protéines, trois possèdent la même cystéine catalytique que les GSTO humaines et une possède une asparagine à la

place. La structure cristallographique de cette dernière (bmGSTO3-3 ) a été publiée récemment (Chen *et al.*, 2011).



La structure cristalline de la GST Oméga humaine GSTO1-1 a été la première à être déterminée (Board *et al.*, 2000). La hGSTO1-1 adopte comme attendu le repliement canonique des GST (Illustration 13). Néanmoins, elle présente plusieurs caractéristiques originales. Tout d'abord le dimère est relativement ouvert car la surface d'interface est d'environ 1000 Å<sup>2</sup> plus faible que celle observée dans la plupart des autres classes de GST (1960 Å<sup>2</sup> dans la hGSTO1-1 contre 2700 à 3400 Å<sup>2</sup> dans les autres). De plus, la hGSTO1-1 possède une élongation N-terminale d'une vingtaine de résidus qui contient un segment riche en proline, et trois hélices C-terminales caractéristiques  $\alpha_8$ ,  $\alpha_9$  et  $\alpha_{10}$ . L'élongation N-terminale présente la particularité de se trouver à proximité de la partie C-terminale pour créer un élément structural qui n'avait jamais été observé auparavant dans la famille des GST (Board *et al.*, 2000). La structure cristallographique révèle un pont disulfure entre la cystéine catalytique (C32) et le glutathion. Alors que la hGSTO1-1 présente un profil catalytique du type glutarédoxine, son activité GST de conjugaison du glutathion sur les substrats classiques n'est que résiduelle ou nulle. La mutation de la cystéine catalytique par une sérine augmente fortement l'activité de conjugaison. Board *et al.* (2000) proposent que la mutation de la cystéine en sérine,

empêchant la formation du pont disulfure, permet au GSH d'adopter une forme thiolate stable qui participe aisément aux réactions de conjugaison.

La hGSTO1-1 possède un site H putatif relativement grand et moins hydrophobe que ceux des autres GST. En effet la participation de plusieurs résidus polaires a été notée, tels que l'arginine 183 de l'hélice  $\alpha_6$  et située au fond de la poche. Cette particularité, ainsi que la grande ouverture entre les sous-unités, suggèrent que le substrat des GSTO n'a pas besoin d'être particulièrement hydrophobe et que sa taille peut être grande et donc de nature protéique par exemple.

La forme hGSTO1-1 $\Delta$ Glu155 de la hGSTO1-1 a été étudiée d'un point de vue biochimique et structurale (Zhou *et al.*, 2011). Ces analyses montrent que la délétion de Glu155 semble avoir plus d'impact sur la structure que sur les fonctions de l'enzyme. Dans la forme de type sauvage, Glu155 se trouve au niveau C-terminal de l'hélice  $\alpha_5$ . Sa délétion dans la forme hGSTO1-1 $\Delta$ Glu155 entraîne une interruption prématurée de l'hélice  $\alpha_5$  et une nouvelle conformation pour la boucle  $\alpha_5\alpha_6$ . Dans le cristal de la forme hGSTO1-1 $\Delta$ Glu155 trois monomères sont présents dans l'unité asymétrique. Des différences notables entre les monomères indépendants sont présentes au niveau des hélices  $\alpha_8$ ,  $\alpha_9$  et  $\alpha_{10}$  montrant une certaine variabilité au niveau du site H putatif de l'enzyme. En ce qui concerne les activités thiol transférase, déshydroascorbate et méthylarsonate réductases, dans tous les essais *in vitro*, la forme hGSTO1-1 $\Delta$ Glu155 est plus efficace que la forme de type sauvage. *In cellulo*, il avait été montré que la délétion du Glu155 diminue l'activité catalytique de 50 % par rapport à l'activité normale avec envers le 4-nitrophenacylglutathion qui est un substrat spécifique de la hGSTO1-1. La déficience de la hGSTO1-1 *in vivo*, s'explique par un manque de stabilité de la forme hGSTO1-1 $\Delta$ Glu155 dû aux différences structurales par rapport à l'enzyme de type sauvage (Zhou *et al.*, 2011).

L'enzyme hGSTO2-2 partage 64 % d'identité avec la hGSTO1-1, cependant les deux enzymes possèdent des substrats spécifiques différents. La hGSTO1-1 réagit plutôt avec de plus grands substrats que la hGSTO2-2. La hGSTO2-2 est la plus instable même si elle possède la meilleure activité déshydroascorbate réductase (DHAR). Plusieurs formes de l'enzyme hGSTO2-2 ont été étudiées structurellement : la forme de type sauvage et une forme dans laquelle les cystéines autres que la cystéine catalytique ont été mutées en sérines pour éviter la formation de ponts disulfures intermoléculaires (Zhou *et al.*, 2012). De plus une délétion d'une partie de l'extension C-terminale a été réalisée dans cette forme. La structure de la hGSTO2-2 est superposable à celle de la hGSTO1-1 où les hélices caractéristiques des GST Oméga  $\alpha_8$  et  $\alpha_9$  sont conservés. Les différences observées comprennent le remplacement de l'hélice  $\alpha_2$  qui participe à la formation du site G par une hélice  $3_{10}$  qui n'interagit pas avec le glutathion. Le plus surprenant est la conformation *cis* du glutamate 85 dans la forme de l'enzyme mutée avec les délétions induisant des changements structuraux locaux. Dans la forme de type sauvage, le lien peptidique a une conformation *trans* et la

chaîne latérale présente une double conformation. Quelle que soit la conformation, le glutamate 85 de la forme de type sauvage stabilise le groupe amine du  $\gamma$ -glutamyl du GSH. La mutation E85A diminue de 50 % l'activité DHAR de la hGSTO2-2 envers le déshydroascorbate. La liberté conformationnelle de ce glutamate est supposée nécessaire pour interagir avec le déshydroascorbate lors de sa réduction. En effet, cette étude révèle que l'acide ascorbique, le produit de la réduction du déshydroascorbate, se fixe de façon inattendue dans le site G de l'enzyme. De plus, la mutation Y34A élimine complètement l'activité DHAR. En effet la tyrosine 34 interagit avec l'acide ascorbique en formant des liaisons  $\pi$  aromatiques dans la structure du complexe (Zhou *et al.*, 2012).

L'analyse des séquences montre que BmGSTO3-3 partage 29 % d'identité avec les autres isoformes BmGSTO1-1, BmGSTO2-2, BmGSTO4-4 qui elles, partagent 40 % d'identité entre elles (Chen *et al.*, 2011). Par contre la BmGSTO3-3 partage jusqu'à 39 % d'identité avec les GST Oméga des autres insectes. Les études biochimiques de l'enzyme BmGSTO3-3 ont révélé une faible activité catalytique due à l'absence de la cystéine catalytique remplacée par une asparagine (N29). L'analyse tridimensionnelle de la BmGSTO3-3 montre qu'elle présente la même topologie structurale que les autres GST Oméga. Néanmoins quelques changements importants ont été observés : une extension N-terminale plus courte, une rotation de l'hélice  $\alpha_2$  et une extension C-terminale non repliée en hélice et localisée à l'intérieur du site G. La mutation de l'asparagine 29 en cystéine ne permet pas de conférer à l'enzyme une activité et une capacité à fixer le GSH. Ceci peut être expliqué par l'obturation du site G par l'extension C-terminale (Illustration 13). Chen *et al.* (2011) supposent que les isoformes BmGSTO1-1, BmGSTO2-2 et BmGSTO4-4 possèdent les mêmes particularités structurales que BmGSTO3-3 car elles fixent également difficilement le glutathion. Pour restaurer cette capacité à fixer le GSH, il a fallu effectuer une délétion des résidus de l'extension C-terminale en plus de la mutation de l'asparagine en cystéine. De plus l'enzyme modifiée présente une activité de conjugaison avec le CDNB. *In vivo*, le rôle des ces enzymes reste inconnu car elles ne présentent probablement pas d'activité de conjugaison. Toutefois les activités DHAR et thiol transférase n'ont pas été testées.

## **A.II.2    Les CLIC**

Les membres de la famille des GST CLIC (ChLoride Intracellular Channel) sont des protéines pouvant fonctionner comme des canaux d'ions chlorures. Ces protéines sont exprimées dans une grande variété de tissus chez les organismes multicellulaires. Six représentants existent chez l'homme. Ils jouent un rôle dans divers processus tels que la résorption des os et l'apoptose (Berry & Hobert, 2006). Les CLIC existent à la fois sous forme soluble et fixées à la membrane, le processus d'insertion restant mal connu (Achilonu *et al.*, 2012). Les formes soluble et membranaire présentent des structures tertiaires et quaternaires différentes. Les CLIC possèdent une signature



similaire au motif catalytique des glutarédoxines et thiorédoxines du type CXXS ou CXXC. Il a été montré que les CLIC2 et CLIC3 ont la capacité de former un pont disulfure intramoléculaire entre les deux cystéines de la signature (Littler *et al.*, 2010).

La forme soluble est monomérique et présente le repliement classique des GST. Le monomère garde la capacité à fixer le GSH comme les GST Oméga et Bêta en formant un pont disulfure. Mais à ce jour aucune activité de conjugaison n'a été détectée chez les CLIC (Oakley, 2011). Ces protéines présentent une longue boucle originale par rapport aux GST, nommée « foot-loop » et insérée entre les hélices  $\alpha_5$  et  $\alpha_6$ . Cette boucle de 15 résidus environ est relativement flexible et montre une faible conservation de séquence entre les CLIC (Littler *et al.*, 2010).

Lors de leur oxydation, les CLIC subissent une transition structurale de la forme monomérique (réduite) à une forme dimérique (oxydée). Le dimère est non covalent et chaque monomère, tout hélice, a perdu le feuillet  $\beta$  du domaine thiorédoxine. L'oxydation a lieu entre la première cystéine de la signature (C24 dans la CLIC1 humaine) et la cystéine 59 localisée dans l'hélice  $\alpha_2$ . Il a été supposé que cette transition permet l'exposition de résidus hydrophobes à la surface de la protéine pour faciliter son insertion dans la membrane (Oakley, 2005).

### **A.II.3 Les GST Bêta**

Les GST de la classe Bêta sont spécifiques au règne des bactéries, où elles sont largement distribuées chez les procaryotes à Gram négatif, et sous-représentées chez les organismes à Gram positif (Allocati *et al.*, 2012). Elles sont impliquées dans de nombreux processus de détoxification tels que l'élimination d'agents antimicrobiens et la dégradation de composés monocycliques et polycycliques (Allocati *et al.*, 2009). La première GST Bêta à avoir été caractérisée est celle de *Proteus mirabilis* (PmGSTB1-1, Rossjohn *et al.*, 1998b). Elle présente une activité GSH transférase sur le CDNB, ainsi qu'une activité du type glutarédoxine du fait de la présence de la cystéine dans le site actif (Di Ilio *et al.*, 1988).

Aujourd'hui quatre structures de GST Bêta ont été décrites dans la littérature, à savoir celles de *Proteus mirabilis* (PmGSTB1-1, Rossjohn *et al.*, 1998b), d'*E. coli* (EcGSTB1-1, Nishida *et al.*, 1998), de *Burkholderia xenovorans* (BxGSTB1-1 ou BphK, Tocheva *et al.*, 2006) et d'*Ochrobactrum anthropi* (OaGSTB1-1, Federici *et al.*, 2007). Six autres structures de GST déposées dans la Protein Data Bank provenant de programmes de génomique structurale appartiennent potentiellement aussi à la classe Bêta.

Les superpositions des structures des GST Bêta montrent une très bonne conservation du repliement GST dans cette classe (Rmsd moyen  $< 1.5$  Å) (Allocati *et al.*, 2009). Le motif « lock-and-key » observé dans les structures des GST mammifères est absent dans les GST Bêta, par

contre le motif « N-capping Box » est conservé. De plus, un motif structural supplémentaire a été identifié comme spécifique à cette classe Bêta. Ce motif est formé par un réseau de liaisons hydrogène, situé à proximité du site G, qui ferme l'hélice du domaine C-terminal avec le début de l'hélice du domaine N-terminal. La rupture partielle de ce réseau a des conséquences dramatiques sur la stabilité et la fonctionnalité de la OaGSTB1-1 (Federici *et al.*, 2007). Les GST Bêta fixent le GSH dans une forme étendue qui est stabilisée par plusieurs interactions hydrophobes et polaires, incluant une liaison hydrogène avec un résidu aspartate ou glutamate de l'hélice  $\alpha_4$  de l'autre monomère comme dans les GST Alpha, Mu et Pi. Le glutathion forme un pont disulfure avec la cystéine du site actif dans la structure de la PmGSTB1-1. Cet état oxydé, toutefois, n'est pas présent dans les structures des OaGSTB1-1 et BxGSTB1-1, qui elles-mêmes possèdent un glutathion réduit.

#### **A.II.4 Les GST Lambda et les DHAR**

La recherche des GST de la classe Oméga chez *A. thaliana* a conduit à l'identification de deux nouvelles classes de GST à cystéine : les GST Lambda et les DésHydroAscorbate Réductases (DHAR) (Dixon *et al.*, 2002). Comme les CLIC (voir ci-dessus), les GST Lambda et les DHAR présentent la singularité d'être monomériques. *A. Thaliana* possède 3 gènes codants pour des DHAR (deux cytosoliques, AtDHAR1 et AtDHAR2, et une chloroplastique, AtDHAR3) et deux gènes codants pour des GST Lambda (une cytosolique, AtGSTL1, et une chloroplastique, AtGSTL2). Ces GST ne présentent pas d'activité de conjugaison ni d'activité peroxydase. Les DHAR ont un profil catalytique similaire aux glutarédoxines en présentant des activités DHAR et thiol transférase. Seule cette dernière activité est retenue dans les GST Lambda. L'AtDHAR3 chloroplastique, possédant une signature du type CXXC, est la seule à avoir la capacité de former un pont disulfure intramoléculaire. Les études biochimiques ont permis de montrer que les autres se lient de manière covalente au glutathion. A ce jour, aucune structure de GST Lambda ou de DHAR n'a été déposée dans la banque de données des structures de protéines (PDB).

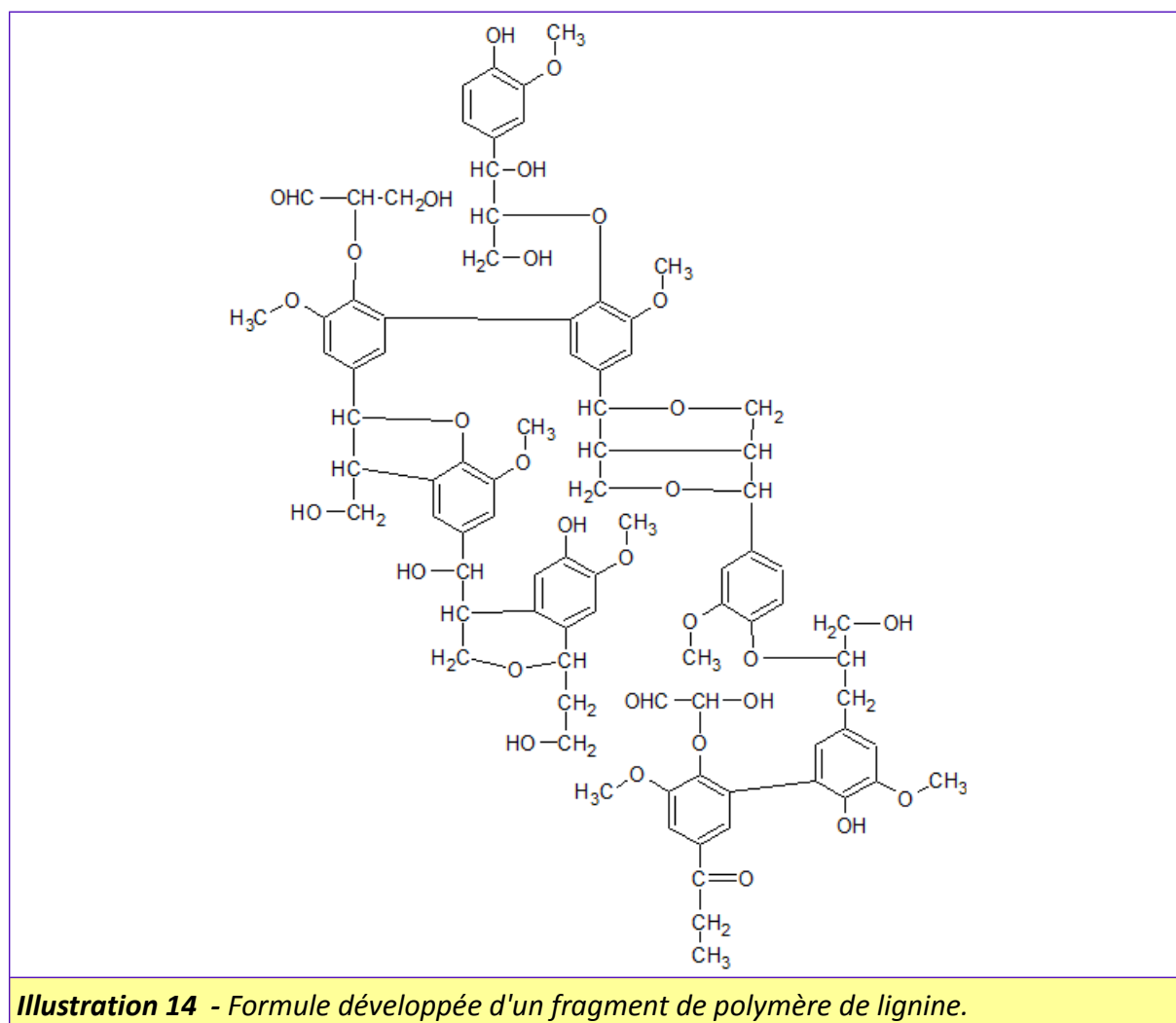
Il a été alors proposé que les GST Lambda et les DHAR puissent avoir une fonction alternative en réponse au stress oxydant (Dixon & Edwards, 2010). En effet, les GST Lambda cytosoliques sont fortement exprimées chez *A. thaliana*, le maïs et le blé dans la réponse aux traitements chimiques (Dixon & Edwards, 2010). Des études récentes ont mis en évidence un rôle des GST Lambda dans le métabolisme des flavonoïdes (Dixon & Edwards, 2010). Par exemple, l'enzyme TaGSTL1 du blé (*Triticum aestivum*) reconnaît sélectivement la taxifoline avec un  $K_d$  de 25 nM. La co-fixation des flavonols comme la quercétine à la TaGSTL1 thiolée représente probablement une étape intermédiaire dans le métabolisme des quinones glutathionylées dont

dérivent les flavonols libres. Ceci suggère que les GST Lambda jouent un rôle nouveau en maintenant les flavonoïdes sous des conditions de stress.

### A.III Les GST et la dégradation du bois

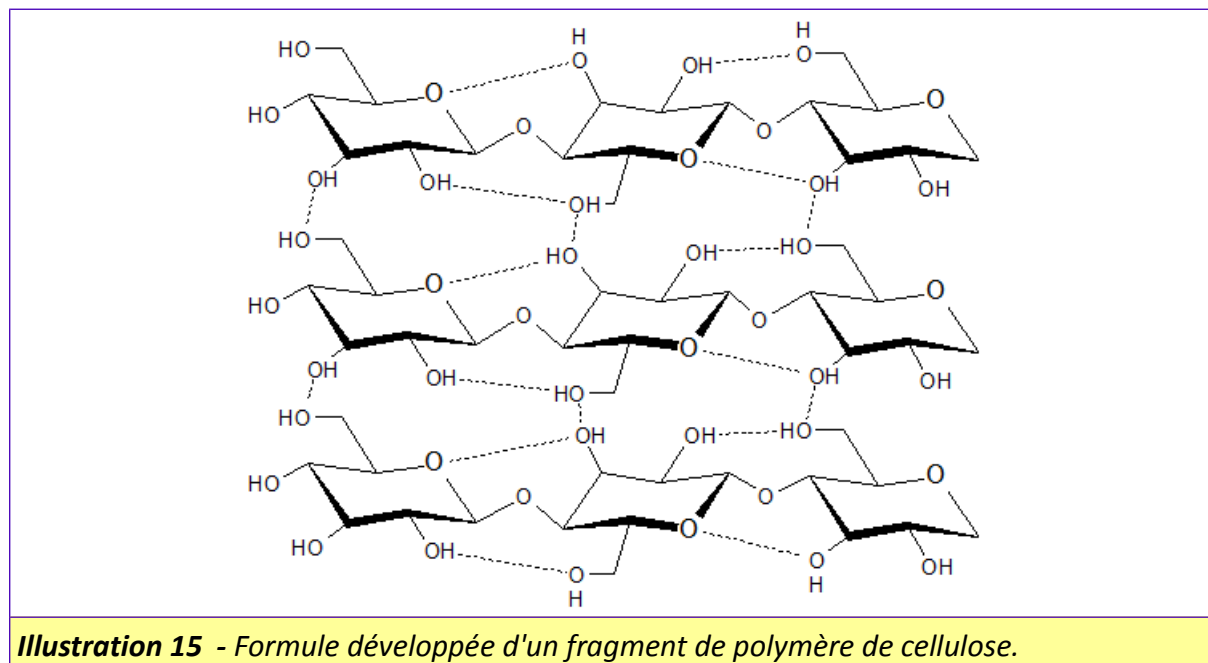
#### A.III.1 La lignine

Le bois, aussi appelé xylème secondaire, est constitué d'un ensemble de tissus dont les parois des cellules sont lignifiées. Sa composition chimique est complexe et varie selon l'essence de l'arbre (Illustration 14).



Un réseau complexe de polymères composés de cellulose (plus de 45 % du poids du bois sec), d'hémicellulose (25-30 %) et de lignine (15-30 %) représente plus de 90 % de la matière sèche du bois (Perez *et al.*, 2002 ; Otsuka *et al.*, 2003). La lignine agit comme un composé structural qui supporte les plantes en donnant de la force et de la rigidité à la paroi cellulaire et donc à la plante

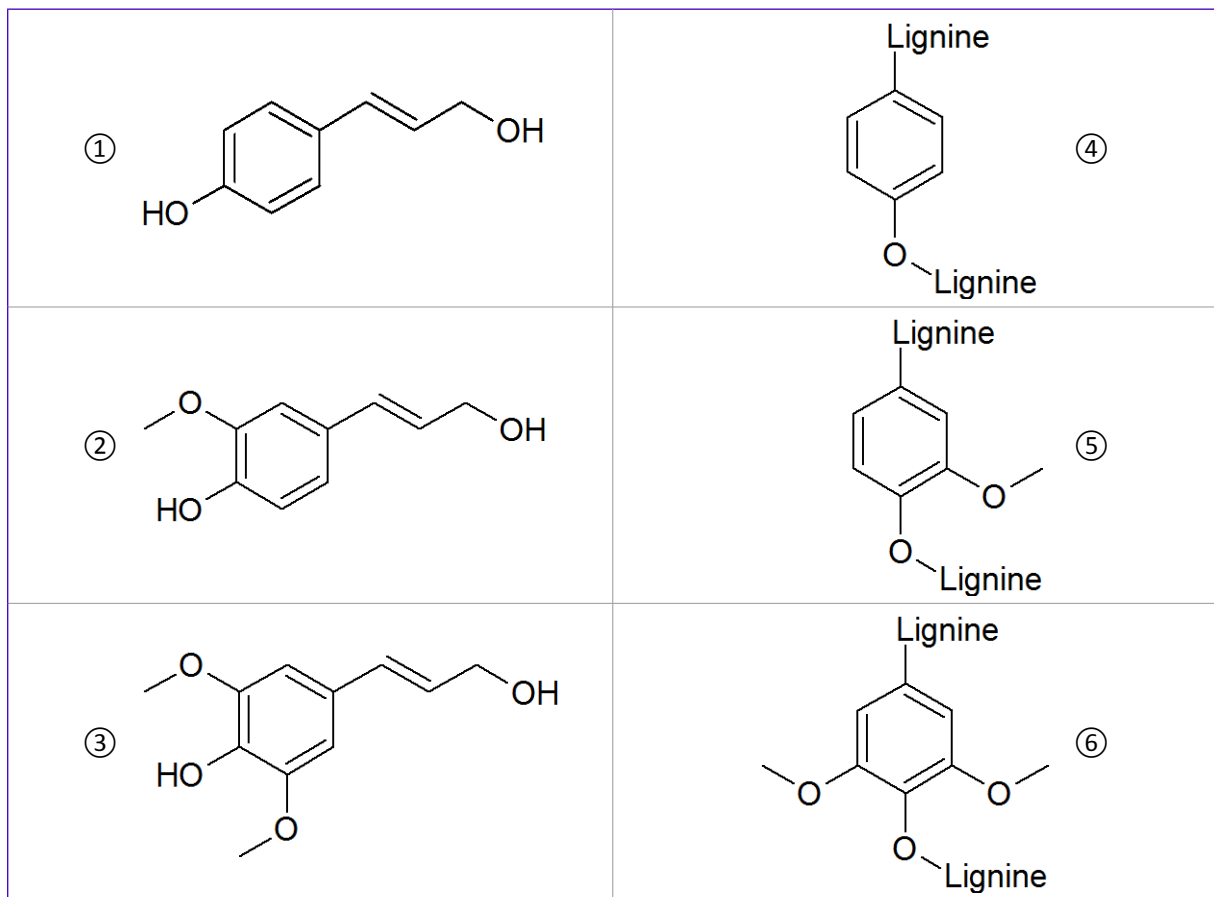
elle-même. La lignine joue également un rôle de protection car son polymère constitutif complexe et varié est extrêmement résistant à divers agents chimiques et à la dégradation biologique. La lignine devient une barrière servant à protéger la cellulose (qui est un homopolymère de glucose, Illustration 15) et les autres carbohydrates.



La lignine est la plus abondante source de composés aromatiques dans la nature et sa décomposition est essentielle au cycle du carbone de la terre. Les champignons saprophytes de la pourriture blanche sont supposés être impliqués dans la dégradation initiale de la lignine native, tandis que les bactéries jouent un rôle principal dans la minéralisation des composés dérivés de la lignine de faible poids moléculaire (Masaï *et al.*, 2007).

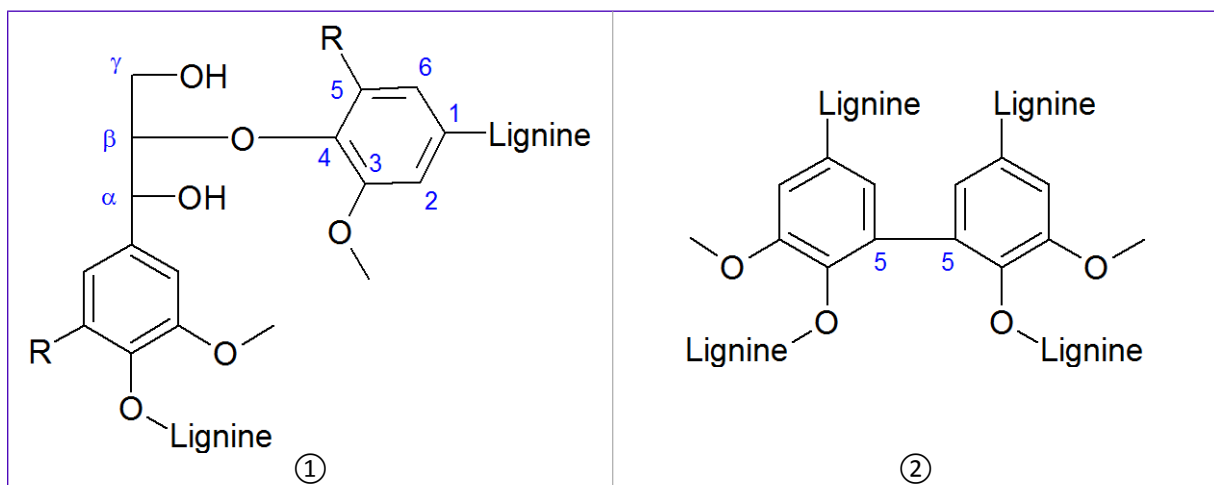
La libération de la cellulose par la lignine est une étape limitante dans l'extraction du glucose pour des applications telle que la production d'éthanol dérivé de la biomasse des plantes. La biodégradation enzymatique de la lignine a pour objectif de permettre une production d'éthanol en grande quantité tout en évitant la création de nombreux déchets industriels.

La lignine est un hétéropolymère complexe résultant du couplage radicalaire biocatalysé de trois types principaux d'unités monolignols que sont l'alcool coumarylique (ou p-coumaryle), l'alcool conyférylique (ou coniféryle) et l'alcool synapylique (Illustration 16). Les principales liaisons entre ces unités trouvées dans les lignines sont les liens  $\beta$ -O-4- qui sont les plus abondants ainsi que les liens  $\beta$ - $\beta$ -,  $\beta$ -5-, 5-5-, et 5-O-4- (Illustration 17).



**Illustration 16** - Formules développées de trois monolignols et de leurs unités lignifiées.

① : alcool paracoumarylique ; ② : alcool coniférylique ; ③ : alcool sinapylique ; ④ : unité p-hydroxyphényle ; ⑤ : unité guaïacyle ; ⑥ : unité syringyle.



**Illustration 17** - Deux modes de liaison entre deux unités monomériques de lignine.

① : mode de liaison inter-unité labile ( $\beta$ -O-4) ; ② : mode de liaison inter-unité résistant (biphényle).

### **A.III.2 La dégradation de la lignine**

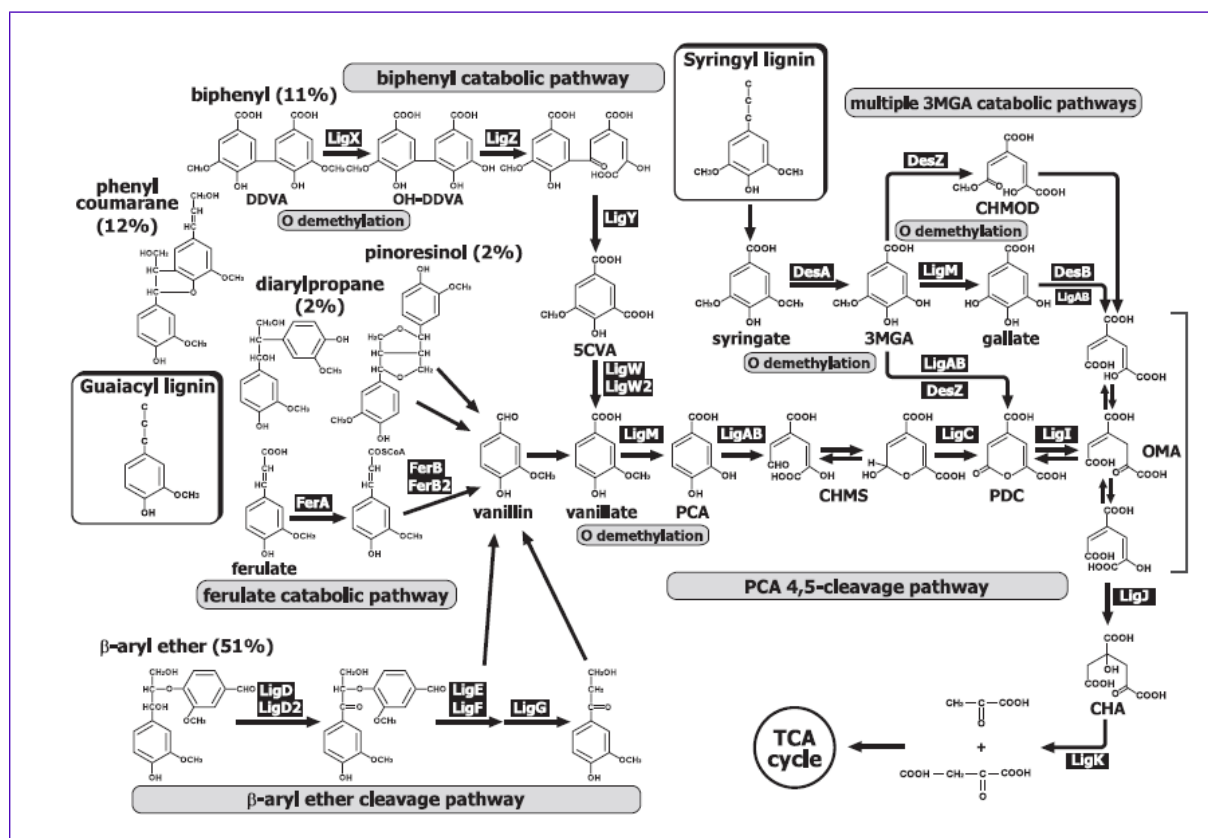
Dans la nature, plusieurs systèmes sont impliqués dans la dégradation de la lignine. Initialement les basidiomycètes secrètent des peroxydases et/ou des laccases pour dégrader la lignine polymérique à l'état brut. Les champignons lignolytiques les plus efficaces sont ceux de la pourriture blanche tels que *Phanerochaete chrysosporium* et *Tramètes versicolor* (ou *Coriolus versicolor*). *P. chrysosporium* produit deux types de peroxydase, la manganèse peroxydase (MnP) et la lignine peroxydase (LiP), et *T. versicolor* produit généralement des laccases (Otsuka *et al.*, 2003). La laccase réagit avec les polyphénols incluant la lignine et d'autres composés aromatiques dérivés de la lignine qui, à leur tour, peuvent être à la fois polymérisés ou dépolymérisés. L'enzyme MnP peut oxyder  $Mn^{2+}$  en  $Mn^{3+}$  qui, à son tour, est capable d'oxyder un très grand nombre de substrats phénoliques incluant la lignine. L'enzyme LiP oxyde directement une variété de composés aromatiques phénoliques et non phénoliques en formant des radicaux libres. Bien que ces radicaux provoquent principalement la rupture des liaisons C $\alpha$ -C $\beta$  et  $\beta$ -O-4 de la lignine (Illustration 17), ils induisent également une dépolymérisation aléatoire de la lignine (Otsuka *et al.*, 2003). De nombreux fragments de lignine de faible masse moléculaire produits par ces peroxydases et/ou laccases sont décomposés en eau et dioxyde de carbone par des enzymes bactériennes très spécifiques, telles que les enzymes de fission du noyau aromatique, les déméthylases et les  $\beta$ -étherases (Otsuka *et al.*, 2003). La dégradation de la lignine est ainsi accomplie par les actions coopératives de systèmes enzymatiques fongiques et bactériens.

### **A.III.3 La coupure de liens $\beta$ -aryl éther catalysée par des GST de *Sphingobium sp. SYK-6***

L'équipe du Dr Masai a isolé en 1991 une bactérie capable de dégrader de nombreux composés de lignine dimérique (Katayama *et al.*, 1987). Cette bactérie alors connue sous le nom de *Pseudomonas paucimobillis* SYK-6 s'appelle aujourd'hui *Sphingobium sp. SYK-6*. Cette bactérie utilise une voie métabolique particulière impliquant la protocatechuate 4,5-dioxygénase pour dégrader des lignines dimériques dont les produits sont impliqués dans le cycle de l'acide tricarboxylique (Masai *et al.*, 1991).

*Sphingobium sp. SYK-6* a la capacité de se développer sur une grande variété de lignines dimériques avec des liaisons intermonomériques présentes dans la lignine naturelle, incluant les composés arylglycérol- $\beta$ -aryl-éther, phénylcoumarane, biphényle, diarylpropane, et pinosésinol (Illustration 17) (Masai *et al.*, 1991). Ces composés dimériques sont transformés en vanillate ou syringate qui sont dégradés par les voies cataboliques du protocatechuate (PCA) et du 3-O-

methylgallate (3MGA) (Illustration 18) (Masai *et al.*, 2007). Dans ce processus de dégradation de la lignine, la coupure de la liaison  $\beta$ -aryl-éther est indispensable car elle représente entre 50 et 70 % des liaisons entre les monomères. Le Dr Masai et ses collaborateurs ont isolé le groupe de gènes *ligDFEG* de *Sphingobium* sp. SYK-6, codant pour des enzymes capables de couper le lien  $\beta$ -aryl-éther d'un fragment dimérique de lignine modèle et appartenant à une famille de déshydrogénase et de réductase (Illustration 19) (Sato *et al.*, 2009). Ce groupe de gènes forme une seule unité transcriptionnelle. L'activité  $\beta$ -éthérase avait été mise en évidence dans une fraction de la membrane de *Sphingobium* sp. SYK-6 (Masai *et al.*, 1989).

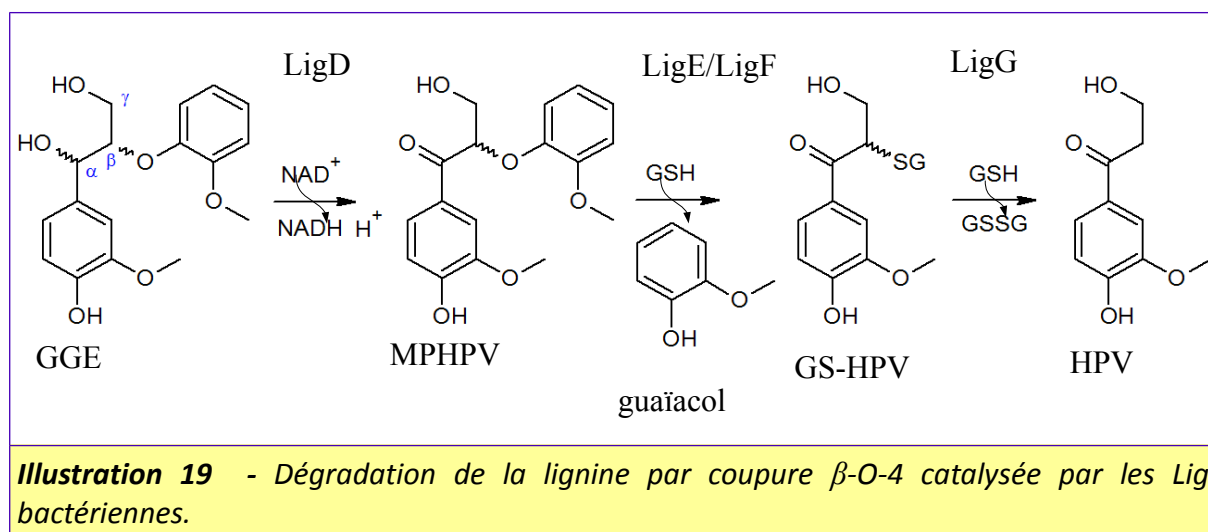


**Illustration 18** - Voies cataboliques pour la dégradation des composés aromatiques dérivés de la lignine par *Sphingobium* sp. SYK-6 (d'après Masai *et al.*, 2007).

Abréviations : DDVA, 5,5'-désydrodivanilate ; OH-DDVA, 2,2',3-trihydroxy-3'-méthoxy-5,5'-dicarboxybiphényl ; 5CVA, 5-carboxyvanillate ; PCA, protocatechuate ; CHMS, 4-carboxy-2-hydroxy-muconate-6-semialdéhyde ; PDC-pyrone-4,6-dicarboxylate ; OMA, 4-oxalamesaconate ; CHA, 4-carboxy-4-hydroxy-2-oxoadipate ; 3MGA 3-O-methylgallate;CHMOD, 4carboxy-2hydroxy-6-méthoxy-6-oxohexa-2,4-dienoate ; TCA, acide tricarboxylique.

Le gène *ligD* code pour la  $\alpha$ -déshydrogénase (LigD) qui oxyde le guaiacylglycérol- $\beta$ -guaiacyl éther (GGE) en  $\alpha$ -(2-méthoxyphénoxy)- $\beta$ -hydroxypropiovanillone (MPHPV). Cette enzyme LigD présente des similarités avec les alcool deshydrogénases de courtes chaînes (Illustration 19) (Sato *et al.*, 2009).

Le lien éther du composé MPHPV est coupé par les enzymes LigE et LigF en présence de glutathion. LigE et LigF partagent 20 % d'identité avec l'enzyme DiChloroMéthane Déhalogénase (DCMA) qui est connue pour être un membre de la superfamille des GST (Masaï *et al.*, 1993). Ces deux enzymes appartenant alors à la famille des glutathion transférases catalysent l'attaque nucléophile du glutathion sur l'atome de carbone en position  $\beta$  du MPHPV pour produire la  $\beta$ -hydroxypropiovanillone glutathionylée (GS-HPV) et du guaiacol (Sato *et al.*, 2009). LigE et LigF sont également actives avec le composé GGE (groupement alcool sur le carbone en position  $\alpha$ ). Des études par dichroïsme circulaire réalisées par le Dr Masai et ses collaborateurs ont permis de révéler que les enzymes LigF et LigE attaquent des énantiomères différents de MPHPV (respectivement MPHPV-I et MPHPV-II). L'enzyme LigG a une activité lyase spécifiquement sur le produit de la réaction catalysée par LigF, à savoir GS-HPV-I. LigG appartient également à la famille des GST et présente une homologie de séquences d'environ 20 % avec les enzymes de la classe Oméga. Ceci implique que le résidu catalytique putatif de LigG est une cystéine comme les GST Oméga (Masaï *et al.*, 1993). Aucune étude structurale de ces enzymes n'avait été reportée au départ de ce travail de thèse.



## A.IV Les GST de *Phanerochaete chrysosporium*

### A.IV.1 Présentation du champignon

*Phanerochaete chrysosporium* est un champignon qui possède ce nom lorsqu'il est dans sa forme de reproduction sexuée appelé également état téléomorphe. Ce champignon possède également un état anamorphe (forme à reproduction asexuée), dans ce cas l'espèce s'appelle *Sporotrichum pulverulentum* (Illustration 20) (Burdshall *et al.*, 1981).



*P. chrysosporium* appartient au phylum *basidiomycota* (basidiomycètes) et est connu pour être un champignon de la pourriture blanche (d'appellation anglaise « white-rot-fungus »). Il est capable de dégrader les différentes parties du bois des feuillus (ou « hardwood » qui contient une lignine avec des unités guaiacyles et syringyles, principalement, Illustration 16) et des résineux (ou « softwood » qui contient une lignine avec des unités guaiacyles essentiellement) que l'on trouve dans les forêts tempérées de l'Amérique du Nord, de l'Europe et de l'Iran (Burdshall, 1985). Ce champignon possède la propriété unique de pousser de manière optimale à 40°C. Il produit de nombreuses spores asexuelles.

*P. chrysosporium* excrète principalement deux types de peroxydases pour dégrader la lignine, les lignine peroxydases (LiP; E.C.1.11.1.14) et les manganèse peroxydases (MnP; E.C.1.11.1.13). Par ailleurs, ce basidiomycète sécrète d'autres enzymes pour dégrader de nombreux composés organiques complexes (Singh & Chen, 2008) (Tableau 2).



**Illustration 20** - Le champignon *Phanerochaete chrysosporium* poussant sur des copeaux de bois.

Les LiP dégradent seulement les composés non-phénoliques et agissent avec le peroxyde d'hydrogène (Hatakka, 2001). Les réactions catalysées par les MnP nécessitent un apport en manganèse qui permet à l'enzyme de catalyser l'oxydation de Mn(II) en Mn(III), en présence de ligands stabilisant le Mn(III). Les complexes Mn(III) en résultant peuvent alors oxyder les substrats organiques (Hatakka, 2001) et ainsi permettre la dégradation de la lignine. La Mn peroxydase peut également catalyser l'oxydation de plusieurs colorants aromatiques et de phénols monoaromatiques (Barr & Aust, 1994).

Composés dégradés	Enzymes produites
Lignines alcalines de pin et de paille	Lignine peroxydase
Pulpe et eau de déchets de papier	Manganèse-peroxydase
décolore l'eau de déchets textiles	peroxydase
Congo red, Amaranth, Atrazine, colorants azotés	Cellulases Succinimides
Alcool veratryl et ses methyl éther	Glyoxal oxydase
Trichlophénol, pentachlorophénol, toluène	Xylanase
DDT, toluène, benzène, xylène	Aryl-alcool deshydrogénase
Acide humique	Cellobiose deshydrogénase, Agent solubilisant le charbon

**Tableau 2** - Les enzymes de *P. chrysosporium* impliquées dans la dégradation de différents composés (Singh & Chen, 2008).

#### A.IV.2 Les GST fongiques

Dès le début des années 1970, l'implication de GST fongiques dans la résistance aux fongicides avait été mise en évidence chez *Saccharomyces pastorianus* (Tillman et al., 1973). Même si des GST issues d'autres organismes ont ensuite été caractérisées, ces études restent peu nombreuses. Deux études phylogénétiques ont tenté de recenser les GST fongiques (McGoldrick *et al.*, 2005 ; Morel *et al.*, 2009).

La première étude, basée sur 21 espèces, propose cinq gros clusters de GST ne pouvant pas être rangés dans les classes connues de GST en 2005 (McGoldrick *et al.*, 2005). Deux clusters n'avaient jamais été identifiés auparavant et les trois autres appartiennent respectivement aux familles Ure2p et eEF1B $\gamma$  (voir paragraphe A.I.2.b), et aux MAK16. Un sixième petit cluster comprenant seulement deux membres avait déjà été identifié auparavant comme une nouvelle classe spécifique des champignons, appelée Gamma (Cha *et al.*, 2002). D'autres clusters plus petits ont été identifiés et correspondent à des classes de GST connues telles que les classes Alpha, Mu, Pi (spécifiques aux mammifères), Phi et Tau (spécifiques aux plantes), et même les MAPEG. Toutefois ces petits clusters ne semblent pas ubiquitaires dans le règne fongique.

Dans la seconde étude, la diversité des GST fongiques a été étudiée à partir de 30 génomes d'ascomycètes, de zygomycètes et de basidiomycètes. Ces organismes ont été également différenciés suivant qu'ils sont saprophytes, symbiotiques ou pathogènes (Morel *et al.*, 2009). Dans cette analyse phylogénétique, il apparaît dans que le nombre total de séquences de GST putatives diffère fortement d'une espèce à l'autre (de 7 à 46 séquences). Ces différences semblent plutôt reliées à la physiologie du champignon. Les organismes contenant un nombre élevé de GST sont ceux qui correspondent aux ascomycètes et aux basidiomycètes de grande complexité pouvant

dégrader beaucoup de composés organiques (c'est-à-dire saprophytes tels que *P. chrysosporium* avec 27 séquences). Les levures sont les organismes présentant le moins de séquences de GST, tels que *S. cerevisiae* avec 11 séquences.

Cette étude a mis en évidence sept classes de GST : GTT1, GTT2, EFB $\gamma$ , MAK16, URE2p-like, Oméga et une nouvelle classe GSTFuA (initialement appelée GTE). Ces trois dernières classes sont particulièrement étendues dans les organismes présentant un nombre élevé de GST.

Les classes GTT1 et GTT2 ont été caractérisées dans la levure *S. cerevisiae* qui possède une GTT1 (ScGTT1) et une GTT2 (ScGTT2). Ces deux protéines sont des enzymes qui présentent une activité avec le CDNB et semblent être impliquées dans des processus de détoxification (Morel *et al.*, 2009). Des études cristallographiques et biochimiques de la ScGTT2, publiées récemment, révèlent qu'elle ne possède aucun des résidus catalytiques classiques des GST (tyrosine, sérine ou cystéine) (Ma *et al.*, 2009). Il a été supposé qu'une molécule d'eau stabilisée par deux résidus polaires de l'hélice  $\alpha_4$  permet l'activation du glutathion.

Les protéines des classes EFB $\gamma$  et MAK16 n'ont pas d'activité GST. L'EFB $\gamma$  est un facteur d'élongation et la protéine MAK16 est une protéine nucléaire impliquée dans le cycle de progression de la cellule et dans la biogenèse des sous-unités ribosomiques 60S (McGoldrick *et al.*, 2005). La classe Ure2p a été définie en référence à l'Ure2p de *S. cerevisiae* qui est un facteur de transcription dans la répression de la voie d'utilisation de l'uréidosuccinate. Chaque sous-unité de cette protéine dimérique présente une extension N-terminale (résidus 1-93) qui est un domaine impliqué dans la formation de fibrilles amyloïdes (Bousset *et al.*, 2001). Les résidus 93 à 354 présentent une structure de type GST avec une extension particulière entre les hélices  $\alpha_5$  et  $\alpha_6$ . Les enzymes YghU et YfcG d'*E. coli* appartiennent également à cette classe Ure2p sans posséder l'extension N-terminale de l'Ure2p de *S. cerevisiae*. Ces deux enzymes bactériennes sont des thioltransférases et présentent la particularité de fixer deux molécules de glutathion par monomère (Wadington *et al.*, 2009 ; Stourman *et al.*, 2011). L'analyse phylogénétique des GST fongiques révèle que les Ure2p de *P. chrysosporium* ne possèdent pas l'extension N-terminale de l'Ure2p de *S. cerevisiae*. Elles peuvent être divisées en deux sous-classes, une sous-classe homologue à l'Ure2p de *S. cerevisiae* (Ure2pA) et une sous-classe homologue aux Ure2p bactériennes (Ure2pB) (Thuillier *et al.*, 2011).

Le génome de *P. chrysosporium* possède huit gènes codants pour des GST Oméga (Tableau 3). Les huit GST Oméga appelées PcGSTO1 à PcGSTO8 ont été divisées en trois sous-classes : la sous-classe I comprend les PcGSTO1 et PcGSTO2, la sous-classe II les PcGSTO3 à PcGSTO5 et la sous-classe III les PcGSTO6 à PcGSTO8. Les sous-classes II et III sont homologues aux GST Lambda (spécifiques des plantes) et aux GST Oméga de mammifères (voir respectivement les

paragraphes A.II.4 et A.II.1). Les PcGSTO de la sous-classe I présentent des homologues chez les bactéries et les plantes. L'homologue bactérien est une tetrachlorohydroquinone déhalogénase de structure inconnue (McCarthy *et al.*, 1996).

Le génome de *P. chrysosporium* possède cinq gènes codants pour la nouvelle classe GSTFuA. Les GSTFuA ont comme plus proche homologue la protéine bactérienne LigE de *Sphingobium* sp. SYK-6, connue pour être impliquée dans la dégradation de la lignine (voir paragraphe A.III.3 et Masai *et al.*, 1991). L'identité de séquence, variant de 16,2 à 24,7%, reste toutefois modeste (Morel *et al.*, 2009). Les génomes des hémiascomycètes (levure et champignons apparentés aux levures) étudiés ne présentent pas de gènes codant pour les GSTFuA. Cette classe semble plus présente dans les champignons saprophytes (3 à 14 gènes) que dans les pathogènes (1 à 4 gènes). Ces analyses suggèrent que les GSTFuA pourraient être impliquées dans la dégradation de la matière organique récalcitrante.

Classe	Sous-classe	Nom	Total
Oméga	I	PcGSTO1, PcGSTO2	8
	II	PcGSTO3 à PcGSTO5	
	III	PcGSTO6 à PcGSTO8	
GTE	I	PcGSTFuA1 à PcGSTFuA4	5
	II	PcGSTFuA5	
Ure2p	A	PcUre2p1	9
	B	PcUre2p2 à PcUre2p9	
GTT2	I	Phchr6683, Phchr6766	3
	II	Phchr7971	
MAK16	-	-	1
EF1By	-	-	1
Total			27

**Tableau 3** - Les GST de *P. chrysosporium*

## A.V Objectifs de la thèse

Les travaux cristallographiques exposés dans ce manuscrit ont été motivés par deux objectifs principaux. Le premier avait l'ambition d'étudier dans le règne fongique la diversité structurale des glutathion transférases (GST) encore peu décrites par rapport à celles provenant d'organismes des règnes animal, végétal et bactérien. En collaboration avec l'équipe du professeur

E. Gelhaye du laboratoire Interactions Arbre-Microorganismes, notre deuxième objectif concernait l'étude des relations entre la structure et la fonction de plusieurs glutathion transférases du champignon saprophyte *Phanerochaete chrysosporium*. Nous avons en particulier tenté de découvrir si cet organisme fongique avait une activité ligninolytique en impliquant des glutathion transférases comme cela a été observé dans la bactérie du sol *Sphingobium* sp. SYK-6.

Les GST forment une superfamille d'enzymes présentes dans la plupart des organismes vivants. Elles sont bien connues pour leur rôle dans la phase II de la détoxification cellulaire où elles transfèrent le glutathion sur des xénobiotiques afin de faciliter leur élimination ou stockage. Toutefois, la grande pluralité des GST fait qu'elles sont impliquées dans de nombreux processus biologiques en ne jouant pas toujours un rôle catalytique (Oakley, 2005). Les premières structures cristallographiques ont été déterminées au début des années 90 (Reinemer *et al.*, 1992; Ji *et al.*, 1992). Les GST cytosoliques sont majoritairement homodimériques avec un repliement structural unique malgré de faibles identités de séquences. De nombreuses structures ont été déterminées visant des applications notamment dans le domaine de la santé car les GST sont impliquées à la fois dans la détoxification d'agents cancérigènes mais aussi dans l'élimination d'agents chimiothérapeutiques. Les GST de champignons ont été beaucoup moins étudiées probablement parce que la plupart se distinguent de leurs homologues procaryotes. La première étude structurale est celle de l'Ure2p de *Saccharomyces cerevisiae* qui est particulière d'une part car elle est impliquée dans une cascade de transduction du signal qui régule le catabolisme de l'azote dans la levure (Umland *et al.*, 2001). D'autre part, elle possède une extension N-terminale capable de former des fibrilles d'amyloïdes telle la protéine prion chez certains vertébrés (Bousset *et al.*, 2001).

McGoldrick et collaborateurs (2005) ont tenté une première description de la variété des GST dans le règne fongique à partir de peu d'organismes partiellement séquencés. Plus récemment nos collaborateurs ont pu faire une description plus exhaustive de cette diversité grâce au programme '1000 Fungal Genomes Project' lancé par un consortium international d'équipes de recherche en collaboration avec l'institut de séquençage 'Joint Genome Institute' (JGI) du département de l'énergie des Etats-Unis d'Amérique. Ce programme vise le séquençage de deux représentants des 500 familles de champignons reconnues. Sept classes de GST fongiques cytosoliques ont été définies : les MAK16, les EF1B $\gamma$ , les Ure2p, les Oméga (divisées en deux sous-classes : les S-glutathionyl-(chloro)-hydroquinone réductases (GHR), et les homologues à la GSTO humaine), les GTT et les GTE pour 'Glutathione Transférase Etherase', (Morel *et al.*, 2009). Notre attention s'est portée sur les GHR et sur les GTE car l'analyse des structures primaires des isoformes de ces deux classes ne décelait pas d'homologue structural proche. Les tentatives de construction de modèles par homologie produisaient des structures avec de mauvais facteurs de

confiance. De plus, les données biochimiques concernant les représentants de ces classes étaient peu nombreuses.

La classe des GTE paraissait particulièrement intéressante pour plusieurs raisons. Tout d'abord aucun représentant de cette classe n'avait été caractérisé auparavant. Par ailleurs, cette classe est distribuée de façon hétérogène dans le règne des champignons en étant quasiment uniquement étendue dans les champignons interagissant avec le bois. Son plus proche homologue caractérisé est LigE qui est une enzyme de la bactérie du sol *Sphingobium sp.* SYK-6 connue avec LigF et LigG, pour leur capacité à couper un lien éther sur un modèle de petits fragments de lignines (Masai *et al.*, 2003). Le basidiomycète *P. chrysosporium* est un champignon de la pourriture blanche ayant la capacité à dégrader la lignine qui possède 5 isoformes de cette nouvelle classe GTE. Notre objectif était dans un premier temps d'étudier les relations structure-fonction d'au moins une isoforme. Il nous a également semblé nécessaire de tenter les études cristallographiques des enzymes LigE, LigF et LigG dont l'activité ligninolytique avait été décrite (Masai *et al.*, 2003). La compréhension des mécanismes moléculaires de la dégradation de ce polymère complexe qu'est la lignine reste très importante à plusieurs titres. En effet, des efforts sont engagés pour valoriser la lignine dans les bioraffineries afin d'utiliser toutes les parties de la plante. Dans le domaine de la bioremédiation, la découverte de nouveaux systèmes capables de dégrader la lignine peut être appliquée pour dégrader des molécules organiques complexes qualifiées de récalcitrantes. Enfin, le contrôle des dégâts occasionnés par les champignons saprophytes sur le bois mort reste un enjeu important pour la filière bois (protection des charpentes et autres bois de construction par exemple).

## **B Résultats**

Ce chapitre présente les résultats des travaux que j'ai réalisés durant mon stage de doctorat. Leur objectif général était de contribuer à la caractérisation des GST de *Phanerochaete chrysosporium*, en réalisant l'analyse structurale de membres représentatifs de différentes classes présentes dans cet organisme, et en exploitant ces résultats en relation avec les connaissances biochimiques acquises sur ces mêmes GST par l'équipe du Pr Gelhaye de l'UMR 1136.

La première partie concerne les PcGHR, dont la structure de la PcGHR1 (initialement appelée PcGSTO1) a été déterminée. Son analyse a été réalisée à la lumière des analyses biochimiques concernant les PcGHR1 et PcGSTO3. Ces résultats ont donné lieu en 2011 à une publication dans *Journal of Biological Chemistry* (volume 286, pages 9162 à 9173) intitulée « Glutathione Transferases of *Phanerochaete chrysosporium* : S-glutathionyl-p-hydroquinone reductase belongs to a new structural class » et dont les auteurs sont E. Meux, P. Prosper, A. Ngadin, C. Didierjean, M. Morel, S. Dumarçay, T. Lamant, J.-P. Jacquot, F. Favier et E. Gelhaye. Le paragraphe B.I donne une courte introduction de cet article et ses principaux résultats avant de présenter l'article lui-même. Par ailleurs, les travaux sur PcGHR1 ont été poursuivis avec la préparation de complexes entre l'enzyme et des substrats glutathionylés. Ces essais et l'analyse de leurs résultats provisoires sont présentés au paragraphe B.II.

La deuxième classe de GST de *P. chrysosporium* étudiée est celle des PcGSTFuA. La partie B.III donne un résumé des travaux réalisés. La structure d'un premier représentant, la PcGSTFuA1, a été établie, puis complétée d'une étude étendue aux autres membres de la classe. Deux publications en ont découlé. La première, écrite par Y. Mathieu, P. Prosper, M. Buée, S. Dumarçay, F. Favier, E. Gelhaye, P. Gérardin, L. Harvengt, J.-P. Jacquot, T. Lamant, E. Meux, S. Mathiot, C. Didierjean et M. Morel et publiée en 2012 dans *Journal of Biological Chemistry* (volume 287, pages 39001 à 39011) sous le titre « Characterization of a *Phanerochaete chrysosporium* glutathione transferase reveals a novel structural and functional class with ligandin properties », relate les caractéristiques originales de PcGSTFuA1. Elle est présentée dans la partie B.III.1. La seconde, acceptée dans PLOS One le 1<sup>er</sup> Octobre 2013 et rédigée par Y. Mathieu, P. Prosper, F. Favier, L. Harvengt, C. Didierjean, J.-P. Jacquot, M. Morel-Rouhier et E. Gelhaye sous le titre « Diversification of fungal specific class A glutathione transferases in saprotrophic fungi », s'appuie sur PcGSTFuA2 et PcGSTFuA3, très proches, pour compléter les connaissances structurales et biochimiques sur cette famille (voir paragraphe B.III.2).

Les 5 GSTFuA de *P. chrysosporium* ont pour plus proche homologues la protéine bactérienne LigE de *Sphingobium* sp. SYK-6. Le système Lig de cette bactérie présenté au paragraphe A.III.3 voit deux autres de ses membres (LigF et LigG) appartenir également à la superfamille des GST. L'étude structurale de ces trois protéines paraissait intéressant dans le cadre de cette thèse, en particulier par comparaison aux PcGSTFuA pour LigE et LigF. Les paragraphes B.IV.1 et B.IV.2 présentent les travaux menés sur ces deux enzymes assurant énantiosélectivement la coupure de la liaison aryl éther  $\beta$ -O-4 des lignines. Aucune de ces deux structures n'a pu être totalement finalisée. Une structure préliminaire de la LigF a toutefois été obtenue. Le paragraphe B.IV.3 relate ensuite les résultats acquis sur LigG. Cette dernière protéine correspondant au cluster de gènes *ligDFEG* de *Sphingobium* sp. SYK-6 exerce une activité lyase sur les produits glutathionylés issus des réactions catalysées par LigE et LigF. L'analyse de la structure de LigG a donné lieu à une publication d'E. Meux, P. Prosper, E. Masai, G. Mulliert, S. Dumarçay, M. Morel, C. Didierjean, E. Gelhaye et F. Favier, parue en 2012 dans *FEBS Letters* (volume 586, pages 3944 à 3950) sous le titre « *Sphingobium* sp. SYK-6 LigG involved in lignin degradation is structurally and biochemically related to the glutathione transferase omega class ». Un bref résumé de cet article est présenté avant l'article lui-même.

## **B.I Etude structurale de la PcGHR1 (ou PcGSTX1)**

### **B.I.1 Introduction**

Parmi les huit GST de *P. chrysosporium* initialement apparentées aux GST Oméga sur la base de leur séquence, PcGSTO1 et PcGSTO2 (plus tard renommées PcGHR1 et PcGHR2) définissent la sous-classe I. Aucun des homologues les plus proches, observé chez les plantes ou les bactéries, n'en a une structure connue. Au contraire, les PcGSTO3 à 5 de la sous-classe II et les PcGSTO6 à PcGSTO8 de la sous-classe III ont pour homologues les plus proches les GST Oméga, deux d'entre elles ayant leur structure connue chez l'homme (hGSTO1-1 et hGSTO2-2), et une chez l'insecte *Bombyx mori* (BmGSTO3-3) (voir paragraphe A.II.1).

Afin d'amorcer la caractérisation des GST à cystéine de *P. chrysosporium* - la classe des GST la plus sur-représentée chez ce champignon saprophyte en regard des champignons moins complexes (voir paragraphe A.IV.2) - un représentant de chacun des deux groupes a été sélectionné : le choix s'est porté sur PcGSTO1 pour la sous-classe I, et sur PcGSTO3 pour les sous-classes II et III.

Lorsque cette étude a débuté, PcGSTO1 s'annonçait posséder une activité potentielle de type glutaredoxine, par comparaison aux GSTO de classe I chez *S. cerevisiae*, bien que ces



dernières présentent avec PcGSTO1 des différences de séquence au niveau du site H. Pendant la détermination de la structure de PcGSTO1, Xun et collaborateurs (2010) ont prolongé les travaux sur les GST impliquées dans la dégradation du pentachlorophénol chez cette bactérie, et proposé par extension que les GSTO de sous-classe I formaient en réalité une nouvelle classe, les GHR, ainsi nommées du fait de leur activité S-glutathionyl-(chloro)hydroquinone réductase. Cette activité devait donc être investiguée chez PcGSTO1.

Concernant PcGSTO3, la structure n'en a pas été déterminée du fait de l'homologie initialement constatée avec HsGSTO1-1 et HsGSTO2-2 de structures connues en 2010 (codes PDB 1EEM et 3LFL pour la première, 3Q18 et 3Q19 pour la seconde).

### **B.I.2 Résultats et discussion**

La séquence de PcGSTO1 possède, en amont de la partie GST à proprement parler, une longue extension dont la présence a posé quelques problèmes lors de la production de l'enzyme. Il semble en effet que cette partie présente des sites de clivage reconnus par la machinerie enzymatique d'*E. coli* utilisée comme système d'expression, de sorte que les échantillons produits consistaient en des mélanges de formes de longueur variable. Ce résultat, nuisible à l'homogénéité requise pour l'étape de cristallogénèse, a nécessité la production d'une forme tronquée de 46 acides aminés N-terminaux, utilisée pour l'analyse cristallographique. De plus, compte tenu de l'objectif structural, à savoir la détermination de la première structure d'un représentant de la sous-classe I des GSTO, l'enzyme a été sélénée au niveau de ses méthionines pour permettre le phasage par la méthode SAD. Des cristaux monocliniques contenant 2 monomères de 352 acides aminés chacun par unité asymétrique ont été exploités jusqu'à une résolution de 2,3 Å. A cette limite, 70 % du modèle a pu être reconstruit automatiquement dans la densité électronique par ARP/wARP (Perrakis *et al.*, 1999), le complément ayant été progressivement reconstruit manuellement au fur et à mesure des cycles d'affinement. Le modèle final comporte une chaîne A dont les 37 premiers acides aminés n'ont pas été observés, et qui présente une densité nette attribuée à une molécule de glutathion bien que cette entité n'ait pas été ajoutée dans le milieu de cristallisation, et une chaîne B dont manquent 53 acides aminés du côté N-terminal, le site actif ne comportant cette fois pas de molécule de GSH.

L'analyse du repliement de PcGSTO1 révèle plusieurs points remarquables. Le premier concerne l'élongation N-terminale, relativement à l'entrée du brin  $\beta 1$  qui débute au résidu 78. Bien qu'aucun des monomères n'en montre la totalité, A et B permettent toutefois d'observer une large portion de cette élongation, et révèlent l'absence de structure secondaire dans cette région qui se positionne de façon étendue en courant tant contre le domaine thiorédoxine que le domaine tout hélice, du côté des sites G et H. La seconde spécificité est la grande longueur de la boucle connectant  $\beta 2$  à  $\alpha 2$ , comportant 20 résidus. Et la dernière caractéristique de repliement concerne

l'extrémité C-terminale qui présente, par rapport au repliement canonique des GST, une hélice  $\alpha_9$  additionnelle, suivie d'une boucle de quelques résidus. L'hélice  $\alpha_9$  n'est pas spécifique de PcGSTO1, elle est en particulier observée dans les structures de GSTO, en revanche le positionnement des résidus qui la suivent est propre à PcGSTO1, et provient notamment de l'élongation de la boucle reliant  $\beta_2$  à  $\alpha_2$ . Ces quelques éléments conduisent à la spécificité majeure de PcGSTO1 : le dimère classique est remplacé par un nouvel assemblage, lui aussi dimérique, dans lequel l'association des monomères se fait par interaction directe du domaine tout hélice avec lui-même, via des interactions aussi bien polaires qu'hydrophobes. Ce mode de dimérisation est retrouvé dans une structure de GST potentielle de *Corynebacterium glutamicum*, déposée dans la PDB (code 3M1G) dans le cadre d'un programme de génomique structurale et non publiée. L'ensemble de ces éléments structuraux nous a conduits à proposer que PcGSTO1 et cette dernière GST forment une nouvelle classe structurale baptisée Xi, qui regrouperait notamment toutes les GSTO de sous-classe I. De plus, celles-ci définissent une nouvelle classe fonctionnelle appelée GHR (Xun *et al.*, 2010) présentant une activité S-glutathionyl-(chloro)hydroquinone réductase (voir paragraphe précédent).

L'activité S-glutathionyl-hydroquinone réductase a effectivement été confirmée par l'équipe du Pr. Eric Gelhaye, face à des substrats tels que le S-glutathionyl-p-benzohydroquinone ou la ménadione (2-méthyl-1,4-naphtoquinone) glutathionylée, conduisant à des produits déglutathionylés et la formation de glutathion oxydé GSSG. Cette activité GHR s'ajoute à l'activité thiol transférase de type glutaredoxine observée vis-à-vis de substrats tels que l'hydroxyéthyl-disulfide (HED) ou le déshydroascorbate (DHA). En revanche, aucune activité glutathion transférase n'a été constatée, avec le substrat classique CDNB. Toutes ses observations résultent de la présence du résidu catalytique cystéinyle (Cys86), situé en N-terminal de l'hélice  $\alpha_1$ , au fond du site G dans lequel le glutathion du monomère A est observé. Sa fixation classique est néanmoins renforcée par quelques interactions supplémentaires mettant en jeu les éléments additionnels du repliement de PcGSTO1. L'ensemble résulte en un GSH relativement enfoui, et un site H potentiel très peu accessible et petit. L'accès de ce site à des substrats de type ménadione glutathionylée nécessite probablement un mouvement de l'extrémité N-terminale, ou une structuration à l'état physiologique différente de celle observée dans la structure cristalline.

Ces premiers travaux structuraux et biochimiques sur PcGSTO1 ont donc permis de conclure à l'existence d'une nouvelle classe structurale de GST, dénommée Xi et notamment caractérisée par un nouveau mode d'assemblage du dimère, dont les GSTO de sous-classe I présentes chez les champignons et les bactéries sont des représentants ayant une fonction GHR. Sur le plan fonctionnel, la PcGSTO1 doit donc être renommée PcGHR1 ; d'un point de vue structural, elle est le premier membre caractérisé de la famille des GST Xi et, en accord avec la nomenclature usuellement utilisée pour les classes, elle doit être nommée PcGSTX1.

Par comparaison, les études biochimiques sur PcGSTO3 révèlent une incapacité à déglutathionyler des composés tels que la *S*-glutathionyl-*p*-hydroquinone, c'est-à-dire l'absence d'activité GHR, quand l'enzyme est capable de déglutathionyler des composés tels que le *S*-(phenylacetophenone)glutathion, ce que ne peut faire PcGHR1. Plus généralement, PcGSTO3 présente un profil d'activité comparable aux GST Oméga humaines. Les différences de séquences entre ces enzymes et la comparaison des structures de PcGSTX1 et HsGSTO1-1 expliquent l'origine de ces différences d'activités.

Toutes ces informations sont précisées dans le texte de la publication de Meux *et al.* parue en 2011 dans *J. Biol. Chem.* (vol. 286, p. 9162-9173), dont le texte est présenté ci-après.

### **B.I.3      Publication**

Voir page suivante.

# Glutathione Transferases of *Phanerochaete chrysosporium* *S*-GLUTATHIONYL-*p*-HYDROQUINONE REDUCTASE BELONGS TO A NEW STRUCTURAL CLASS<sup>\*§</sup>

Received for publication, October 14, 2010, and in revised form, December 17, 2010. Published, JBC Papers in Press, December 22, 2010, DOI 10.1074/jbc.M110.194548

Edgar Meux<sup>†</sup>, Pascalita Prosper<sup>§</sup>, Andrew Ngadin<sup>†</sup>, Claude Didierjean<sup>§</sup>, Mélanie Morel<sup>†</sup>, Stéphane Dumarçay<sup>¶</sup>,  
Tiphaine Lamant<sup>‡</sup>, Jean-Pierre Jacquot<sup>‡</sup>, Frédérique Favier<sup>§1</sup>, and Eric Gelhaye<sup>‡2</sup>

From the <sup>†</sup>UMR 1136 INRA-UHP "Interactions Arbres/Micro-Organismes," IFR 110 "Ecosystèmes Forestiers, Agroressources, Bioprocédés et Alimentation," the <sup>§</sup>CRM2, Equipe Biocristallographie, UMR 7036 CNRS-UHP, Institut Jean Barriol, and the <sup>¶</sup>Laboratoire d'Études et de Recherches sur le Matériau Bois, EA UHP 4370, Nancy Université, Faculté des Sciences et Techniques, BP 70239, 54506 Vandoeuvre-les-Nancy, France

The white rot fungus *Phanerochaete chrysosporium*, a saprophytic basidiomycete, possesses a large number of cytosolic glutathione transferases, eight of them showing similarity to the Omega class. PcGSTO1 (subclass I, the bacterial homologs of which were recently proposed, based on their enzymatic function, to constitute a new class of glutathione transferase named *S*-glutathionyl-(chloro)hydroquinone reductases) and PcGSTO3 (subclass II related to mammalian homologs) have been investigated in this study. Biochemical investigations demonstrate that both enzymes are able to catalyze deglutathionylation reactions thanks to the presence of a catalytic cysteinyl residue. This reaction leads to the formation of a disulfide bridge between the conserved cysteine and the removed glutathione from their substrate. The substrate specificity of each isoform differs. In particular PcGSTO1, in contrast to PcGSTO3, was found to catalyze deglutathionylation of *S*-glutathionyl-*p*-hydroquinone substrates. The three-dimensional structure of PcGSTO1 presented here confirms the hypothesis that it belongs not only to a new biological class but also to a new structural class that we propose to name GST xi. Indeed, it shows specific features, the most striking ones being a new dimerization mode and a catalytic site that is buried due to the presence of long loops and that contains the catalytic cysteine.

*Phanerochaete chrysosporium* is considered as the model organism to study the physiology of white rot fungi, the only known microorganisms able to completely break down lignin

\* This work was supported by Agence Nationale de la Recherche Research Grants ANR-06-BLAN-0386 and ANR-09-BLAN-0012, Ministère de l'Enseignement Supérieur de la Recherche et de la Technologie, Institut National de la Recherche Agronomique, and CNRS.

This work is dedicated to the memory of the late Dr. André Aubry.

§ The on-line version of this article (available at <http://www.jbc.org>) contains supplemental Table 1, Figs. 1 and 2, and additional references.

The atomic coordinates and structure factors (code 3PPU) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (<http://www.rcsb.org/>).

<sup>1</sup> To whom correspondence may be addressed: CRM2, UMR 7036 CNRS-UHP, Nancy Université, Faculté des Sciences et Techniques, BP 70239, 54506 Vandoeuvre-les-Nancy Cedex, France. Tel.: 33-3-83-68-48-79; Fax: 33-3-83-40-64-92; E-mail: Frederique.Favier@crm2.uhp-nancy.fr.

<sup>2</sup> To whom correspondence may be addressed: UMR 1136 INRA-UHP, IFR110, Nancy Université, Faculté des Sciences et Techniques, BP 70239, 54506 Vandoeuvre-les-Nancy, France. Tel.: 33-3-83-68-42-28; E-mail: Eric.Gelhaye@lcb.uhp-nancy.fr.

to carbon dioxide and water (1–3). Analysis of the genome of this basidiomycete revealed a complex superfamily of cytochrome P450 monooxygenases representing around 1% of this genome (1, 4). The large size of this gene family is correlated with the extraordinary ability of this fungus to degrade and metabolize numerous recalcitrant compounds such as pesticides, polyaromatic hydrocarbons, halogenated aromatics, and textile dyes. Besides extracellular oxidative systems, including lignin and manganese peroxidases, cytochrome P450 monooxygenases are thought to be involved in the phase I of detoxication processes catalyzing the oxidation of various substrates (5–7). The oxidation of recalcitrant compounds is often followed by glutathione conjugation catalyzed by glutathione transferases (GSTs). The glutathione conjugates formed are less toxic and lipophilic than the parent compounds (8).

The distribution of GSTs in fungi has been recently investigated (9), highlighting that *P. chrysosporium* exhibits a large subfamily of cytosolic GSTs (at least 27 isoforms) in comparison with *Saccharomyces cerevisiae* for instance (11 isoforms). Based on primary structure similarity, these cytosolic GSTs belong to seven different classes as follows: Ure2p-like, GTT1, GTT2, EFB $\gamma$ , MAK16, etherases, and Omega (GSTO)<sup>3</sup> (9). Among the cytosolic GSTs, *P. chrysosporium* exhibits eight putative proteins showing homology with GSTs of the Omega class, named PcGSTO1 to PcGSTO8. In a previous phylogenetic analysis of fungal GSTs, we found that GSTOs fall into three different subclasses. Subclass 1 encloses PcGSTO1 and PcGSTO2, these latter being related to uncharacterized plant GST Lambda and bacterial GSTs (9). Members of subclasses II and III are related to the well characterized human HsGSTO1-1 (10, 11).

A major feature of GSTOs is the presence of a cysteinyl residue in the catalytic site present in the glutathione binding domain (G-site), explaining the deglutathionylation activities exhibited by these proteins. For instance, HsGSTO1-1 has been shown to remove glutathione from *S*-(phenacyl)glutathione (12), as this activity requires the presence of the catalytic Cys<sup>32</sup> (13). In addition, GSTOs have been shown to be involved in the reduction of methyl and dimethyl arsonate, an intermediate in

<sup>3</sup> The abbreviations used are: GSTO, glutathione transferase Omega; DHA, dehydroascorbate; GHR, *S*-glutathionyl-(chloro)hydroquinone reductase; HED, hydroxyethyl disulfide; PDB, Protein Data Bank; PAP-SG, *S*-(phenacylacetylphenone)glutathione; Se, selenium; CDNB, 1-chloro-2,4-dinitrobenzene.

the arsenic biotransformation (14, 15). Bacterial GSTOs, which belong to subclass I, have been shown to be involved in pentachlorophenol catabolism acting as *S*-glutathionyl-(chloro)hydroquinone reductase; this activity was also related to the presence of a catalytic cysteine (16, 17). In fungi, few studies have been devoted to this GST class except in *S. cerevisiae* (ScGSTO). The three ScGSTOs belong to the subclass I and have been partially characterized particularly at the protein level. The three proteins exhibit glutaredoxin-like activities because they are able to reduce both hydroxyethyl disulfide and dehydroascorbate in the presence of glutathione. In addition, the transcription of the three genes is induced by oxidative stress (18, 19). Moreover, ScGSTO1 has been shown to be involved in sulfur metabolism and located in peroxisomes. PcGSTO1 and PcGSTO2, nevertheless, differ from their yeast counterparts particularly at the H-site (the domain thought to interact with the hydrophobic substrate).

From phylogenetic analysis and biochemical analysis, Xun *et al.* (17) recently proposed that the Omega subclass I members constitute a new class of GSTs named *S*-glutathionyl-(chloro)hydroquinone reductase (GHR) due to their activity with *S*-glutathionyl-hydroquinones.

The aim of this study was the investigation of the PcGSTOs, a class of GSTs that is over-represented in *P. chrysosporium*. Recombinant proteins of isoforms that belong to the GSTO subclasses I and II have been biochemically characterized, exhibiting substrate specificity. Furthermore, the three-dimensional structure of PcGSTO1 has been solved, allowing the first description of a *S*-glutathionyl-*p*-hydroquinone reductase at the structural level. From these biochemical and structural data, the potential functions of the different isoforms are discussed.

## EXPERIMENTAL PROCEDURES

**Materials**—Hydroxyethyl disulfide (HED) and 5,5'-dithiobis-2-nitrobenzoic acid were from Aldrich and Pierce, respectively. All other reagents were from Sigma.

**Cloning and Construction of *Pcgsto1* and *Pcgsto3* Mutants by Site-directed Mutagenesis**—The open reading frame sequences encoding *P. chrysosporium* PcGSTO1 and PcGSTO3 were amplified from *P. chrysosporium* cDNA using *Pcgsto1* and *Pcgsto3* forward and reverse primers (see supplemental Table 1) and cloned into the NcoI and BamHI restriction sites (underlined in the primers) of pET3d (Novagen). For PcGSTO3, the amplified sequence encoded a protein in which an alanine has been added during cloning. Using two complementary mutagenic primers, the two cysteines of PcGSTO3 were individually substituted into serines, and the primers are listed in supplemental Table 1. For PcGSTO1, due to the presence of the signal peptide, a first construction has been made, and the resulting protein starts with the N-terminal sequence MATTICLRH. Because we obtained a mixture of different proteins after production and purification (see under "Results"), a new construction has been made. The resulting protein with an additional alanine started thus with the N-terminal sequence MASFT-TGST (see supplemental Fig. 1). PcGSTO1 used for structure determination consists also of the 352 last residues of the full-length protein.

**Expression and Purification of the Recombinant Proteins**—For protein production, the *Escherichia coli* BL21(DE3) strain that contained the pSBET plasmid was co-transformed with the different recombinant plasmids (20). Cultures were progressively amplified up to 2.4 liters in LB medium supplemented with ampicillin and kanamycin at 37 °C. Protein expression was induced at exponential phase by adding 100 μM isopropyl β-D-thiogalactopyranoside for 4 h at 37 °C. The cultures were centrifuged for 15 min at 4400 × *g*. The pellets were resuspended in 30 ml of TE NaCl buffer (30 mM Tris-HCl, pH 8.0, 1 mM EDTA, 200 mM NaCl), and the suspension was stored at −20 °C.

Cell lysis was performed by sonication (three times for 1 min with intervals of 1 min), and the soluble and insoluble fractions were separated by centrifugation for 30 min at 27,000 × *g*. The soluble part was then fractionated with ammonium sulfate in two steps, and the protein fraction precipitated between 40 and 80% of the saturation contained the recombinant protein, as estimated by 15% SDS-PAGE. The protein was purified by size exclusion chromatography after loading on an ACA44 (5 × 75 cm) column equilibrated in TE NaCl buffer. The fractions that contained the protein were pooled, dialyzed by ultrafiltration to remove NaCl, and loaded onto a DEAE-cellulose column (Sigma) in TE buffer (30 mM Tris-HCl, pH 8.0, 1 mM EDTA). The proteins were eluted using a 0–0.4 M NaCl gradient. Finally, the fractions of interest were pooled, dialyzed, concentrated by ultrafiltration under nitrogen pressure (YM10 membrane; Amicon), and stored in TE buffer at −20 °C. Purity was checked by SDS-PAGE. Protein concentrations were determined spectrophotometrically using a molar extinction coefficient at 280 nm of 64,860 M<sup>−1</sup>·cm<sup>−1</sup> for the PcGSTO1 WT, C86S, and 51340 M<sup>−1</sup>·cm<sup>−1</sup> for PcGSTO3 WT and C37S.

*E. coli* strain B121 containing the pET-*Pcgsto1* plasmid was cultured as described previously (21) and was used as a source of selenomethionine-substituted form of PcGSTO1 (SeMet-PcGSTO1). The purification of SeMet-PcGSTO1 was performed following a procedure similar to that described for PcGSTO1.

**Mass Spectrometry Analysis**—PcGSTOs WT and mutated were analyzed by quadrupole-TOF MS as described by Koh *et al.* (22).

**Fluorescence Properties of Wild-type and Mutated PcGSTOs**—The fluorescence characteristics of PcGSTOs in the reduced and oxidized forms were recorded with a spectrofluorometer (Cary Eclipse; VARIAN) in TE buffer at a protein concentration of 10 μM.

**Determination of Free Thiol Groups**—The number of free thiol groups in either untreated, denatured, or reduced conditions was determined spectrophotometrically with 5,5'-dithiobis-2-nitrobenzoic acid, as described by Koh *et al.* (22). All thiol titrations were performed on enzymes either as purified or after SDS treatment or after dithiothreitol (DTT) reduction. In the latter case, proteins were first reduced by incubation with 10 mM DTT for 30 min at room temperature, then precipitated with 1 volume of 20% (w/v) trichloroacetic acid (TCA), and finally stored on ice for 30 min. The mixture was then centrifuged for 10 min at 13,000 × *g* and the pellets were washed twice with 2% (w/v) TCA. The precipitate was resuspended in 30 mM Tris-HCl, pH 7.0, 1 mM EDTA, and 1% (w/v) SDS. The

## GSTs of *P. chrysosporium*

concentration was estimated by UV spectrophotometry. For denatured proteins, samples were incubated in the presence of 2% SDS treatment during 30 min before 5,5'-dithiobis-2-nitrobenzoic addition.

**Activity Measurements**—The activity measurements of WT or mutant PcGSTOs in the HED assay or for reduction of dehydroascorbate (DHA) were performed as described by Couturier *et al.* (23). Dimethylarsinate reductase and glutathionyl-phenylacetophenone assay were performed as described previously (12, 13). The GST activity was measured with 0.2 mM CDNB as the substrate in 30 mM Tris buffer, pH 8.0, and the continuous absorbance at 340 nm was monitored. *S*-Glutathionyl-*p*-hydroquinone reductase activity was monitored as follows: 1 mM 1,4-benzoquinone in 30 mM Tris-HCl, pH 8, was used as base line for an absorbance spectrum (230–400 nm). After 2 min, 1 mM reduced glutathione (GSH) was added to the reaction mixture, and the spectra were monitored every minute during 5 min. At last, purified recombinant proteins were added in the reaction mixture, and spectra were recorded every minute. In complementary experiments, 1 mM 1,4-benzoquinone in 30 mM Tris-HCl, pH 8, was incubated in the presence of 1 mM GSH, 200  $\mu$ M NADPH, 0.5 IU glutathione reductase before the addition of the protein of interest. The activity was followed by monitoring the decrease in absorbance arising from NADPH oxidation in this coupled enzyme assay system that showed the formation of glutathione disulfide (GSSG).

In complementary experiments, *S*-(phenylacetophenone)-glutathione (PAP-SG) and 2-methyl-*S*-glutathionyl-naphthoquinone (thiodione) were prepared as described previously (24, 25), and the purity and the nature of the products were verified by mass spectrometry. Activities of the WT and variant enzymes were assayed by following the appearance of de-glutathionylated menadione (at 425 nm) and phenylacetophenone (at 290 nm). The reactions were performed in 30 mM Tris-HCl, pH 8, containing GSH (5 mM) and substrates (from 15 to 2 mM) in a final volume of 650  $\mu$ l and initiated by adding WT or mutant enzyme. The concentration of enzyme added was adjusted based on preliminary activity experiments, ranging from 0.1 and 1  $\mu$ M. The reactions were stopped by adding 350  $\mu$ l of ethanol, then strongly mixed by vortex, then centrifuged at 14,500 rpm during 15 min, and analyzed by HPLC using a Gemini 5- $\mu$ m C18 column (150 mm long  $\times$  4.6 mm inner diameter) at 22  $^{\circ}$ C.

To visualize the de-glutathionylation reactions, PAP-SG or thiodone (15  $\mu$ M) was added in a 650- $\mu$ l reaction mixture containing 30 mM Tris-HCl, pH 8, and 15  $\mu$ M enzyme. After stopping the reaction by adding ethanol, samples were analyzed by HPLC as described above.

**Size Exclusion Chromatography**—Size exclusion chromatography experiments have been performed using a Superose 12HR column connected to a fast protein liquid chromatography (FPLC) system. The column was equilibrated with 30 mM Tris-HCl buffer, pH 8, containing 0.15 M NaCl, and the proteins were eluted in the same buffer at a flow rate of 0.25 ml $\cdot$ min $^{-1}$ .

**Crystallization and X-ray Diffraction Data Collection**—Prior to crystallization, protein samples were concentrated to 40 mg $\cdot$ ml $^{-1}$  PcGSTO1, in TE buffer. Crystals of the SeMet-PcGSTO1 were obtained by using the microbatch method. 1.5

**TABLE 1**  
Statistics of X-ray diffraction data collection and model refinement

Data collection	
Space group	C2
Cell dimensions	<i>a</i> , 166.8 Å, <i>b</i> , 70.3 Å, <i>c</i> , 72.6 Å; $\beta$ , 98.8 $^{\circ}$
Resolution	46.66 to 2.30 Å (2.43 to 2.30 Å) <sup>a</sup>
$R_{\text{merge}}$	3.7% (18.8%)
$I/\sigma I$	19.8 (4.8)
Completeness	98.9% (96.0%)
Redundancy	4.3
Refinement	
Resolution	46.66 to 2.30 Å
No. of reflections	36,622
$R_{\text{all}}/R_{\text{free}}^b$	19.2/22.8
No. of atoms	
Protein	2572 (monomer A) + 2451 (monomer B)
Ligand	1 glutathione (20 atoms) in monomer A
Water	223
<i>B</i> -factors	
Protein	37.9 (monomer A); 41.6 (monomer B)
Ligand	49.4
Water	38.6
Ramachandran statistics	
Residues in preferred regions	95.7%
Residues in allowed regions	4.3%
Outlier residues	0.0%
Root mean square deviations	
Bond lengths	0.011 Å
Bond angles	1.25 $^{\circ}$

<sup>a</sup> Values in parentheses are for highest resolution shell.

<sup>b</sup>  $R_{\text{all}}$  is determined from all the reflections (working set + test set), and  $R_{\text{free}}$  corresponds to a subset of 5% of reflections (test set).

$\mu$ l of the PcGSTO1 solution at various concentrations and 1.5  $\mu$ l of various crystallization solutions were deposited in a 72-well Microbatch plate (Hampton Research) pre-filled with paraffin oil and stored at 4  $^{\circ}$ C. Preliminary results have been obtained from two commercial screening solutions as follows: the condition B2 from the JBScreen classic kit 2 (Jena Bioscience) and the Crystal Screen solution 6 from Hampton Research. Further trials aimed to improve reproducibility and crystal quality gave the best results for PcGSTO1 solutions at 30–40 mg $\cdot$ ml $^{-1}$  and crystallization solutions composed of 20–30% polyethylene glycol 4000 (PEG 4000), 0 to 0.2 M magnesium or calcium chloride, and 0.1 M Tris-HCl, pH 8.5. However, polycrystals grew frequently so that only rare portions of them were suitable for x-ray data collection. No crystal of non-selenomethionylated PcGSTO1 suitable for data collection has been obtained.

The diffraction data were collected at 100 K on the synchrotron beamline FIP-BM30A (26) at ESRF (France) with a wavelength of 0.9805 Å corresponding to the peak of the Se K-edge, from a crystal grown with 40 mg $\cdot$ ml $^{-1}$  PcGSTO1, 25% PEG 4000, and no salt, and flash-frozen after a quick immersion in the crystallization solution mixed with 20% glycerol. The data sets were processed with XDS (27) and scaled with SCALA from the CCP4 Package (28). Statistics are summarized in Table 1.

**Structure Solution and Quality of the Model**—The PcGSTO1 structure was solved using the single wavelength anomalous dispersion protocol of Auto-Rickshaw at the EMBL-Hamburg automated crystal structure determination platform (29). The input diffraction data were prepared and converted for use in Auto-Rickshaw using programs of the CCP4 suite (28). Nine heavy atoms among the 10 requested were found using the program SHELXD (30). Initial phases were calculated after density modification using the program SHELXE (31), at 3.0 Å resolution. The initial phases were improved using density modifica-

tion and phase extension to 2.30 Å resolution using the program DM (32). 70% of the model was built by ARP/wARP (33, 34). It was progressively completed and refined by using iteratively COOT (35) and REFMAC5 (36), including diffraction data up to 2.3 Å resolution. An electron density further attributed to one GSH bound in the active site appeared in monomer A during refinement, although no GSH was added to crystallize PcGSTO1. On the contrary, the corresponding area of monomer B only displayed a few poor spherical densities that more probably corresponded to water molecules. Final model statistics are shown in Table 1.

Most residues corresponded to well defined electron density, except for the N-terminal part of both monomers that appeared too disordered to be observed. Thus, models of monomers A and B contained residues Glu<sup>38</sup>–Asp<sup>351</sup> and residues Phe<sup>54</sup>–Asp<sup>351</sup>, respectively. Furthermore, the residues Glu<sup>38</sup>–Ser<sup>53</sup> only observed in monomer A displayed a weaker electron density and higher *B* factors with respect to the rest of the structure. Figs. 4 and 5 were prepared by using PyMOL.

## RESULTS

PcGSTO1 belongs to the first subclass of GSTs Omega displaying a CPWATR potential active site compatible with the CP(W/F)(A/T)(H/Q)R characteristic of the members of this subclass (Fig. 1, *top* and *bottom*) (9). According to the prediction softwares, PcGSTO1 possesses a signal peptide targeting to mitochondria. Different attempts to produce heterologously the full-length protein in *E. coli* resulted in a mixture of truncated proteins and suggested that this N-terminal extension contained putative cleavage sites recognized by the *E. coli* machinery or an intrinsic instability of the precursor. These data were in accordance with the presence of a signal peptide in the N-terminal part of the protein as shown previously for different other redox proteins (37, 38). To obtain a homogeneous protein preparation, a shorter form of PcGSTO1 was produced and purified removing the N-terminal extension of the protein (46 amino acids). Nevertheless, the measured enzymatic activities and specificities (full-length and without signal peptide) described in this study were similar for both forms (data not shown).

PcGSTO3 belongs to the subclass II of the GSTs Omega with an FCPFVQ active site quite similar to the FCPFVE of the HsGSTO1-1 (Fig. 1, *bottom*). A second cysteinyl residue is present at the same position into both proteins (Cys<sup>94</sup>, PcGSTO3 numbering). Different attempts have been made to produce heterologously the mutated protein PcGSTO3 C94S, but they remained unsuccessful because the resulting protein was fully insoluble.

**Thiol Content Determination**—To investigate the putative role of the conserved cysteines in the catalytic mechanism of both proteins, site-directed mutagenesis has been used to produce the recombinant proteins in *E. coli*, PcGSTO1 C86S and PcGSTO3 C37S. The thiol content of these proteins has been determined in different conditions as follows: native oxidized corresponding to nonreducing and nondenaturing conditions; denatured oxidized corresponding to a titration in the presence of SDS; reduced and denatured corresponding to a reduction followed by a protein precipitation (see “Experimental Procedures”). A summary of these data is shown in Table 2. Nearly

three and two thiols per protein were titrated for the reduced and denatured PcGSTO1 WT and PcGSTO3 WT proteins, respectively, in accordance with the expected thiol content. Under denaturing conditions and after reduction, the obtained values for PcGSTO1 C86S (around two thiols per protein) and PcGSTO3 C37S (around one thiol per protein) were in agreement with the theoretical values. Comparison of values obtained for denatured PcGSTO1 WT and C86S suggested that the sulfur of Cys<sup>86</sup> was involved at least partially in a DTT-reducible link. For PcGSTO3 WT, one cysteinyl residue was probably involved in a disulfide bridge, whereas for PcGSTO3 C37S, one additional thiol group seemed to be noncovalently bound in the protein.

**Fluorescence Spectrometry and Mass Analysis**—PcGSTO3 contains two tryptophans adjacent to the active site. Given that proximity, we have investigated whether the intrinsic fluorescence of PcGSTO3 could change under reducing (DTT) or oxidizing (GSSG or diamide) conditions (Fig. 2). The native protein displayed an emission spectrum with a maximum at 350 nm characteristic of a fluorescence signal strongly dominated by Trp. Adding 400 μM DTT to native PcGSTO3 immediately led to a strong increase of the fluorescence signal (Fig. 2). The addition of GSSG (10 mM) on previously DTT-reduced PcGSTO3 strongly decreased the fluorescence signal, whereas addition of diamide (10 mM) did not change the reduced PcGSTO3 fluorescence spectrum (data not shown). Similar experiments have been performed with PcGSTO3 C37S; however, in this case no changes in the fluorescence spectrum have been obtained whatever the reductant and oxidant used. In accordance with thiol titrations, these data suggested that the Cys<sup>37</sup> of the native PcGSTO3 WT was glutathionylated and that the tryptophan environment strongly changed after glutathionylation of this active site cysteine.

To confirm this hypothesis, PcGSTO3 WT and PcGSTO3 C37S were analyzed by mass spectrometry. A Q-TOF analysis of the DTT-reduced PcGSTO3 revealed a single protein with a molecular mass of 29,615 Da, which was consistent with a protein where the N-terminal methionine was cleaved. After treatment with GSSG, the mass of the protein increased by around 305 Da, a feature consistent with the formation of one glutathione adduct. This obtained value of 29,921 Da also corresponded to the mass of the native protein, and thus confirmed this latter also possessed a glutathione adduct on Cys<sup>37</sup>. For the mutant PcGSTO3 C37S, the analysis revealed a single protein with a molecular mass of 29,599 Da, which was consistent with a protein without the first methionine and which harbored the Cys to Ser mutation. This result combined with the thiol titration was in agreement with the noncovalent binding of one reduced glutathione to PcGSTO3 C37S.

Similar experiments (fluorescence and mass spectrometry) have been conducted with PcGSTO1 and PcGSTO1 C86S. Even when a tryptophan residue was present near the putative catalytic cysteine, the fluorescence spectrum remained unchanged whatever the reductant (GSH/DTT) or the oxidant (GSSG/diamide) used. A Q-TOF analysis of the DTT-reduced protein revealed a single protein with a molecular mass of 40,187 Da, which was consistent with a protein where the methionine was cleaved. After treatment with GSSG, the mass of the reduced protein increased

## GSTs of *P. chrysosporium*

ScPcpF	-----MGLLIDGVWRDAWYDTKSSG	20
EcYqjG	-----MGQLIDGVWHDWYDTKSTG	20
PcGSTO1	<i>MTTICLRHLSSGFRRQLLQNNTRVYRSFPARMSFTTGSTLQHQEAQLQAKDETVTQLEQ</i>	29
ScPcpF	-----GRFVRKESQYRGGLDAG-----FRGEPGRYHLY	48
EcYqjG	-----GKQFRSASAFRNWLTADGAPGTGTGGFIAEKDRYHLY	58
PcGSTO1	RAAGGASHELQSDISKMKTEDDGSFKRKAASFRNWIQPNG-----DFTPEKGRYHLY	81
	* * * . : : * . : . . * * * . * * * * *	
ScPcpF	AGFACPWHRVLIIMRALKGLEEMISVSMVNAYMGENGWTFPLPGDD-VVPDS--INGADYL	105
EcYqjG	VSLACPWAHRTLIMRKLKGLEPFISVSVNPLMLENGWTFDDSFPGATGDT--LYQNEFL	116
PcGSTO1	VSYACPWATRLLIVRKLKGLEDFIGVTVVSPRMGSNGWPFANVDFPPAADSDPLNNAQHV	143
	.. * * * * * . * * : * * * * * : * * : : * . * * * * * . * * : : : . : :	
ScPcpF	YQVYTAADPTYTGRVTIPILWDKVEKRILNNESEIIRILNSAFDDVG--ALPGDYYP	162
EcYqjG	YQLYLHADPHYSGRVTVPVLWDKKNHTIVSNESAEIIRMFNTAFDALG--AKAGDYYP	173
PcGSTO1	KDLYLKVKPDYDGRFTVPVLWDKHTGTIVNNESEIIRMFNTAFNHLHPEDKAKLDLYPE	203
	: : * . . * * * * * : * * * * * : * * * * * : * * * * * : : * * *	
ScPcpF	EFRPEIDRINARVYETLNGVYRSGFATTQEAYEEAFYPLFDTLDWLEEHLTGREWLVGD	222
EcYqjG	ALQTKIDELNGWIYDVTNNGVYKAGFATSQEAYDEAVAKVFESLARLEQILGQHRYL	233
PcGSTO1	SLRAKIDEVNDWVYDVTNNGVYKSGFASTQKAYEAAVILPFESLDRLEKMLEGQDYLIGG	263
	: : : * * * : * * : * * * * * : * * * * * : * * : * * * * * : * * * * *	
ScPcpF	RLTEADIRLFPFLVRFDAIYHGHFKCNLRRRIAD-YPNLSRLVVGKLAS-HERVAPTINLRH	280
EcYqjG	QLTEADIRLWTTLVRFDPVYVTHFKCDKHRISD-YLNLYGFLRDIYQ-MPGIAETVNF	291
PcGSTO1	QLTEADIRLFTVIRFDPVYVTHFKCNLRTIRDGYPNLHRWMLKLYWGNPAFKDTCNFEH	323
	: * * * * * : * * * * * : * * * * * : * * * * * : : . * * : *	
ScPcpF	AKAHYGSHPVSNPTGIVPVGPAQPLPGLTLQS--- 313	
EcYqjG	IRNHYFRSHKTINPTGIISIGPWQDLDEPHGRDVRFG 328	
PcGSTO1	IKTHYFWSHTFINPHRIVPIGPIPDILPLDA----- 354	
	: * * : * * : * * * * * : : * * * * * :	
PcGSTO3	MPIPDEKIFPHATGAAAKTVEQHQPQELVYAGWFCPFVQRTWIALEERGIPIYQYKEVN	60
PCGSTO4	MSVPDVQIFPHATGAAAKTVEKHQAPQDLVYSGWFCPYVQRTWIALEERNIPYQYVEVN	60
HsGSTO1-1	MSGESARSLGKGSAPPGPVPEGS-----IRIYSMRFCPFAERTRLVLKAKGIRHEVIN	55
	* . . : : : . . . * : * : * * * * * : * * : * * : * * : * * :	
PcGSTO3	PYKKEKHFLDINPKGLVPAIEY-KGKALYESIILCEFLEDAYPNYKPKLLPEDPFERAYA	119
PCGSTO4	PYKKEKHFLDINPKGLVPAIEY-KGRALYESIILCEFLEETYPHTPHLLPADTVERAYA	119
HsGSTO1-1	LKNKPEWFFKKNPFGLVPLENSQQLIYESAITCEYLDEAYPGK--LLPDDPYEKACQ	113
	: * : * . * * * * * : * : * * * * * : * * * * * . * * * * * :	
PcGSTO3	RIWLDYISKS--IIPANFRLIQAQTPEKQEEALADFNKALKQFAEKIKGPYFLGEQFSLV	177
PCGSTO4	RIWIDFVTKN--IVPNFMRLVQAQTTEKQDAARADLYRGLRTFAEKVGRPYFLGDDFSLV	177
HsGSTO1-1	KMILELFSKVPVSLVGSFIRSQNKEDYAGLKEEFRKEFTKLEEVLTNKKTTFFGGNSISMI	173
	: : : : * : : * : : . * . : : : * * : * * :	
PcGSTO3	DIAIVPWIVRDYIIAENRGFKREDIGSKWVEYASKLEKRDVARTSS--EHYAEIYGRYL	235
PCGSTO4	DVALAPWVVRDWWIAENRGYDRAAVGDVWVKRYAEIVEQRDSIRKTSERQHFTDIYGRYL	237
HsGSTO1-1	DYLWIPWFERLEAMKLNEDVHTPKLKLWMAAMKEDPTVSALLTSEKDWQGFLELYLQNS	233
	* : * * . * : * . : . * . . : : . : : : * * :	
PcGSTO3	RDEAQSEAAKATRAGRDIIP 254	
PCGSTO4	RDEAQSEAAKATREGRAIP 256	
HsGSTO1-1	PEACDYGL----- 241	
	: . :	

FIGURE 1. Alignment of PcGSTO1 (top) and PcGSTO3 (bottom) with homologs. Amino acids involved in the catalytic site are labeled. NCBI accession numbers are as follows: PcGSTO1 (EU791894), PcGSTO3 (EU 791893), PcGSTO4 (JGI accession number 7168), *Sphingobium chlorophenolicum* PcpF (AAM 96671), *Escherichia coli* YqjG (NP\_417573), and *Homo sapiens* GSTO1-1 NP\_004823. For PcGSTO1, the putative signal peptide is in *italic*.

TABLE 2

### Number of free thiols in WT and mutated PcGSTOs under various redox conditions

The native column is indicative of the protein thiol content measured after purification. Proteins were denatured using SDS treatment or reduced by DTT and subsequently precipitated using TCA. The thiol content per protein was quantified by 5,5-dithiobis(nitrobenzoic acid). Data are represented as mean  $\pm$  S.D. ( $n = 4$ ).

	Native	Denatured	Reduced and denatured
PcGSTO1	0.46 $\pm$ 0.12	2.50 $\pm$ 0.25	3.0 $\pm$ 0.2
PcGSTO1 C86S	0.15 $\pm$ 0.03	2.0 $\pm$ 0.1	2.04 $\pm$ 0.12
PcGSTO3	0.15 $\pm$ 0.03	1.00 $\pm$ 0.03	1.80 $\pm$ 0.13
PcGSTO3 C37S	0.92 $\pm$ 0.09	1.78 $\pm$ 0.12	1.00 $\pm$ 0.04

by around 306 Da, a feature consistent with the formation of one glutathione adduct. This obtained value of 40,493 Da corresponded also to the mass of the native protein and confirmed this latter also possessed a glutathione adduct. These data combined with thiol titration results were consistent with the potential glutathionylation of Cys<sup>86</sup>, because no glutathione adduct was observed on PcGSTO1 C86S.

**Substrate Specificity between PcGSTO1 and PcGSTO3**—Both PcGSTOs were found to be active in the two classical glutaredoxin assays as follows: the reduction of DHA and HED assays.



The activity of the PcGSTOs in these two assays increased linearly with increasing protein concentrations in the 0–750 and 0–100 nM concentration ranges, respectively. As expected, these observed “thiol transferase” activities are related to the presence of a cysteine in their catalytic site, because the mutants PcGSTO1 C86S and PcGSTO3 C37S are fully inactive in these assays (data not shown). The kinetic analyses revealed catalytic efficiency values ( $k_{\text{cat}}/K_m$ ) similar between both PcGSTOs against DHA, whereas PcGSTO1 exhibited a slightly stronger activity against HED than PcGSTO3 (Table 3). Interestingly, whereas PcGSTO1 C86S remained inactive in all tested assays, PcGSTO3 C37S exhibited a classical glutathione transferase activity detected using CDNB as the substrate. Both PcGSTOs WT were fully inactive against CDNB.

The HsGSTO1.1, ortholog of PcGSTO3 has been shown to possess activity against various substrates, including dimethylarsinate and 2-bromo-4'-phenylacetophenone (12, 13). In both cases, a glutathione adduct occurred followed by a deglutathionylation reaction catalyzed by the enzyme, preventing access to the initial substrate concentrations and also to kinetic parameters. As expected, PcGSTO3 was active in these tests in the presence of glutathione, whereas PcGSTO3 C37S, PcGSTO1 WT, and PcGSTO1 C86S were fully inactive (data not

shown). To determine the catalytic parameters, PAP-SG was synthesized and purified, and the activity against this substrate was followed by analytical HPLC in the presence of GSH (Table 3). In these conditions, only PcGSTO3 is active with a catalytic efficiency of around  $4 \cdot 10^5 \text{ M}^{-1} \text{ s}^{-1}$ , a value greater than that obtained for HsGSTO1-1 ( $10^4 \text{ M}^{-1} \text{ s}^{-1}$ ) (39). The incubation of previously reduced PcGSTO3 in the presence of PAP-SG without addition of GSH led to the formation of phenylacetophenone detected by analytical HPLC (supplemental Fig. 2A). This deglutathionylated product is not observed either in the presence of native purified protein (*i.e.* with a glutathione adduct) or in the presence of PcGSTO3 C37S. These results are in accordance with the involvement of Cys<sup>37</sup> in the deglutathionylation leading to the formation of a disulfide bridge with the glutathione removed from PAP-SG. This hypothesis was confirmed by mass spectrometry analysis of the resulting protein showing the presence of a glutathione adduct (data not shown).

In additional experiments, we tested quinones previously incubated with reduced glutathione as substrates for PcGSTOs. In these conditions, Michael addition occurred leading to the formation of glutathione adducts corresponding to *S*-glutathionyl-*p*-hydroquinones (40, 41). As shown in Fig. 3, addition of reduced glutathione strongly modified the absorbance spectrum of 1,4-benzoquinone with the appearance of a peak maximum at 305 nm in accordance with the formation of *S*-glutathionyl-*p*-benzohydroquinone as described previously (42). Addition of PcGSTO1 led to a strong modification of the absorbance spectrum, with the disappearance of the 305-nm peak and the formation of a peak at 290 nm corresponding to *p*-benzohydroquinone. In contrast, no effect was observed in the presence of PcGSTO1 C86S and PcGSTO3 (data not shown). All these data were consistent with a deglutathionylation reaction of *S*-glutathionyl-*p*-benzohydroquinone. As expected, the activity of PcGSTO1 against *S*-glutathionyl-*p*-benzohydroquinone led to the production of GSSG as revealed by a coupled enzymatic test performed using glutathione reductase (data not shown). Similar data were obtained using menadione (2-methyl-1,4-naphthoquinone), the glutathionylated product (*S*-glutathionyl-naphthoquinone) being characterized by an absorbance peak at 425 nm (data not shown) (43). The obtained results suggested a preference of PcGSTO1 for aromatic substrates in contrast to PcGSTO3. To obtain catalytic parameters,

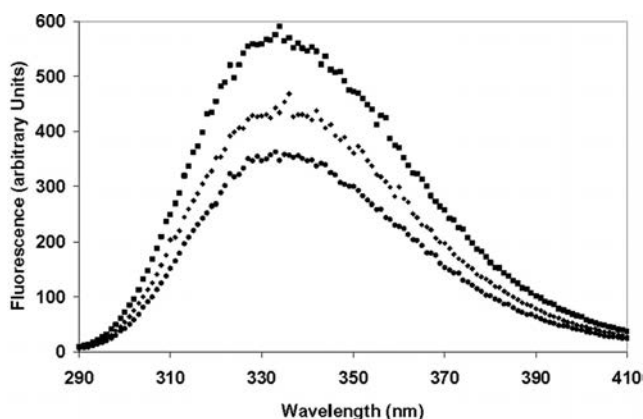


FIGURE 2. Fluorescence spectra of PcGSTO3 under different redox states. Emission spectra of native ( $\blacklozenge$ ), DTT-reduced ( $\blacksquare$ ), and subsequently GSSG-oxidized ( $\bullet$ ) PcGSTO3 (excitation at 290 nm) were recorded with 10  $\mu\text{M}$  protein at 25  $^{\circ}\text{C}$  in TE buffer, pH 8.0. The reduced PcGSTO3 was obtained by a treatment of 10  $\mu\text{M}$  protein with 400  $\mu\text{M}$  DTT. PcGSTO3 was oxidized using 10 mM GSSG after DTT reduction.

TABLE 3  
Kinetic parameters of PcGSTO1 and PcGSTO3 in HED, DHA, and CDNB activity assays

The apparent  $K_m$  value for GSH in the DHA assay was determined using a GSH concentration range of 0.2–4 mM in the presence of 2 mM DHA for PcGSTO1 and PcGSTO3. The apparent  $K_m$  value for GSH was determined for PcGSTO3 C37S in the CDNB assay using a GSH concentration range of 0.1–2 mM in the presence of 2 mM CDNB. The apparent  $K_m$  value for DHA was determined using a concentration range of 0.1–4 mM in the presence of 3 mM GSH. The apparent turnover values ( $k_{\text{cat}}$ ) for HED were determined in the presence of 4 mM GSH and 4 mM HED. The apparent  $K_m$  value for CDNB was determined using a concentration range of 0.1–2 mM in the presence of 2 mM GSH. Concerning PAP-SG and thiodione, the apparent  $K_m$  values were determined using a substrate concentration range of 0.015–4 mM. The apparent  $K_m$  and  $k_{\text{cat}}$  values were calculated by nonlinear regression using the Michaelis-Menten equation. Data are represented as mean  $\pm$  S.D. ( $n \geq 3$ ). ND means not detected. The detection limit was estimated at 0.5 mIU for spectrophotometric tests, whereas 0.5 nmol of either menadione or phenylacetophenone was detected using analytical HPLC.

	GSH	HED	DHA	CDNB	PAP-SG	Thiodione
$K_m$ (mM)						
PcGSTO1	1.65 $\pm$ 0.33		1.84 $\pm$ 0.40	ND	ND	2.00 $\pm$ 0.62
PcGSTO3	1.44 $\pm$ 0.25		0.33 $\pm$ 0.07	ND	0.47 $\pm$ 0.19	ND
PcGSTO3 C37S	0.18 $\pm$ 0.04			0.97 $\pm$ 0.27	ND	ND
$k_{\text{cat}}$ ( $\text{s}^{-1}$ )						
PcGSTO1		2.67 $\pm$ 0.01	1.81 $\pm$ 0.17		ND	19.75 $\pm$ 2.31
PcGSTO3		15.75 $\pm$ 0.26	0.97 $\pm$ 0.06		19.01 $\pm$ 2.89	ND
PcGSTO3 C37S		ND	ND	18.43 $\pm$ 1.90	ND	ND

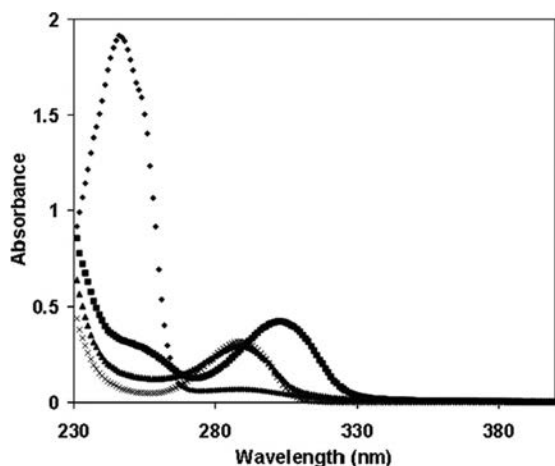


FIGURE 3. PcGSTO1 activity against 5-glutathionyl-*p*-benzoquinone. Absorption spectra have been recorded adding successively in 1 ml at 25 °C in TE buffer, pH 8.0, 1  $\mu$ M *p*-benzoquinone ( $\blacklozenge$ ), 5  $\mu$ M reduced GSH ( $\blacksquare$ ), and 0.28  $\mu$ M PcGSTO1 ( $\blacklozenge$ ). In a separate experiment, as control, absorption spectrum of 1  $\mu$ M *p*-benzoquinone was recorded at 25 °C in TE buffer, pH 8.0.

*S*-glutathionyl-naphthoquinone, also called thiodione, was synthesized and purified, and the activity against this substrate was followed by analytical HPLC (Table 3). The activity of PcGSTO1 against thiodione leads to the formation of a product that elutes with the same retention time as menadione (data not shown). The  $k_{\text{cat}}$  value with this substrate is around 10-fold higher than the values obtained with HED and DHA. No activity against thiodione was detected using PcGSTO1 C86S, confirming the importance of the cysteinyl residue in the catalysis of the deglutathionylation reaction. In additional experiments, no glutathione was added in a mixture containing PcGSTO1 and thiodione, the results of the reaction being analyzed by HPLC (supplemental Fig. 2B). In these conditions, addition of reduced PcGSTO1 led to the formation of menadione. Analysis by mass spectrometry of the resulting protein showed the presence of a glutathione adduct (data not shown). No menadione formation was observed when PcGSTO1 C86S was used.

**Dimerization**—Size exclusion chromatography and dynamic light scattering were used to determine the oligomeric nature of the native forms of PcGSTO1 and PcGSTO3. The retention times of the size exclusion chromatography allowed estimation of the following native molecular masses:  $\sim$ 85,000 Da for PcGSTO1 and  $\sim$ 62,000 Da for PcGSTO3. Similar results have been obtained using dynamic light scattering suggesting that both enzymes were dimers.

**Crystal Structure of PcGSTO1**—The structure of PcGSTO1 has been solved to 2.3 Å resolution by single wavelength anomalous dispersion using crystals of SeMet-PcGSTO1 (Table 1). The crystals belonged to the space group C2, and the asymmetric unit consisted of two polypeptide chains (monomer A, residues Glu<sup>38</sup>–Asp<sup>351</sup>; monomer B, residues Phe<sup>54</sup>–Asp<sup>351</sup>; root mean square deviation, 0.348 Å for 298 superimposed Ca atoms), one GSH molecule not covalently bound and 223 water molecules. Crystal packing analysis using PISA (44) suggested that the asymmetric unit content corresponded to a dimer. The two protomers are related to each other by a noncrystallographic 2-fold symmetry axis.

PcGSTO1 adopts the canonical GST fold composed of an N-terminal thioredoxin-like domain and a C-terminal  $\alpha$ -helix domain. It also possesses several additional features as follows: a long N-terminal coil of 77 residues; a loop that connects  $\beta$ 2 to  $\alpha$ 2 elongated by 20 residues compared with typical GSTs; and a C-terminal extension corresponding to a ninth  $\alpha$ -helix followed by a coil of 20 residues (Fig. 4 and supplemental Fig. 1).

This helix  $\alpha$ 9 (Phe<sup>319</sup> to Ile<sup>333</sup>) is the only secondary structure among these supplementary features. It adds to the  $\alpha$ -helix domain, near the C-terminal end of helix  $\alpha$ 6, so that it also interacts with the thioredoxin-like domain via the short loop that connects the strand  $\beta$ 1 and the  $\alpha$ -helix  $\alpha$ 1 and via the long loop that goes from  $\beta$ 2 to  $\alpha$ 2. Helix  $\alpha$ 9 is a structural characteristic of GST Omega (10) and is also present in GST Tau (45) and GST Delta (46). In these GSTs, the few residues that follow  $\alpha$ 9 (when they exist) go toward the thioredoxin-like domain to interact with the loops  $\beta$ 1- $\alpha$ 1 and  $\beta$ 2- $\alpha$ 2. In PcGSTO1, the long loop observed between  $\beta$ 2 and  $\alpha$ 2 hinders this interaction, so the 20 residues that follow  $\alpha$ 9 in PcGSTO1 form a coil that runs at the opposite side of the thioredoxin-like domain, antiparallel to  $\alpha$ 9.

This C-terminal coil allows the formation of a dimer that completely differs from the usual GST dimer, where the two PcGSTO1 monomers associate exclusively via their  $\alpha$ -helix domains (Fig. 4), with a large buried area (3100 Å<sup>2</sup>) as determined by PISA (44). After  $\alpha$ 9, the 20 C-terminal residues of one monomer mainly interact with residues of the helix  $\alpha$ 5 N-terminal end (H-bonds to the side chains of Gln<sup>231</sup> and Tyr<sup>234</sup>), next with a residue of the  $\alpha$ 6 C-terminal end (Thr<sup>283</sup>), and then with a residue of the loop that follows (Asn<sup>288</sup>), in the other monomer. These direct hydrogen bonds only concern the main chain atoms of the C-terminal end and add to several hydrophobic contacts. Is this dimer a crystal artifact or a physiological entity? The large surface of contact between the protomers and the involvement of the C-terminal coil, specific to PcGSTO1 and related proteins, in the inter-monomer interactions argue for a physiological dimer, as do the dynamic light scattering and size exclusion chromatography results. Furthermore, this dimer, different from the canonical GST dimer, is similar to the dimer assigned by Cuff *et al.* (Midwest Center for Structural Genomics) in the crystal structure of a putative glutathione *S*-transferase from *Corynebacterium glutamicum* recently released by the Protein Data Bank (PDB code 3M1G).<sup>4</sup> The similar dimerization mode appears clearly related to a monomer structure found by DALI (47) to be the closest to the PcGSTO1 structure (root mean square deviation, 1.092 Å for 233 superimposed residues).

PcGSTO1 also possesses an N-terminal extension with respect to other GSTs, by far much longer than the 20 residues observed in GSTs, because 77 residues precede the entrance to  $\beta$ 1. No electron density was observed for the 37 first residues, which were thus supposed to be highly disordered. Residues Glu<sup>38</sup> to Ser<sup>53</sup> were only observed in monomer A (see under “Experimental Procedures”). The main chain of residues 38–53 first runs between the long loop that connects  $\beta$ 2 to  $\alpha$ 2 and the

<sup>4</sup> M. E. Cuff, N. Marshall, G. Cobb, and A. Joachimiak, unpublished results.

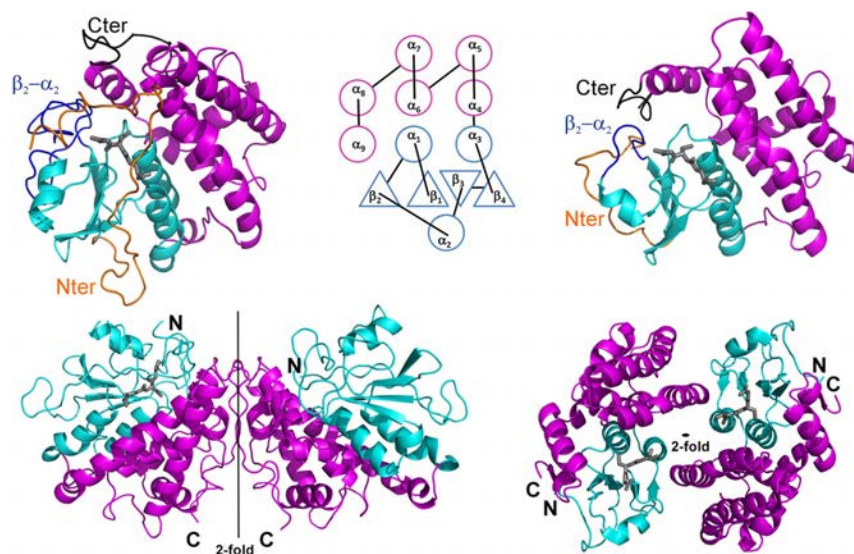


FIGURE 4. Schematic drawing of the PcGSTO1 (left) and HsGSTO1-1 (right) structures. The monomers are displayed at the top separated by their common topology diagram mentioning the secondary structure numbering; the dimers are shown at the bottom, with orientations chosen to render their assembling at most. In each case, the thioredoxin-like domain is colored in cyan, and the  $\alpha$ -helix domain is in purple. In monomers, the other colors emphasize the features that distinguish PcGSTO1 from HsGSTO1-1 as follows: the N-terminal coil in orange, the loop that connects the strand  $\beta_2$  to the  $\alpha$ -helix  $\alpha_2$  in blue, and the C-terminal coil in black. The gray sticks correspond to glutathione, only present in one monomer of PcGSTO1 and in the two monomers of HsGSTO1-1.

loop that goes from  $\alpha_2$  to  $\beta_3$  in the N-terminal thioredoxin-like domain. Then it interacts with the short loop between  $\alpha_4$  and  $\alpha_5$  in the  $\alpha$ -helix domain, where it forms a turn before it enters the region Phe<sup>54</sup>–Arg<sup>77</sup> observed in both monomers. This segment adopts an extended conformation to return toward the N-terminal domain. There it interacts with the loops that connect  $\alpha_2$  to  $\beta_3$  and  $\beta_4$  to  $\alpha_3$ , parallel to it. Then it enters  $\beta_1$ . The conformation of the PcGSTO1 N-terminal end confers it with the original property to form part of the active site, composed of the G-site for GSH binding and the H-site, which receives the hydrophobic part of the glutathionylated substrate.

**The G-site**—Only monomer A displayed an electron density consistent with modeling of a GSH molecule (see under “Experimental Procedures”). However, no major structural difference between monomers A and B was observed that could account for an asymmetric binding of GSH. Nevertheless, in this structure only residues that belong to monomer A contributed to GSH binding.

No disulfide bridge exists between the GSH cysteine side chain and the catalytic residue Cys<sup>86</sup> situated at the N-terminal end of helix  $\alpha_1$  (Fig. 5). The sulfur to sulfur distance is 3.9 Å; the electron density is unambiguous, and careful analysis of the diffraction data reveals no radiation damage that could have disrupted a pre-existing disulfide bond. Superposition of the structures of PcGSTO1 and HsGSTO1-1 (where a disulfide bridge exists) shows a conserved orientation of the catalytic cysteine side chain. On the contrary, a displacement of the GSH S $\gamma$ 2 atom is observed in PcGSTO1, which results from slight differences in the rotation angles around the bond between C $\gamma$ 1 and C $\delta$ 1 in the GSH glutamyl residue and around the bond between C $\alpha$ 2 and C $\beta$ 2 in the GSH cysteinyl side chain. The latter points toward three tyrosine residues (Fig. 5), the hydroxyl groups of which form a hydrogen bond network; in

particular, the GSH S $\gamma$ 2 atom lies at 3.25 Å from the Tyr<sup>223</sup> OH group tightly hydrogen-bonded to the Tyr<sup>326</sup> OH.

In addition to the usual interactions with the main chain of a residue (Val<sup>158</sup>) that precedes the conserved *cis*-proline at the entrance of the strand  $\beta_3$ , the GSH cysteine main chain forms an additional hydrogen bond between its CO group and the N $\epsilon$ 1 atom of Trp<sup>119</sup>, situated in the long loop between  $\beta_2$  and  $\alpha_2$  absent in most GSTs (Fig. 5). In the same way, besides the usual stabilization of the GSH glutamyl residue (main chain by residues Glu<sup>173</sup> and Ser<sup>174</sup> from the loop between  $\beta_4$  and  $\alpha_3$  and aliphatic part of the side chain by hydrophobic contact with Trp<sup>88</sup> from the N-terminal end of  $\alpha_1$ ), an additional hydrogen bond was observed between the GSH glutamyl atom O $\epsilon$ 1 and the N $\eta$ 1 atom of Arg<sup>56</sup> in the long N-terminal extension of PcGSTO1. Finally, the oxygen atoms of the carboxylic group of the GSH glycyl residue form two hydrogen bonds with the side chain of Arg<sup>155</sup> situated in the loop between  $\alpha_2$  and  $\beta_3$  slightly longer than in the other GSTs (Fig. 5), although usually only one bond is observed with a basic residue (48).

To summarize, some of the additional features that PcGSTO1 displays with respect to other GSTs (the long N-terminal end, the long loop between  $\alpha_2$  and  $\beta_2$ , and the few supplementary residues between  $\alpha_2$  and  $\beta_3$ ) contribute to GSH binding. They squeeze GSH so much that only 45 Å<sup>2</sup> of its surface is accessible to solvent (Fig. 5). As a comparison, this value is 148 Å<sup>2</sup> in HsGSTO1.1. The active site seems so enclosed that it raises the question of its accessibility not only to the GSH part of the glutathionylated substrate but also to its hydrophobic part.

**The Putative H-site**—The lack of a structure of PcGSTO1 with a bound *S*-glutathionylated substrate forbids direct experimental delineation of the H-site. The PcGSTO1 surface shows no clear accessible cavity that could accommodate the hydrophobic part of the aromatic *S*-glutathionylated substrate (Fig.

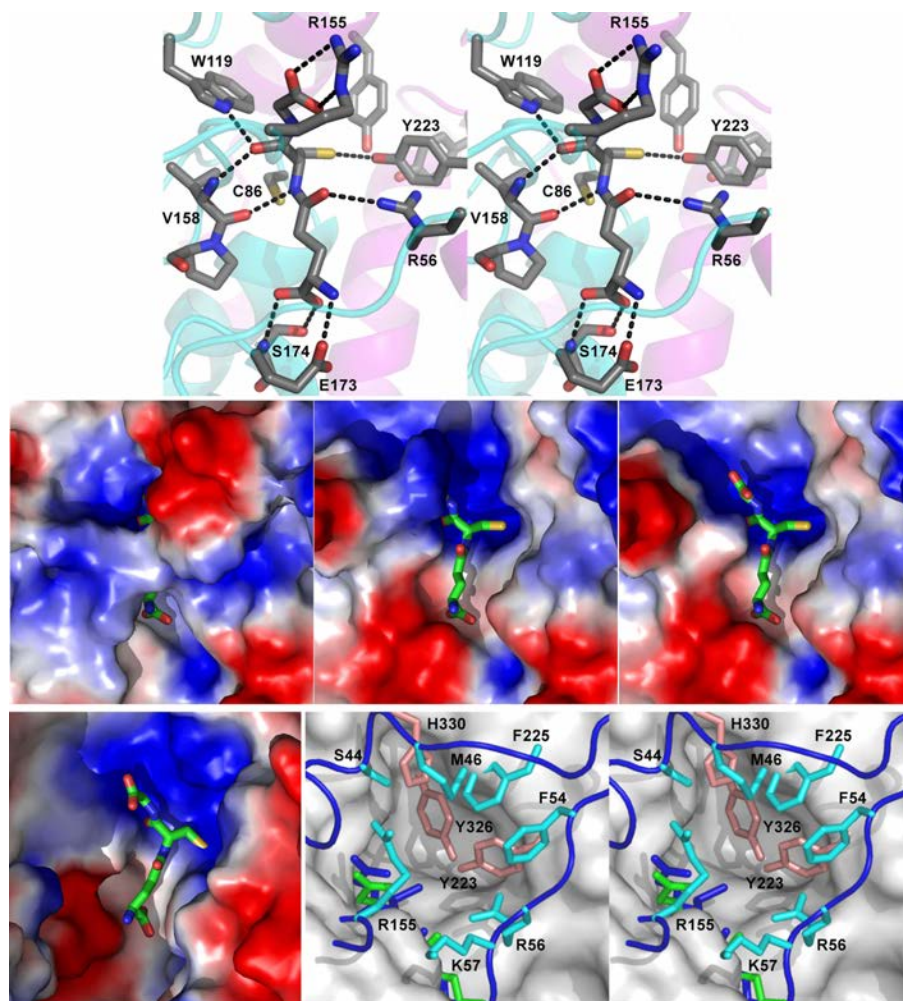


FIGURE 5. **Substrate-binding sites of PcGSTO1 and HsGSTO1-1 (10).** *Top*, stereo view of the hydrogen bond network formed by the active site residues around GSH in PcGSTO1. The schematic trace of the polypeptide main chain is suggested in cyan (thioredoxin-like domain) and purple ( $\alpha$ -helix domain). *Middle*, PcGSTO1 electrostatic surface (blue, positive charge; red, negative charge) of PcGSTO1 in the region of GSH (shown as sticks). *Left panel*, calculated with all the residues observed for monomer A in the crystallographic model (Glu<sup>38</sup>–Asp<sup>351</sup>); *middle panel*, calculated without the N-terminal coil (i.e. with residues Tyr<sup>78</sup>–Asp<sup>351</sup>); *right panel*, same as previous but without Arg<sup>155</sup>. *Bottom*, HsGSTO1-1 active site compared with PcGSTO1. *Left panel*, electrostatic surface calculated for the HsGSTO1-1 dimer, in the region of GSH (shown as sticks). *Right panel*, stereo view of the accessible surface of HsGSTO1-1 (gray) in the region of GSH (green), superimposed to the PcGSTO1 active site containing the GSH (blue sticks). Residues Tyr<sup>223</sup>, Tyr<sup>326</sup>, and His<sup>330</sup> fill the pocket defined by Board *et al.* (10) as the HsGSTO1-1 active site. Residues in cyan and in orange form a second and a third layers of residues that almost close the PcGSTO1 active site.

5). Comparison with HsGSTO1-1 emphasizes the absence of a large binding pocket; the bottom of the HsGSTO1-1 putative H-site (formed by Phe<sup>31</sup>, Cys<sup>32</sup>, Pro<sup>33</sup>, Arg<sup>183</sup>, Trp<sup>222</sup>, and nearby residues (10)) is filled, in PcGSTO1, by Tyr<sup>223</sup> and Tyr<sup>326</sup>, which surround the GSH cysteinyl side chain, and by Arg<sup>56</sup>, Cys<sup>86</sup>, and His<sup>325</sup> (Figs. 1 and 5). Ser<sup>44</sup>, Met<sup>46</sup>, Phe<sup>54</sup>, Lys<sup>57</sup>, Arg<sup>155</sup>, Phe<sup>227</sup>, and His<sup>330</sup> form a narrow, deep cylindrical hole, incompatible with access and binding of a bulky substrate.

Searching in the PDB for X-ray structures of GSTs in complex with *S*-glutathionylated substrates, 1-(*S*-glutathionyl)-(2,4-dinitrobenzene) was found to be the closest analog to *S*-glutathionyl-*p*-benzohydroquinone. It has been used as a substrate by the GSTs of the classes Mu from rat (PDB code 5GST (49)) and human (PDB code 1XWK (50)), Pi from human (PDB code 18GS (51), Sigma from squid (PDB code 1GSQ (52)),

and Alpha from chicken (PDB code 1VF3).<sup>5</sup> In each of these structures, 1-(*S*-glutathionyl)-(2,4-dinitrobenzene) leans its benzene ring against the hydrophobic side chains of the residues that belong to the  $\alpha 4$  C-terminal end, and it often interacts with the first N-terminal residue of  $\alpha 1$  (using PcGSTO1 numbering of secondary structures). Comparable interactions could be obtained using Trp<sup>88</sup>, Tyr<sup>223</sup>, and Phe<sup>227</sup>, but they would require a slightly different positioning of the 1-(*S*-glutathionyl)-(2,4-dinitrobenzene) glutathionyl part with respect to the current position of GSH in PcGSTO1. Indeed, in the structure, GSH appears too shifted toward the  $\alpha$ -helix domain, with its cysteinyl sulfur atom exactly aligned on the  $\alpha 1$  axis. The current position of the long N-terminal loop shows that residues such as Met<sup>46</sup>, Phe<sup>54</sup>, or Arg<sup>56</sup> could also be involved, but once again,

<sup>5</sup> S. C. Lin, Y. C. Lo, M. F. Tam, and Y. C. Law, unpublished results.

the main question concerns access of the substrate to its binding site, which would likely require a movement of the N-terminal loop.

## DISCUSSION

Compared with other organisms, the GSTO class of enzymes is considerably larger in saprotrophic fungi and particularly in *P. chrysosporium*. From primary structure homology, this class could be divided in three different subclasses in fungi (9), one related to plant Lambda GSTs and bacterial GSTs, whereas the second and the third subclasses are related to animal isoforms. In accordance with our first phylogenetic analysis (9), Xun *et al.* (17) have recently proposed from a study of the bacterial homologs of PcpF that the proteins belonging to the Omega subclass I constitute a new class of GSTs named *S*-glutathionyl-(chloro)hydroquinone reductases.

The present x-ray structure of PcGSTO1 confirms that it should be considered as the member of a new class of GSTs. Indeed, the GST canonical fold (a thioredoxin-like domain followed by a  $\alpha$ -helix domain) is complemented by additional specific features as follows: the long N-terminal coil (77 residues) that covers the active site and participates in glutathione binding, as does the long loop (about 30 residues) between the strand  $\beta$ 2 and the  $\alpha$ -helix  $\alpha$ 2, and finally the C-terminal coil (20 residues after the ninth  $\alpha$ -helix) that allows an original dimerization mode where the monomers interact via their  $\alpha$ -helix domain (also observed for a putative GST from *C. glutamicum*).<sup>4</sup> All these features clearly contrast with the HsGSTO1-1 structure (10).

As shown for members of the Omega class, the main characteristic of GHRs is the presence of a conserved cysteine in their catalytic site. From a biochemical point of view, as shown previously (10) and in this study, GHRs and GSTOs share the ability to reduce various substrates such as HED and DHA using glutathione. Concerning PcGSTO1 and PcGSTO3, we propose from fluorescence experiments (for PcGSTO3) and free thiol and mass measurements that the catalytic cysteine is involved in the deglutathionylation activity leading to the formation of a disulfide bridge with the removed glutathione from their substrate. This hypothesis is in accordance with experiments performed using thiodione or PAP-SG as substrates, because the appearance of the deglutathionylated products is observed in the absence of added glutathione, a glutathione-adduct being detected in these conditions on the resulting proteins. Furthermore, these activities require the presence of the catalytic cysteine. This common mechanism has been described at least for monothiol glutaredoxin (23), human GSTOs (10), and bacterial GHRs (16). In PcGSTO3, as shown previously for HsGSTO1-1, a change of the catalytic cysteine to a serine allows reversal of the activity of the protein allowing the mutated protein to catalyze glutathione transfer, and it also confirms the presence of a disulfide bridge between Cys-37 and glutathione in PcGSTO3 WT. For PcGSTO1, biochemical data are also in accordance with the formation of a disulfide bridge between the catalytic Cys-86 and glutathione. Similar results have been obtained for the bacterial ortholog PcpF (16). However, this disulfide bridge is not present in the crystal structure of PcGSTO1. X-rays could have disrupted the bond, but this should have taken place very

rapidly at the beginning of the radiation exposure because no residual electron density corresponding to this bond was observed and the sulfur to sulfur distance is large (3.9 Å), contrary to documented cases where such a phenomenon has been described. At the moment, the reasons for the discrepancy between the PcGSTO1 x-ray structure and the results of thiol titrations concerning the presence of a disulfide bond remain to be elucidated.

As defined by Xun *et al.* (17), the main biochemical characteristic of GHRs is the ability to remove glutathione from *S*-glutathionyl-(chloro)hydroquinone substrates. In this study, we have shown that in contrast to PcGSTO3, PcGSTO1 is active against *S*-glutathionyl-*p*-hydroquinone. Binding of *S*-glutathionyl-*p*-hydroquinone can be modeled from the PcGSTO1 structure, one monomer of which was observed in complex with GSH. Glutathione adopts a conformation very close to that observed in HsGSTO1-1. However, on the contrary to the latter, GSH is more buried in PcGSTO1. It results from new interactions with residues that belong to the features characteristic of PcGSTO1 as follows: Arg<sup>56</sup> in the N-terminal coil, Trp<sup>119</sup> in the loop  $\beta$ 2- $\alpha$ 2, and Arg<sup>155</sup> in the loop  $\alpha$ 2- $\beta$ 3. The additional coils close the cleft usually present between the Trx-like domain and the  $\alpha$ -helix domain, near the loop  $\beta$ 1- $\alpha$ 1 in which is the catalytic residue Cys<sup>86</sup>. Furthermore, the side chains of Tyr<sup>223</sup> and Tyr<sup>326</sup> here shut up the pocket present near the catalytic cysteine in HsGSTO1-1. Even if conformational adaptation occurs upon substrate binding, it would hardly concern these residues because they belong to  $\alpha$ -helices. Hence, a substrate-binding site different from the putative pocket in HsGSTO1-1 has to be considered for the hydroquinone adduct. It could involve Trp<sup>88</sup>, Tyr<sup>223</sup>, and Phe<sup>227</sup>. All the PcGSTO1 residues involved either in the G-site or the putative H-site are conserved in the bacterial homologs (ScPcpF and EcYqjG, see Fig. 1), although they are different or absent in other GSTs (see supplemental Fig. 1). Once again, comparison of the PcGSTO1 and HsGSTO1-1 structures in the light of their substrate-binding site emphasizes the need to consider GHRs as having a new GST structure, different from the Omega class.

In agreement with our previous phylogenetic analysis (9), the biochemical and structural data presented in this study and the recent proposition of Xun *et al.* (17), we propose to change the name of PcGSTO1 in PcGHR1 with regard to the enzymatic function and to define a new structural class of GSTs named GST xi, characterized by their canonical GST fold complemented by specific features and their new dimerization mode. Thus, this paper describes this newly defined *S*-glutathionyl-(chloro)hydroquinone reductase from *P. chrysosporium* as a member of the new structural class xi of GSTs, which would also contain the structure of the putative GST from *C. glutamicum*.

As for PcpF (16), PcGHR1 in *P. chrysosporium* could be involved in the catabolism of pentachlorophenol. In this fungus, the reductive dehalogenation of pentachlorophenol requires indeed at least two enzymes, a glutathione transferase for glutathionylation and an enzyme called glutathione conjugate reductase for deglutathionylation of the resulting *S*-glutathionyl-*p*-hydroquinones (53–55). From the biochemical data presented in this study, we postulated that PcGHR1 or the second

isoform PcGHR2 (formerly named PcGSTO2 (9)) is responsible for this latter reaction. The wide presence of GHRs in bacteria, fungi, and plants suggests an additional physiological role beyond the degradation of the recently introduced pentachlorophenol. Concerning the white rot fungi, their main physiological feature is their ability to mineralize wood components and in particular lignin. It has been suggested recently that chlorination of lignin is a phenomenon occurring through, at least in part, the activity of chloroperoxidase (56, 57). As a result, soils and decayed plant litter contain significant quantities of chlorinated aromatic polymers, which could lead after oxidation via various oxidative systems to the formation of chlorinated quinones. GHRs could play a central role in the intracellular detoxication pathways of such molecules. On the other hand, the significance of the Omega class extension in saprotrophic fungi remains unclear, and the function of the different GSTO isoforms in white rot fungi remains to be elucidated.

*Acknowledgments*—We thank the staff of the BM30A beamline at ESRF for their kind assistance during data collection, Arnaud Gruez for numerous discussions, and Jean-Michel Girardet for technical assistance.

## REFERENCES

- Martinez, D., Larrondo, L. F., Putnam, N., Gelpke, M. D., Huang, K., Chapman, J., Helfenbein, K. G., Ramaiya, P., Detter, J. C., Larimer, F., Coutinho, P. M., Henrissat, B., Berka, R., Cullen, D., and Rokhsar, D. (2004) *Nat. Biotechnol.* **22**, 695–700
- Morel, M., Kohler, A., Martin, F., Gelhaye, E., and Rouhier, N. (2008) *New Phytol.* **180**, 391–407
- Morel, M., Ngadin, A. A., Jacquot, J. P., and Gelhaye, E. (2009) *Adv. Bot. Res.* **52**, 153–186
- Doddapaneni, H., and Yadav, J. S. (2005) *Mol. Genet. Genomics* **274**, 454–466
- Kasai, N., Ikushiro, S., Hirose, S., Arisawa, A., Ichinose, H., Wariishi, H., Ohta, M., and Sakaki, T. (2009) *Biochem. Biophys. Res. Commun.* **387**, 103–108
- Kasai, N., Ikushiro, S., Shinkyo, R., Yasuda, K., Hirose, S., Arisawa, A., Ichinose, H., Wariishi, H., and Sakaki, T. (2010) *Appl. Microbiol. Biotechnol.* **86**, 773–780
- Matsuzaki, F., and Wariishi, H. (2004) *Biochem. Biophys. Res. Commun.* **324**, 387–393
- Schröder, P., Scheer, C. E., Diekmann, F., and Stampfl, A. (2007) *Environ. Sci. Pollut. Res. Int.* **14**, 114–122
- Morel, M., Ngadin, A. A., Droux, M., Jacquot, J. P., and Gelhaye, E. (2009) *Cell. Mol. Life Sci.* **66**, 3711–3725
- Board, P. G., Coggan, M., Chelvanayagam, G., Easteal, S., Jermin, L. S., Schulte, G. K., Danley, D. E., Hoth, L. R., Griffor, M. C., Kamath, A. V., Rosner, M. H., Chrunyk, B. A., Perregaux, D. E., Gabel, C. A., Geoghegan, K. F., and Pandit, J. (2000) *J. Biol. Chem.* **275**, 24798–24806
- Mannervik, B., Board, P. G., Hayes, J. D., Listowsky, I., and Pearson, W. R. (2005) *Methods Enzymol.* **401**, 1–8
- Board, P. G., Coggan, M., Cappello, J., Zhou, H., Oakley, A. J., and Anders, M. W. (2008) *Anal. Biochem.* **374**, 25–30
- Whitbread, A. K., Masoumi, A., Tetlow, N., Schmuck, E., Coggan, M., and Board, P. G. (2005) *Methods Enzymol.* **401**, 78–99
- Burmeister, C., Lüersen, K., Heinick, A., Hussein, A., Domagalski, M., Walter, R. D., and Liebau, E. (2008) *FASEB J.* **22**, 343–354
- Zakharyan, R. A., Sampayo-Reyes, A., Healy, S. M., Tsaprilis, G., Board, P. G., Liebler, D. C., and Aposhian, H. V. (2001) *Chem. Res. Toxicol.* **14**, 1051–1057
- Huang, Y., Xun, R., Chen, G., and Xun, L. (2008) *J. Bacteriol.* **190**, 7595–7600
- Xun, L., Belchik, S. M., Xun, R., Huang, Y., Zhou, H., Sanchez, E., Kang, C., and Board, P. G. (2010) *Biochem. J.* **428**, 419–427
- Garcerá, A., Barreto, L., Piedrafita, L., Tamarit, J., and Herrero, E. (2006) *Biochem. J.* **398**, 187–196
- Herrero, E., Ros, J., Tamarit, J., and Belli, G. (2006) *Photosynth. Res.* **89**, 127–140
- Schenk, P. M., Baumann, S., Mattes, R., and Steinbiss, H. H. (1995) *Bio-Techniques* **19**, 196–198, 200
- Ramakrishnan, V., Finch, J. T., Graziano, V., Lee, P. L., and Sweet, R. M. (1993) *Nature* **362**, 219–223
- Koh, C. S., Navrot, N., Didierjean, C., Rouhier, N., Hirasawa, M., Knaff, D. B., Wingsle, G., Samian, R., Jacquot, J. P., Corbier, C., and Gelhaye, E. (2008) *J. Biol. Chem.* **283**, 23062–23072
- Couturier, J., Koh, C. S., Zaffagnini, M., Winger, A. M., Gualberto, J. M., Corbier, C., Decottignies, P., Jacquot, J. P., Lemaire, S. D., Didierjean, C., and Rouhier, N. (2009) *J. Biol. Chem.* **284**, 9299–9310
- Nickerson, W. J., Falcone, G., and Strauss, G. (1963) *Biochemistry* **2**, 537–543
- Vince, R., Daluge, S., and Wadd, W. B. (1971) *J. Med. Chem.* **14**, 402–404
- Roth, M., Carpentier, P., Kaikati, O., Joly, J., Charrault, P., Pirocchi, M., Kahn, R., Fanchon, E., Jacquamet, L., Borel, F., Bertoni, A., Israel-Gouy, P., and Ferrer, J. L. (2002) *Acta Crystallogr. D Biol. Crystallogr.* **58**, 805–814
- Kabsch, W. (1993) *J. Appl. Crystallogr.* **26**, 795–800
- Collaborative Computational Project No. 4 (1994) *Acta Crystallogr. D Biol. Crystallogr.* **50**, 760–763
- Panjikar, S., Parthasarathy, V., Lamzin, V. S., Weiss, M. S., and Tucker, P. A. (2005) *Acta Crystallogr. D Biol. Crystallogr.* **61**, 449–457
- Schneider, T. R., and Sheldrick, G. M. (2002) *Acta Crystallogr. D Biol. Crystallogr.* **58**, 1772–1779
- Sheldrick, G. M. (2002) *Z. Kristallogr.* **217**, 644–650
- Cowtan, K. (1994) *Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography* **31**, 34–38
- Morris, R. J., Zwart, P. H., Cohen, S., Fernandez, F. J., Kakaris, M., Kirillova, O., Vornrhein, C., Perrakis, A., and Lamzin, V. S. (2004) *J. Synchr. Rad.* **11**, 56–59
- Perrakis, A., Morris, R., and Lamzin, V. S. (1999) *Nat. Struct. Biol.* **6**, 458–463
- Emsley, P., and Cowtan, K. (2004) *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132
- Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) *Acta Crystallogr. D Biol. Crystallogr.* **53**, 240–255
- Gelhaye, E., Rouhier, N., Gérard, J., Jolivet, Y., Gualberto, J., Navrot, N., Ohlsson, P. I., Wingsle, G., Hirasawa, M., Knaff, D. B., Wang, H., Dizen-gremel, P., Meyer, Y., and Jacquot, J. P. (2004) *Proc. Natl. Acad. Sci. U.S.A.* **101**, 14545–14550
- Gelhaye, E., Rouhier, N., Laurent, P., Sautière, P. E., Martin, F., and Jacquot, J. P. (2002) *Physiol. Plant.* **114**, 165–171
- Board, P. G., and Anders, M. W. (2007) *Chem. Res. Toxicol.* **20**, 149–154
- Rao, D. N., Takahashi, N., and Mason, R. P. (1988) *J. Biol. Chem.* **263**, 17981–17986
- Song, Y., Wagner, B. A., Witmer, J. R., Lehmler, H. J., and Buettner, G. R. (2009) *Proc. Natl. Acad. Sci. U.S.A.* **106**, 9725–9730
- Brunmark, A., and Cadenas, E. (1988) *Chem. Biol. Interact.* **68**, 273–298
- Mauzeroll, J., and Bard, A. J. (2004) *Proc. Natl. Acad. Sci. U.S.A.* **101**, 7862–7867
- Krissinel, E., and Henrick, K. (2007) *J. Mol. Biol.* **372**, 774–797
- Thom, R., Cummins, I., Dixon, D. P., Edwards, R., Cole, D. J., and Laphorn, A. J. (2002) *Biochemistry* **41**, 7008–7020
- Oakley, A. J., Harnnoi, T., Udomsinprasert, R., Jirajaroenrat, K., Ketterman, A. J., and Wilce, M. C. (2001) *Protein Sci.* **10**, 2176–2185
- Holm, L., and Rosenström, P. (2010) *Nucleic Acids Res.* **38**, Suppl. W545–W549
- Vararattanavech, A., and Ketterman, A. J. (2007) *Biochem. J.* **406**, 247–256
- Ji, X., Armstrong, R. N., and Gilliland, G. L. (1993) *Biochemistry* **32**, 12949–12954

50. Patskovsky, Y., Patskovska, L., Almo, S. C., and Listowsky, I. (2006) *Biochemistry* **45**, 3852–3862
51. Oakley, A. J., Lo Bello, M., Nuccetelli, M., Mazzetti, A. P., and Parker, M. W. (1999) *J. Mol. Biol.* **291**, 913–926
52. Ji, X., von Rosenvinge, E. C., Johnson, W. W., Tomarev, S. I., Piatigorsky, J., Armstrong, R. N., and Gilliland, G. L. (1995) *Biochemistry* **34**, 5317–5328
53. McCarthy, D. L., Navarrete, S., Willett, W. S., Babbitt, P. C., and Copley, S. D. (1996) *Biochemistry* **35**, 14634–14642
54. Reddy, G. V., and Gold, M. H. (1999) *Biochem. Biophys. Res. Commun.* **257**, 901–905
55. Reddy, G. V., and Gold, M. H. (2001) *Arch. Biochem. Biophys.* **391**, 271–277
56. Ortiz-Bermúdez, P., Hirth, K. C., Srebotnik, E., and Hammel, K. E. (2007) *Proc. Natl. Acad. Sci. U.S.A.* **104**, 3895–3900
57. Ortiz-Bermúdez, P., Srebotnik, E., and Hammel, K. E. (2003) *Appl. Environ. Microbiol.* **69**, 5015–5018

**Supplemental Figure 1: Sequence alignment based on the structural superimposition of PcGSTO1 and GSTs representative of different classes found by DALI (46) to be the closest structural neighbors of PcGSTO1.**

CgGSTO : GST from *Corynebacterium glutamicum* (PDB ID 3M1G, UniProKB ref. Q8NR03, Midwest Center for Structural Genomics, unpublished results); gmGSTT : a tau class GST from *glycine max* (PDB ID 2VO4 (58), UniProKB ref. O49235); DmGSTD: a delta class GST from *Drosophila melanogaster* (PDB ID 3F6F, UniProKB ref. Q9VGA1, Wongsantichon and coworkers, unpublished results); ScGST, a GST from *Saccharomyces cerevisiae* (PDB ID 3ERF (59), UniProKB ref. Q12390); PmGSTB, a beta class GST from *Proteus mirabilis* (PDB ID 2PMT (60), UniProKB ref. P15214); hGSTO, a omega class GST from human (PDB ID 1EEM (10), UniProKB ref. P78417). The sequence alignment was generated with the programs STRAP (61) and ESPript (62). PcGSTO1 and hGSTO secondary structure elements are shown above and below the sequence alignment, respectively.

58. Axarli, I., Dhavala, P., Papageorgiou, A. C., and Labrou, N. E. (2009) *J. Mol. Biol.* **385**, 984-1002
59. Ma, X. X., Jiang, Y. L., He, Y. X., Bao, R., Chen, Y., and Zhou, C. Z. (2009) *EMBO Rep.* **10**, 1320-1326
60. Rossjohn, J., Polekhina, G., Feil, S. C., Allocati, N., Masulli, M., De Illio, C., and Parker, M. W. (1998) *Structure (Lond.)* **6**, 721-734
61. Gille, C., and Frommel, C. (2001) *Bioinformatics* **17**, 377-378
62. Gouet, P., Robert, X., and Courcelle, E. (2003) *Nucleic Acids Res.* **31**, 3320-3323



PcGSTO1 1 10 20 30 40 50 60 70  
PcGSTO1ASFTTGSTLQHQEAQLQGAKDETVTQLEQRAAGGASHELQSDISKMKTEDDGSFKRKAASFRNWIQPNGD  
CgGSTO . . . . .SNAMANTSSDWAGAPQNASADGEFVRDNTNYIDDRIVADVPAGSEPIAQEDGTFH  
gmGSTT . . . . .  
DmGSTD . . . . .  
ScGST . . . . .MNGRGFLIYNG  
PmGSTB . . . . .  
hgSTO . . . . .MSGESARSILGKGSAPP  
hgSTO . . . . . TT

PcGSTO1 β1 α1 β2  
80 90 100 110 120 130 140  
PcGSTO1FTPEKGRYHLYVSYACPWATRTLIVRKLKGLLEDFIGVTVVSPRMGSGNWPFANVDPFPAADSDPLNNAQH  
CgGSTO WPVEAGRYRLVAARACPAWHRVTITRRLGLLENVISLGLTGPTHVRSWTFDLDPNHLDPLVLIQIPRLQDA  
gmGSTT . . .MQDEVVLLDFWPSPFGRMVRVIALAEKGI..KYEYKEDLRN  
DmGSTD . . . . .MDLYRPGSAPCRSVLMTAKALGVE.FDKKTIINTRA . . . . .REQ  
ScGST GEKMKQKMIIDTPAGYPARVRIALAEKNMSSVQFVRINLWK . . . . .GEH  
PmGSTB . . . . .MKLYYTPGS.CSLSPHIVLRETGL..DFSIERIDLRT . . . . .KKTE  
hgSTO GPVPEGSIRIYSMRFPCFAERTRIVLAKAGI..RHEVININLKN . . . . .  
hgSTO . . . . . TT

PcGSTO1 α2 β3 β4 α3  
150 160 170 180 190 200  
PcGSTO1VVKDLVLYKVPDYDGRFTVPLWDKHVGGTIVNNESSEIIRMFNTAFNHLLEDKAKLDLYPESL...RAKI  
CgGSTO YFNRFPDYPRG...ITVPALVEESSKKVVNTNDYPSITIDFNLEWKQFHREGA..PNITYPAEL...REEM  
gmGSTT KSPLLQMNPNVH...KKIPLVLIHN.GKPIIC..ESLIAVQYIEEVWND...RNDPLPSDYPQ.RAQT  
DmGSTD FTPEYLNINPO...HTIPLTHDH.GFALW...ESRAIMVYLVKEYYK...DDPLPKDVPQK.QALI  
ScGST KKPEFLAKNYS...GTVPVLELDDGTLTA..ECTAITEYIDALD...GTPITLTKTPLE.KGVI  
PmGSTB SGKDFLAINPK...GQVPLQLDNDGDIIT..EGVAIVQYLADLKPD...RNTIAPPKALERYHQ  
hgSTO KPEWFFKKNPF...GLVPLVLENSQQLIY..ESAITCEYLDEAYPG...KKLPLDDPYE.KACQ  
hgSTO . . . . . α2 β3 β4 α3

PcGSTO1 α4 α5  
210 220 230 240 250 260  
PcGSTO1DEVNDWVYDVTNNGVYK...SGFASTQ...KAYEAAVILPFESLDRLEKMLEGGDYLI.GGQLT  
CgGSTO APVMKRIFTEVNNGVYR...TGFAGSQ...EAHNEAYKRLWVALDWLEDRLSTRRYLM.GDHIT  
gmGSTT RFWADVVDKKIYDLGRK...IWTSGE...EKEAAKKEFIEALKLLEEQLGDKT.YFG.GDNLG  
DmGSTD NQRLYFDMGTLTKSFSF...YYYPQIFLKKP...ANENYKKIEVAFEFINTFLLEGQTYSG.GDYS  
ScGST HMMNKRAELELLDPVSVYFHHTATPGLGPEVELYQNKEWGLRQRDKALHGMHYFDTVLRRERPYVA.GDSFS  
PmGSTB IEWLNFLASEVHKGYSP...LFSDDTPE...YLPVVKNKLSKFVYINDVLSK.QKVC.GDHFT  
hgSTO KMILELFSKVPISLVGSI...RSQNK...DYAGLKEEPRKFTKLEEVLTNKKTTFFGGRNISIS  
hgSTO . . . . . α4 α5 TT

PcGSTO1 α6 α7 α8 α9  
270 280 290 300 310 320 330  
PcGSTO1EADIRLFVTVIRFDVPVYVTHFKCNLRTIRDGYPNLHRWMRKLYWGNPAFKDTCNFHEHIKTHYFWSHTFIN  
CgGSTO EADIRLRYPTLVRFDAVYHGFKCGRNKITEMPNLWGYLRDLRFQT..PFGFDTTDFTEIKQHYIITHAEIN  
gmGSTT FVDIALVPPFYT.WFKAYETFGTLNI...ESECPRFIAWAKRCLQKESVAKSLPDQQRVYEFIMDLRKKLG  
DmGSTD LADIAFLATVS.TFDVAGFDKRY...ANVARWYENAKKLT.PGWEEWAGCQEFKRYFDN...  
ScGST MADITVIAGLI.FAAIVKLQVPEEC...EALRAWYKRMQQR.PSVKLLLEIRSKSS...  
PmGSTB VADAYLFTLSQ.WAPHVALDLDL...SHLQDYLARIAQR.PNVHSALVTEGLIKE...  
hgSTO MIDYLLWPWFERLEAMKLNCEVDHT...PKLKLWMAAMKED.ETVSALLTSEKDWQGFLELYLQNS  
hgSTO . . . . . α6 α7 α8 α9

PcGSTO1 α10  
340 350  
PcGSTO1PFRIVPIGPIPDILPLDA...  
CgGSTO PTRIVVGPDLSGFATPHGREKLGSGPFABGVTLPGPIPAGEEVKNPEPFQK  
gmGSTT IE...  
DmGSTD . . . . .  
ScGST . . . . .  
PmGSTB . . . . .  
hgSTO PEACDYGL...  
hgSTO . . . . . α10

**Supplemental Table 1 : Primers used in this study.**

The *Nco*I and *Bam*HI cloning restriction sites are underlined in the primers. The mutagenic codons are in bold.

Name	Sequence
GSTO1 full length for	5' <u>CCCCCATGGCT</u> ACCACCATCTGCCTGCGA 3'
GSTO1 truncated for	5' <u>CCCCCATGGCT</u> AGCTTCACAACCGGCAGC 3'
GSTO1 rev	5' <u>CCCCGGATCCT</u> CAGGCGTCGAGAGGGAG 3'
GSTO3 for	5' <u>CCCCCATGGCT</u> CCCATCCCCGACGAG 3'
GSTO3 rev	5' <u>CCCCGGATCCT</u> ACGGGATATCACGGCC 3'
GSTO1 C86S for	5'TACGTCTCGTACGCGTCCCCTTGGGCCACCCGC 3'
GSTO1 C86S rev	5'GCGGGTGGCCCAAGGGGACGCGTACGAGACGTA 3'
GSTO3 C37S for	5'TACGCTGGCTGGTTCAGTCCGTTTGTCCAGAGG 3'
GSTO3 C37S rev	5'CCTCTGGACAAACGGACTGAACCAGCCAGCGTA 3'

## **B.II**      **Etude structurale de la PcGSTX1-C86S en complexe avec des ligands**

### **B.II.1**      **Introduction**

La structure de PcGSTX1 a révélé un site actif extrêmement resserré, dont l'accessibilité est pratiquement nulle. Cette observation, à l'opposé du caractère souvent vaste et accessible du site actif des GST, provient de l'extrémité N-terminale de la chaîne polypeptidique. Longue de 77 résidus dont seuls les 40 derniers ont pu être observés, elle se positionne en une boucle étirée qui obstrue l'accès au site actif, de sorte qu'une autre conformation soit probablement requise pour permettre la fixation d'un substrat. L'étude structurale de complexes entre PcGSTX1 et différents ligands paraissait donc nécessaire, tant pour résoudre la question de l'accessibilité du site actif, que pour identifier la localisation et les caractéristiques du site H.

Les paragraphes qui suivent présentent les résultats obtenus avec différents ligands : le 1-(S-glutathionyl)-2,4-dinitrobenzène (GSDNB), utilisé dans près de 10 structures de GST à tyrosine catalytique de la PDB en tant que produit de l'activité glutathion transférase de ces enzymes sur le CDNB ; l'hexylglutathion, utilisé dans plus d'une trentaine de structures de GST (principalement à sérine ou tyrosine catalytique) pour le caractère hydrophobe du groupement R lié au glutathion de cette molécule commercialisée, et dont la fixation permet de ce fait une première évaluation du site H ; et la ménadione glutathionylée. Ce dernier choix provient du fait que, contrairement à PcGSTO3 dont les caractéristiques biochimiques (démonstrées) et structurales (supposées) la rapprochent des GST Oméga humaines, PcGSTX1 possède une fonction GHR lui permettant de déglutathionyer les hydroquinones (voir paragraphe B.I.2).

Lorsque ces expériences ont débuté, PcGSTX1 était la seule GST Xi à avoir été caractérisée structuralement et biochimiquement. Toutefois, pendant ces travaux, des études comparables ont été publiées concernant les protéines YqjG d'*E. coli* (codes PDB 3R3E, 4G0I, 4G0K, 4G0L) et PcpF de *S. chlorophenicum* (code PDB 4FQU) (Green *et al.*, 2012). Les résultats obtenus sur PcGSTX1 sont ici analysés à la lumière de ces nouvelles structures.

### **B.II.2**      **Matériel et Méthodes**

#### **B.II.2.a**      **Cristallisations**

Les tests de cristallisation de la protéine modifiée PcGSTX1-C86S ont été réalisés, à 4°C, à partir d'une solution protéique concentrée à 40 mg/mL dans du tampon TE (30 mM Tris-HCl, pH

8,0, 1 mM EDTA). La protéine modifiée ne cristallise pas dans les mêmes conditions que la forme de type sauvage. La protéine PcGSTX1-C86S cristallise dans la condition 22 du kit Crystal Screen I (Hampton Research) contenant 0,2 M d'acétate de sodium trihydrate, 0,1M de TRIS-HCL à pH 8,5 et 30 % w/v de polyéthylène glycol 4000. Cette condition est toutefois voisine des conditions de cristallisation de la forme de type sauvage. Les cristaux ont été obtenus avec la méthode microbatch en mélangeant 2  $\mu$ L de la solution protéique avec la solution de cristallisation en équivolume. Les cristaux sont apparus au bout d'une semaine et ont été conservés à 4°C pour faire diffuser à l'intérieur les différents ligands sélectionnés : l'hexylglutathion (produit commercial, Sigma Aldrich), le 1-(S-glutathionyl)-2,4-dinitrobenzène (GSDNB) (préparé au laboratoire suivant le protocole de Busenlehner *et al.* (2007)), la ménadione glutathionylée (ménadione-SG) (préparé par nos collaborateurs suivant le protocole de Vince *et al.* (1971)).

Pour obtenir des cristaux de la protéine mutée PcGSTX1-C86S complexée avec les ligands, nous avons d'abord tenté de travailler en cocrystallisation, à l'époque uniquement dans le cas de la GS-ménadione. Cependant, les cristaux obtenus n'ont pas révélé la présence du substrat dans le site actif, ce qui a été expliqué a posteriori par une probable erreur de gestion de la concentration de la solution stock de GS-ménadione, sans doute utilisée trop diluée. Dans le même temps, des essais ont été menés en tentant de faire diffuser les ligands 5 fois plus concentrés que la protéine dans les cristaux. Pour cela, une solution à 0,2  $\mu$ L de ligand initialement concentré à 50 mM a été introduite dans les gouttes de cristallisation contenant des gros cristaux. Les solutions d'hexylglutathion, de GSDNB et de ménadione-SG sont respectivement incolores, jaunes et orange. Au bout de quelques heures les cristaux deviennent jaunes et orange suite à l'ajout du GSDNB et de la ménadione-SG, respectivement.

Les cristaux ont ensuite été congelés dans une solution de cryoprotection contenant 40 % de solution de cristallisation, 40 % de tampon TE et 20 % de glycérol (sans ligand).

### **B.II.2.b Collecte et traitement des données de diffraction**

La collecte de données a été effectuée à basse température (100 K) sur la ligne PROXIMA1 du synchrotron SOLEIL à Gif-sur-Yvette à la longueur d'onde de 0,97911 Å. Les données ont été traitées au laboratoire avec le logiciel XDS (Kabsch, 2010) pour l'intégration et SCALA (Evans, 2006) pour la mise à l'échelle. Les phases initiales ont été obtenues par remplacement moléculaire à l'aide du logiciel MOLREP (Vagin & Teplyakov, 2010) de CCP4 (Winn *et al.*, 2011). Les affinements ont été conduits avec PHENIX (Adams *et al.*, 2010) et les constructions réalisées avec COOT (Emsley *et al.*, 2010). Les statistiques correspondant à la résolution de ces différentes structures sont rassemblées dans le Tableau 4.

Caractéristiques du cristal			
Complexe	PcGSTX1C86S GS-hexyl	PcGSTX1C86S GSDNB	PcGSTX1C86S GS-ména
Groupe d'espace	C2	C2	C2
Paramètres de maille (Å)	a=168.36 Å b=68.44 Å c=73.24 Å ; β=98.38°	a=167.73 Å b=68.44 Å c=73.28 Å ; β=98.35°	a=168.25 Å b=68.32 Å c=72.95 Å ; β=98.61°
Collectes de données			
Longueur d'onde (Å)	0.97911	0.97911	0.97911
Resolution (Å)	49.04 -2.06 (2.17-2.06)	46.41-1.59 (1.67-1.59)	46.25-1.6 (1.69-1.60)
Rmerge (%)	4.7 (13.8)	2.9 (46.8)	2.6 (39.0)
Réflexions uniques	51171 (7413)	110995 (16010)	107858 (15703)
Complétude (%)	99.9 (99.2)	99.8 (98.6)	100.0 (100.0)
Redondance moyenne	3.7 (3.6)	3.4 (3.2)	3.7 (3.6)
I/σI moyen	15.3 (9.1)	17.2 (2.1)	24.0 (3.0)
Méthode de phasage	Remplacement Moléculaire		
	Modèle : PcGSTX1-Se, code PDB 3PPU		
Facteur de corrélation (CC) (%)	69.2	74.2	71.8
Affinement			
Rwork (%)	16.6 (24.3)	17.5 (31.1)	16.9 (25.9)
Rfree (%)	19.5 (26.9)	19.7 (35.2)	19.1 (29.6)
RMSD			
Longueur de liaison (Å)	0.008	0.032	0.033
Angle de liaison (°)	1.075	2.425	2.538
Facteur B moyen (Å <sup>2</sup> )	26.42	28.72	26.29
Contenu de l'unité asymétrique			
Monomères	2	2	2
Résidus/monomère	A : 301 (47- 351 ) B : 296 (56- 351)	A : 301 (47- 351 ) B : 296 (56- 351)	A : 293 (59- 351 ) B : 293 (59- 351)
Ligands			0.3 GS-ménadione 0.7 GSH 0.7 ménadione
H <sub>2</sub> O	609	779	764
Distribution de Ramachandran			
Favorable (%)	97.53	98.04	97.45
Accepté (%)	2.47	1.63	2.38
Défavorable (%)	0.0	0.33	0.17
<b>Tableau 4 - Résumé des principales statistiques d'enregistrement, résolution et affinement des structures des complexes entre PcGSTX1-C86S et des ligands potentiels.</b> Les valeurs entre parenthèses concernent les statistiques de la plus haute couronne de résolution.			

### **B.II.3     Résultats**

#### **B.II.3.a    Complexes avec le GS-hexyl et le GSDNB**

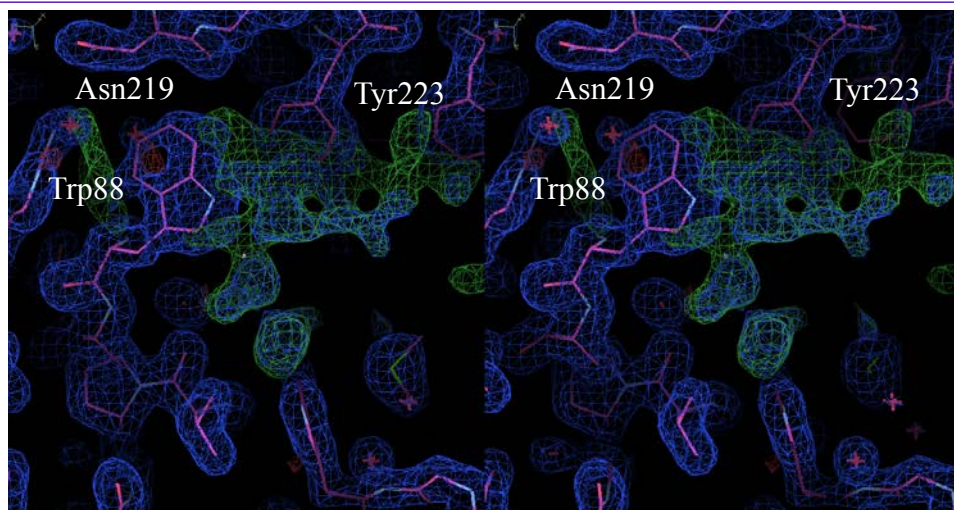
Une étude structurale préliminaire du mutant C86S de PcGSTX1 a été réalisée dans les mêmes conditions que l'enzyme sauvage (données non présentées dans ce manuscrit). Comme attendu de cette mutation ponctuelle classique du résidu catalytique, la structure s'avère comparable, tant au niveau de l'appariement inédit des monomères pour former le nouveau dimère propre à la classe Xi, qu'au niveau du repliement de la chaîne polypeptidique, en particulier la longue extension N-terminale obstruant partiellement le site actif. Contrairement au cas de la forme sauvage qui montrait une molécule de GSH dans le site actif du monomère A, les sites actifs des monomères A et B sont ici vides.

Les cartes de densité obtenues pour PcGSTX1-C86S en complexe avec les substrats hexylglutathion et GSDNB ne révèlent aucunement la présence de densité résiduelle correspondant à ces deux substrats. L'analyse des sites G et H montrent qu'il n'y a ni glutathion ni substrat fixé et qu'il n'y a aucun changement structural au niveau de l'élongation N-terminale.

Ces résultats sont en accord avec les tests biochimiques qui indiquent que PcGSTX1 ne possède aucune activité catalytique de déglutathionylation de ces deux molécules. De même, l'enzyme ne possède pas d'activité glutathion transférase vis-à-vis du CDNB, de sorte que le GSDNB n'en représente pas non plus le produit. PcGSTX1 n'est donc probablement pas suffisamment affine vis-à-vis de ces deux molécules, ne serait-ce que pour permettre le déplacement nécessaire de l'extrémité N-terminale encombrant l'entrée du site actif.

#### **B.II.3.b    Complexe avec la GS-ménadione**

Les cartes de densité électronique de type Fo-Fc et 3Fo-2Fc révèlent de nombreuses modifications de la structure. Tout d'abord, la région du site actif est remplie d'une densité intense très étendue et assez mal définie, correspondant sans ambiguïté à la présence d'un groupement glutathion au niveau du site G d'une part, et à la fixation d'un groupement additionnel très volumineux d'autre part (Illustration 21). Le tryptophane 88 ainsi que les phénylalanines 227 et 332 adoptent une nouvelle position pour accommoder la présence de ces adduits dans le site actif. Conjointement, la densité électronique devient totalement inexistante depuis l'extrémité N-terminale jusqu'au résidu 58.



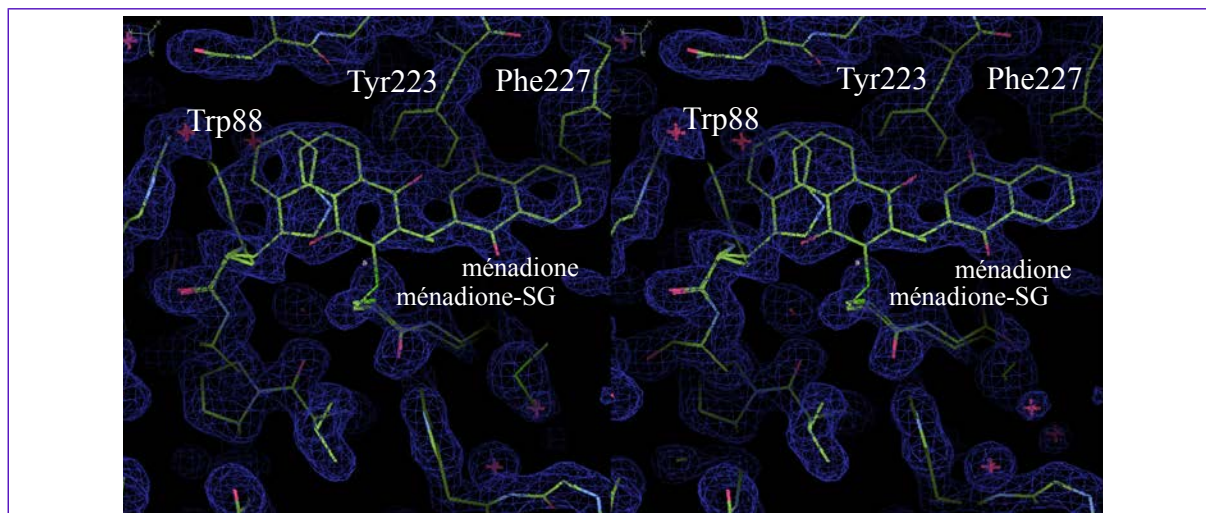
**Illustration 21** - Représentation stéréographique des sites G et H de la PcGSTX1-C86S de *P. chrysosporium* en complexe avec la ménadione glutathionylée.

La densité électronique de type 2Fo-Fc est représentée en bleu (niveau de contour : 1.0  $\sigma$ ), et la densité électronique positive de type Fo-Fc est représentée en vert (niveau de contour 4  $\sigma$ ). Ces densités sont calculées en l'absence de ligand.

Vue réalisée avec le logiciel COOT.

Ces observations valident qu'un complexe a effectivement été obtenu. Cependant, la construction du modèle s'est avérée difficile du fait du caractère très mal défini de la densité électronique au niveau du substrat. Si l'on cherche à maintenir la partie glutathion dans une conformation identique à celle observée chez PcGSTX1, il est impossible de construire la partie ménadione de la molécule, car elle se superposerait partiellement à celle de Trp88. Deux possibilités s'offrent alors, qui ont toutes deux été retenues car elles semblent coexister d'après la densité électronique (Illustration 22). Dans la première situation (correspondant à une occupation de 0,7 pour les éléments concernés), le glutathion est lié d'une manière identique à ce qui a été observé chez PcGSTX1. Trp88 garde sa position, mais alors la partie ménadione ne peut être positionnée liée au glutathion, elle est présente sous forme de ménadione seule qui se place parallèlement aux cycles aromatiques des chaînes latérales de Tyr223 et Phe227. Phe227 présente ainsi une nouvelle conformation, qui entraîne conjointement la nouvelle conformation de Phe332. Cette première situation résulterait de l'instabilité de la GS-ménadione qui se dégrade en quelques heures en solution. Dans la seconde situation (correspondant à une occupation de 0,3), la GS-ménadione est fixée, mais la partie ménadione impose un déplacement de la chaîne latérale de Trp88 dont la nouvelle position empêche la fixation classique de la partie glutathion, qui elle-même adopte une nouvelle conformation au niveau de son résidu  $\gamma$ -glutamyl, entraînant un déplacement de la chaîne latérale de Glu173. La faible qualité de la densité électronique des éléments correspondant à cette deuxième situation ne permet pas le positionnement fiable de la partie ménadione, aussi la construction de cette partie du modèle a été réalisée en s'appuyant sur la structure d'EcYqjG en

complexe avec la GS-ménadione publiée au cours de ces travaux (code PDB 4G0K, Green *et al.*, 2012). La partie ménadione vient alors s'appuyer contre les chaînes latérales de Tyr223 et Pro87.



**Illustration 22** - Représentation stéréographique des sites G et H de la PcGSTX1-C86S de *P. chrysosporium* en complexe avec la ménadione glutathionylée.

La densité électronique de type 2Fo-Fc est représentée en bleu (niveau de contour : 1.0  $\sigma$ ). Ces densités sont calculées en présence du glutathion, de la ménadione, et de la ménadione glutathionylée. Le tryptophane 88 adopterait deux positions possibles (dont une perpendiculaire au plan de la vue représentée ici) selon que la ménadione ou la ménadione glutathionylée est considérée. Vue réalisée avec le logiciel COOT.

La nature du site H est nettement mise en évidence par ces résultats, puisque les parties hydrophobes (aliphatiques ou aromatiques) des résidus Pro87, Tyr88, Tyr215, Asn219, Asn220, Tyr223 et Phe227 forment toute une paroi sur laquelle viennent s'appuyer les cycles de la ménadione ou de la ménadione-SG. Au contraire, à proximité du soufre de la cystéine catalytique 86 (ici mutée en sérine) et du soufre du glutathion se trouve une région fortement polaire, avec les chaînes latérales Tyr223, Tyr281 et Tyr326 (dont les trois hydroxyls partagent des distances interatomiques très courtes), ainsi qu'His325 et His330. Ils permettent la stabilisation d'une molécule d'eau à proximité du soufre du glutathion ou de la ménadione glutathionylée. Ce réseau hydrophile pourrait intervenir lors de la déglutathionylation.

### **B.II.3.c Conclusion**

La préparation des complexes de PcGSTX1 C86S avec différents ligands par méthode de trempage, bien que de succès a priori incertain compte tenu du site actif fermé observé dans l'enzyme natif (voir paragraphe B.I.2), a toutefois connu une certaine réussite, conduisant à des résultats d'une part cohérents avec l'activité de l'enzyme, et d'autre part permettant une première description du site H de PcGSTX1. Ainsi, ni l'hexylglutathion, ni le GSDNB ne se lient à PcGSTX1, en accord avec l'absence d'activité de déglutathionylation de l'enzyme sur ce type de



substrats. Au contraire, l'observation de ménadione-SG dans le site actif est corroborée par l'activité GHR de PcGSTX1.

Comme attendu, la fixation de la ménadione-SG s'accompagne d'un fort bouleversement tant global que local, avec la déstructuration complète d'une grande partie de l'extrémité N-terminale dont la densité électronique disparaît totalement jusqu'au résidu 58, et le déplacement de chaînes latérales s'adaptant à la présence de la partie ménadione du ligand. La mauvaise qualité des cartes de densité électronique ne permet pas l'établissement d'un modèle satisfaisant au niveau du site actif mais les observations montrent un bon accord avec les résultats structuraux publiés par Green *et al.* (2012) sur la protéine YqjG d'*E. coli* en complexe avec la ménadione-SG (code PDB 4G0K).

Pour conclure, à l'heure actuelle, quatre structures de la PDB s'avèrent correspondre à des GST Xi, car elles présentent les éléments structuraux caractéristiques de cette nouvelle classe d'enzymes à cystéine catalytique, en particulier une longue extrémité N-terminale, un allongement de la boucle reliant  $\beta 2$  à  $\alpha 2$ , une hélice additionnelle  $\alpha 9$  suivie d'une boucle C-terminale, et un nouveau mode de dimérisation. L'une d'elle, une GST de *Corynebacterium glutamicum* (code PDB 3M1G), résolue par un consortium de génomique structurale, n'a pas été caractérisée sur le plan biochimique. En revanche, les trois autres (PcGSTX1, la protéine YqjG d'*E. coli*, et PcpF de *Sphingobium chlorophenolicum*) ont des profils d'activité proches, en particulier une activité GHR avérée. Pour deux de ces enzymes (PcGSTX1 dans la présente étude, et EcYqjG (Green *et al.*, 2012)), la structure d'un complexe avec le substrat GS-ménadione a été déterminée. Les observations très proches réalisées au niveau des sites actifs ont permis l'identification des acides aminés définissant le site H, de nature très hydrophobe, ainsi que d'acides aminés intervenants probablement dans l'acte catalytique au niveau de la liaison entre le glutathion et le groupement qu'il porte (triplet de tyrosines et paire d'histidines). Tous ces acides aminés sont soit strictement conservés, soit de nature conservée dans les GST dont les séquences les apparentent à ces trois enzymes étudiés (Green *et al.*, 2012).

## **B.III Etudes structurales des PcGSTFuA**

### **B.III.1 PcGSTFuA1**

#### **B.III.1.a Introduction**

Les glutathion transférases (GST) sont principalement connues pour leur rôle dans les voies de détoxifications des organismes (phase II) où elles catalysent la conjugaison du glutathion avec des molécules toxiques endogènes ou exogènes. Certaines GST peuvent également catalyser la réaction inverse et d'autres présentent la capacité de fixer sélectivement des petites molécules organiques telles que les tétra-pyrroles dans les mammifères et les plantes. Cette dernière propriété est qualifiée de « ligandine ».

Le séquençage massif et récent de génomes de champignons a permis l'analyse de l'étendue des GST dans ce règne, GST peu étudiées en comparaison de celles des règnes des animaux et des plantes. Aujourd'hui au moins sept classes de GST fongiques cytosoliques ont été mises définies : les MAK16, les EF1B $\gamma$ , les Ure2p, les Oméga, les S-glutathionyl-(chloro)-hydroquinone réductases (ou GST Xi), les GTT et les GTE pour 'Glutathion Transférase Ethérase'. Cette dernière et nouvelle classe a été initialement appelée ainsi car son plus proche homologue connu caractérisé est l'éthérase LigE de la bactérie du sol *Sphingobium* sp. SYK-6. Les LigE, F et G ont été étudiées pour leur capacité à dégrader de petits fragments de lignine en coupant une liaison éther qui est le lien le plus abondant entre les monomères de ce polymère complexe constitutif du bois. Cette nouvelle classe de GST fongiques est particulièrement étendue dans les organismes impliqués dans la dégradation du bois tels que *Postia placenta* qui possède 11 isoformes ou *Phanerochaete chrysosporium* avec 5 isoformes. Ce dernier est un champignon saprophyte de la pourriture blanche qui dégrade et minéralise complètement la lignocellulose. Ce travail présente pour la première fois la caractérisation biochimique et cristallographique d'un membre de cette nouvelle classe de GST analogues à l'éthérase LigE bactérienne à savoir la GST5118 de *P. chrysosporium*, ultérieurement renommée PcGSTFuA1.

#### **B.III.1.b Résultats et discussion**

Les structures cristallographiques de la GST5118 de *P. chrysosporium*, en l'absence et en présence de glutathion, ont été résolues grâce à la méthode SAD à partir d'une forme sélénométhionylée de l'enzyme. La GST5118 adopte une structure quaternaire dimérique différente de celle habituellement observée pour cette famille de protéine. Les domaines N- et C-terminaux adoptent respectivement les repliements ( $\beta_1\alpha_1\beta_2\alpha_2\beta_2'\beta_2''\beta_3\beta_4\alpha_3$ ) et ( $\alpha_4\alpha_4'\alpha_5\alpha_6\alpha_7\alpha_8$ ). L'épingle à cheveux  $\beta_2\beta_2''$  est observée pour la première fois dans une GST, celle-ci couvre le site du glutathion (site G)

et change de conformation lors de sa fixation. De plus ce motif structural gêne la formation du dimère canonique des GST. La présence de l'hélice  $\alpha_4$  est remarquable car celle-ci a seulement été observée dans la glutarédoxine 2 qui est considérée comme l'ancêtre commun des GST.

Les tests enzymatiques n'ont pas mis en évidence d'activité éthérase avec le fragment de lignine dégradé par les Lig bactériennes. La GST5118 présente une activité glutathion transférase avec les phénéthyl-isothiocyanate et 4-nitrophényl-butyrates et non avec le substrat classique des GST c'est-à-dire le CDNB. De plus elle possède une activité avec le CMFDA en présence ou en l'absence de glutathion. Le CMFDA est un substrat intéressant car il libère de la fluorescence via une réaction estérase. Le résidu catalytique putatif de l'enzyme S22 a été remplacé soit par une alanine soit par une cystéine. L'enzyme modifiée S22A perd son activité pour le phénéthyl-isothiocyanate et devient active avec le CDNB (activité GSH transférase) et également l'hydroxyéthyl-disulfure (activité thioltransférase) et le cumène hydroperoxyde (activité peroxydase). La forme S22C est inactive avec tous les substrats testés.

L'acide 8-Anilino-1-naphtalène sulfonique (ANS) est une sonde fluorescente qui se fixe dans les poches hydrophobes des protéines. Les expériences réalisées avec cette sonde ont permis de montrer le site actif de la GST5118 accueille l'ANS avec son groupement sulfate probablement située dans la poche anionique mise en évidence dans les structures cristallographiques. De plus et de façon inattendue, il apparaît que l'ANS gêne la fixation de molécules qui se fixent dans le site G (glutathion ou molécules glutathionylées) alors que ce n'est pas le cas pour des molécules qui se fixent uniquement de le site H de l'enzyme. Les sondes fluorescentes ANS et CMFDA ont permis de mettre en évidence une propriété « ligandine » pour la GST5118 envers des composés issus de la dégradation du bois. Comme la sonde ANS, ces composés occupent plutôt le G de la protéine que le site H.

L'ensemble de ces données biochimiques et structurales nous a conduit à définir une nouvelle classe fonctionnelle dans la superfamille des GST. Cette nouvelle classe, spécifique au règne des champignons, a été appelée GSTFuA pour Fungal specific class A.

Toutes ces informations sont précisées dans le texte de la publication de Mathieu *et al.* parue en 2012 dans *J. Biol. Chem.* (vol. 287, p. 39001-39011), dont le texte est présenté ci-après.

### **B.III.1.c Publication PcGSTFuA1**

Voir texte page suivante.

# Characterization of a *Phanerochaete chrysosporium* Glutathione Transferase Reveals a Novel Structural and Functional Class with Ligandin Properties<sup>\*S</sup>

Received for publication, July 20, 2012, and in revised form, September 7, 2012. Published, JBC Papers in Press, September 24, 2012, DOI 10.1074/jbc.M112.402776

Yann Mathieu<sup>†S¶1</sup>, Pascalita Prosper<sup>||\*\*1</sup>, Marc Buée<sup>S</sup>, Stéphane Dumarçay<sup>††</sup>, Frédérique Favier<sup>||\*\*</sup>, Eric Gelhaye<sup>†S</sup>, Philippe Gérardin<sup>††</sup>, Luc Harvengt<sup>¶</sup>, Jean-Pierre Jacquot<sup>†S</sup>, Tiphaine Lamant<sup>†S</sup>, Edgar Meux<sup>†S</sup>, Sandrine Mathiot<sup>||\*\*</sup>, Claude Didierjean<sup>||\*\*2</sup>, and Mélanie Morel<sup>†S3</sup>

From the <sup>†</sup>Université de Lorraine, Interactions Arbre-Microorganismes, UMR 1136, Institut Fédératif de Recherche 110 EFABA, Vandoeuvre-lès-Nancy F-54506, France, <sup>S</sup>Institut National de la Recherche Agronomique, Interactions Arbre-Microorganismes, UMR 1136, Vandoeuvre-lès-Nancy F-54506, France, <sup>||</sup>Université de Lorraine, CRM2, UMR 7036, Vandoeuvre-lès-Nancy F-54506, France, <sup>\*\*</sup>CNRS, CRM2, UMR 7036, Vandoeuvre-lès-Nancy F-54506, France, <sup>††</sup>Université de Lorraine, Laboratoire d'Études et de Recherche sur le Matériau Bois, EA 1093, Vandoeuvre-lès-Nancy F-54506, France, and the <sup>¶</sup>Laboratoire de Biotechnologie, Pole Biotechnologie et Sylviculture Avancée, Institut Technologique Forêt Cellulose Bois-construction Ameublement, Campus Forêt-Bois de Pierroton, 33610 Cestas, France

**Background:** GSTs are detoxification enzymes poorly characterized in fungi.

**Results:** *GSTFuA1* possesses a unique three-dimensional structure and binds wood degradation compounds at or near the glutathione binding pocket.

**Conclusion:** This GST is a new fungal isoform that we name *GSTFuA1*.

**Significance:** GSTs with binding properties could be of great interest in various biotechnological applications.

Glutathione *S*-transferases (GSTs) form a superfamily of multifunctional proteins with essential roles in cellular detoxification processes. A new fungal specific class of GST has been highlighted by genomic approaches. The biochemical and structural characterization of one isoform of this class in *Phanerochaete chrysosporium* revealed original properties. The three-dimensional structure showed a new dimerization mode and specific features by comparison with the canonical GST structure. An additional  $\beta$ -hairpin motif in the N-terminal domain prevents the formation of the regular GST dimer and acts as a lid, which closes upon glutathione binding. Moreover, this isoform is the first described GST that contains all secondary structural elements, including helix  $\alpha 4'$  in the C-terminal domain, of the presumed common ancestor of cytosolic GSTs (*i.e.* glutaredoxin 2). A sulfate binding site has been identified close to the glutathione binding site and allows the binding of 8-anilino-1-naphthalene sulfonic acid. Competition experiments between 8-anilino-1-

naphthalene sulfonic acid, which has fluorescent properties, and various molecules showed that this GST binds glutathionylated and sulfated compounds but also wood extractive molecules, such as vanillin, chloronitrobenzoic acid, hydroxyacetophenone, catechins, and aldehydes, in the glutathione pocket. This enzyme could thus function as a classical GST through the addition of glutathione mainly to phenethyl isothiocyanate, but alternatively and in a competitive way, it could also act as a ligandin of wood extractive compounds. These new structural and functional properties lead us to propose that this GST belongs to a new class that we name GSTFuA, for fungal specific GST class A.

Glutathione *S*-transferases (GSTs) are a superfamily of proteins widespread in animals, plants, fungi, and bacteria. From a functional point of view, GSTs usually catalyze glutathione (GSH) transfer onto hydrophobic molecules (glutathionylation activity) or GSH removal from specific substrates (deglutathionylation) (1). These enzymes have broad substrate acceptance, although each GST possesses its own specific catalytic profile. Most GSTs are dimeric proteins. Each monomer is composed of a conserved thioredoxin domain containing the GSH binding pocket (G site) and a more variable  $\alpha$ -helical domain (H site) containing the binding site for the GSH acceptor substrate (2). Most studies have focused on the GSH-dependent catalytic activities involved both in detoxification processes and endogenous metabolism. However, some GSTs have also been identified as proteins that selectively bind organic anions such as tetrapyrroles in mammals and plants (3–5). This “ligandin” property has been defined as the capacity of the protein to bind nonsubstrate ligands (3). In plants, it could be involved in the intracellular transport of hydrophobic compounds, such as pigments, and in temporary storage of phytohormones (6, 7). Fun-

\* This work was supported by Convention industrielle de formation par la recherche (CIFRE) and Agence Nationale de la Recherche (ANR) Research Grant ANR-09-BLAN-0012; the Ministère de l'Enseignement Supérieur, de la Recherche, et de la Technologie; the Institut National de la Recherche Agronomique (INRA); and CNRS.

<sup>S</sup> This article contains supplemental Table 1 and Figs. S1–S11.

The atomic coordinates and structure factors (codes 4F03 and 4G19) have been deposited in the Protein Data Bank (<http://www.pdb.org>).

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) JQ974949

<sup>1</sup> Both authors contributed equally to this work.

<sup>2</sup> To whom correspondence may be addressed: Université de Lorraine, UMR 7036 CRM2 Institut Jean Barriol, Faculté des Sciences et Technologies, BP 70239, 54506 Vandoeuvre-lès-Nancy Cedex, France. Tel.: 33-3-83-68-48-79; E-mail: Claude.Didierjean@crm2.uhp-nancy.fr.

<sup>3</sup> To whom correspondence may be addressed: Université de Lorraine, UMR 1136 “Interactions Arbres/Micro-organismes,” Faculté des Sciences et Technologies BP 70239, 54506 Vandoeuvre-lès-Nancy Cedex, France. Tel.: 33-3-83-68-42-28; E-mail: mmorel@scbiol.uhp-nancy.fr.

## Ligandin Properties of a New Fungal GST

gal GSTs have been less studied than their plant or animal counterparts. Nevertheless, the recent burst of fungal genome sequencing programs has highlighted evolutionary specificities of GSTs in these organisms (8–10). Currently, at least seven different classes of cytosolic GSTs have been defined in fungi: MAK16, EFB $\gamma$ , Ure2p, Omega, *S*-glutathionyl-(chloro)hydroquinone reductase, GTT, and etherase-like. Although some isoforms from the Ure2p, Omega, *S*-glutathionyl-(chloro)hydroquinone reductase, and GTT have been characterized in yeast (11–13), *Aspergillus* sp. (14, 15), and *Phanerochaete chrysosporium* (9, 16), the etherase-like class has never been studied to date. Etherase-like GSTs are fungal specific, although they show weak homology with LigE, a bacterial protein responsible for the cleavage of  $\beta$ -aryl ether linkages of lignin in *Sphingobium* sp. SYK-6 (8, 17). The etherase-like class is expanded in wood decaying fungi, because, for instance, 11 etherase-like encoding genes are present in the genome of *Postia placenta*. Although less expanded, they are, however, present in pathogenic fungi as *Aspergillus* sp. or *Fusarium* sp. for example (8).

*P. chrysosporium* is a white rot fungus able to completely mineralize lignocellulose by degrading recalcitrant compounds. During the wood decaying process, phenolic acids and phenolic aldehydes are released. These compounds are problematic in biofuel production, but at the same time, they can have beneficial properties for human health. As an example, vanillin possesses antimicrobial activity, antioxidant properties, antimutagenic effects, and chemoprotective activity in inflammation and cancer (for a review, see Ref. 18). Syringaldehyde is also a naturally occurring aromatic compound, which exhibits both a detrimental action on cellulose hydrolysis and fermentation processes and a good antioxidant activity, being 2- and 10-fold more effective than quercetin and trolox, respectively (19).

In this study, we describe for the first time the biochemical and structural characterization of a GST belonging to the previously named “etherase-like” class in *P. chrysosporium* and demonstrate the ligandin property of this enzyme toward compounds derived from lignin degradation.

### EXPERIMENTAL PROCEDURES

**Materials**—Hydroxyethyl disulfide was from Pierce.  $\alpha$ -*O*-Methylumbelliferyl- $\beta$ -hydroxypropiovanillone was a gift from Dr. Masai (Nagaoka University of Technology, Japan). *S*-(phenylacetophenone)-glutathione and 2-methyl-*S*-glutathionyl-naphtoquinone were synthesized as described previously (16). *para*-Nitrophenyl sulfate potassium salt was from Acros Organics. 5-Chloromethylfluorescein diacetate (CMFDA)<sup>4</sup> was from Invitrogen. All other reagents were from Sigma-Aldrich.

**Cloning of GST5118, GST5118-S22A, and GST5118-S22C**—The enzyme studied here is referred to as GST5118 based on its protein identification number in the Joint Genome Institute database. The open reading frame sequence encoding *P. chrysosporium* GST5118 was amplified from a *P. chrysosporium* cDNA library using GST5118 forward and reverse primers (5'-

CCCCCATGGCTCAGCCCATCGTGTT-3' and 5'-CCCGGATCCCTATACATCAACCTGCT3', respectively) and cloned into the NcoI and BamHI restriction sites (underlined in the primers) of pET-3d (Novagen). The amplified sequence encodes a protein in which an alanine has been inserted after the initiator methionine to improve protein production. Two mutants of the putative catalytic Ser-22 into Ala or Cys residues were also generated using two complementary mutagenic primers (supplemental Table 1).

**Expression and Purification of the Recombinant Proteins**—For protein production, the *Escherichia coli* BL21(DE3) strain, containing the pSBET plasmid, was co-transformed with the different recombinant plasmids (20). Cultures were progressively amplified up to 2 liters in LB medium supplemented with ampicillin and kanamycin at 37 °C. Protein expression was induced at exponential phase by adding 100  $\mu$ M isopropyl  $\beta$ -D-thiogalactopyranoside for 4 h at 37 °C. The cultures were then centrifuged for 15 min at 4400  $\times$  g. The pellets were resuspended in 30 ml of TE NaCl (30 mM Tris-HCl, pH 8.0, 1 mM EDTA, 200 mM NaCl) buffer. Cell lysis was performed on ice by sonication (3  $\times$  1 min with intervals of 1 min), and the soluble and insoluble fractions were separated by centrifugation for 30 min at 27,000  $\times$  g at 4 °C. All subsequent steps were performed in the cold. The soluble part was then fractionated with ammonium sulfate in two steps, and the protein fraction precipitating between 40 and 80% of saturation contained the recombinant protein, as estimated by 15% SDS-PAGE. The protein was purified by size exclusion chromatography after loading on an ACA44 (5  $\times$  75-cm) column equilibrated in TE NaCl buffer. The fractions containing the protein were pooled, dialyzed by ultrafiltration to remove NaCl, and loaded onto a DEAE-cellulose column (Sigma) in TE (30 mM Tris-HCl, pH 8.0, 1 mM EDTA) buffer. The proteins were eluted using a 0–0.4 M NaCl gradient. Finally, the fractions of interest were pooled, dialyzed, concentrated by ultrafiltration under nitrogen pressure (YM10 membrane; Amicon), and stored in TE buffer at –20 °C. Purity was checked by SDS-PAGE. Protein concentrations were determined spectrophotometrically using a molar extinction coefficient at 280 nm of 68,870 M<sup>-1</sup> cm<sup>-1</sup>.

For the production of the selenomethionine-substituted GST5118, a methionine auxotroph strain, BL21(DE3)Met<sup>-</sup>, was cotransformed with pET-GST5118 and pSBET. Cultures were done as described previously (21), and the labeled protein was purified following a procedure identical to that described for unlabeled GST5118.

**Activity Measurements**—The activity measurements of GST5118, GST5118-S22C, and GST5118-S22A proteins in thiol transferase activity with a hydroxyethyl disulfide assay or for reduction of dihydroascorbate were performed as described by Couturier *et al.* (22). The GSH transferase activity was assessed with phenethyl isothiocyanate (phenethyl-ITC) prepared in 2% (v/v) acetonitrile, 1-chloro-2,4-dinitrobenzene (CDNB), and 4-nitrophenyl butyrate (PNP-butyrate) prepared in DMSO. For these three substrates, the reactions were monitored at 274, 340, and 412 nm, respectively, following the increase in absorbance arising from the formation of the *S*-glutathionylated adduct. The reactions with CDNB and PNP-butyrate were performed in 100 mM phosphate buffer, pH 7.5, in

<sup>4</sup> The abbreviations used are: CMFDA, 5-chloromethylfluorescein diacetate; ANS, 8-anilino-1-naphthalene sulfonic acid; CDNB, 1-chloro-2,4-dinitrobenzene; ITC, isothiocyanate; PNP-butyrate, paranitrophenyl butyrate.

the presence of GSH (5 mM), whereas the reaction with phenethyl-ITC was performed at pH 6.5 with an identical GSH concentration. Esterase activity was measured in microplates using CMFDA as substrate, which releases fluorescence upon activation by esterases (23). Experiments were performed in 50 mM phosphate buffer, pH 8.0, in a total volume of 200  $\mu$ l. The reactions were started by the addition of the purified enzyme, and fluorescence was measured every minute for 1 h with excitation at 485 nm and emission at 535 nm on a VICTOR<sup>TM</sup> X5 plate reader (PerkinElmer Life Sciences). Catalytic parameters were calculated using the GraphPad<sup>®</sup> software from steady state experiments performed using 0–25  $\mu$ M CMFDA and 5 mM GSH.

Peroxidase activities were monitored as follows. 1 mM peroxide (hydrogen peroxide, *tert*-butyl hydroperoxide, and cumene hydroperoxide) in 30 mM Tris-HCl, pH 8.0, was incubated in the presence of 2 mM GSH, 200  $\mu$ M NADPH, 0.5 IU of glutathione reductase. The activity was followed by monitoring the decrease in absorbance arising from NADPH oxidation in this coupled enzyme assay system, showing the formation of oxidized glutathione (GSSG). The reactions were started by the addition of the purified enzyme and monitored with a Cary 50 UV-visible spectrophotometer (VARIAN). To determine the catalytic properties of the enzyme, steady state assays were performed using variable substrate concentrations (from 10  $\mu$ M to 10 mM), and the catalytic parameters were calculated using the GraphPad<sup>®</sup> software. Etherase activity was assessed according to Masai *et al.* (17). Sulfatase activity was performed with *para*-nitrophenyl sulfate potassium salt as described by Kim *et al.* (24).

**Mass Spectrometry Analysis**—GST5118 wild type and mutants were analyzed by Q-TOF MS as described by Koh *et al.* (25). When oxidation or reduction was required, proteins were treated at 25 °C for 60 min, whatever the reductant (dithiothreitol (DTT)) or the oxidant (GSSG) used. Excess of oxidant/reductant was removed using a Sephadex<sup>®</sup> G25 (GE Healthcare) gel filtration column.

**8-Anilino-1-naphthalene Sulfonic Acid (ANS) Binding and Ligand Screening**—All binding and competition experiments were performed in TE buffer, pH 8.0. ANS binding onto GST5118 and GST5118-S22C (DTT-treated or not) was investigated by monitoring fluorescence upon the addition of 50  $\mu$ M ANS to 3  $\mu$ M protein using a Cary Eclipse (VARIAN) fluorescence spectrophotometer. The excitation wavelength was set at 385 nm, and the emission spectra were recorded from 400 to 700 nm. Additional tryptophan-based fluorescence experiments were performed with GST5118 and GST5118-S22C in the presence or absence of 50  $\mu$ M ANS. After excitation at 290 nm, emission spectra were recorded from 305 to 560 nm. The ANS binding site was investigated in the same conditions with 1 mM GSH, 5 mM phenethyl-ITC, or 50  $\mu$ M CMFDA.

The correlation between ANS concentration and fluorescence yield was obtained by incubating 100  $\mu$ M GST5118 with 1  $\mu$ M ANS, giving the maximal value for 1  $\mu$ M ANS bound onto the protein. Fluorescence from samples containing ANS alone was subtracted as a background. GST5118 was then incubated at a concentration of 3  $\mu$ M with increasing concentrations of ANS ranging from 0 to 1 mM. To determine the dissociation

constant of ANS, the following equation (Equation 1) was applied.

$$y = B_{\max} \times [\text{Free ANS}] / (K_d + [\text{Free ANS}]) \quad (\text{Eq. 1})$$

where  $y$  represents the specific binding (*i.e.* the concentration of ANS bound/protein) and  $B_{\max}$  is the number of maximum specific binding sites.

Competition experiments were performed in a total volume of 250  $\mu$ l containing 3  $\mu$ M enzyme, 50  $\mu$ M ANS, and various putative ligands (from 1  $\mu$ M to 10 mM). Fluorescence emission from protein-free samples containing ANS and potential ligand were subtracted as background. A decrease in fluorescence, indicating competition between ligand and ANS, was measured after 10 min of incubation with excitation at 355 nm and emission at 460 nm on a VICTOR<sup>TM</sup> X5 plate reader (PerkinElmer Life Sciences).  $IC_{50}$  values were obtained by fitting data to Equation 2,

$$Y = Bottom + (Top - Bottom) / (1 + 10^{(\text{Log[Ligand]} - \text{Log } IC_{50})}) \quad (\text{Eq. 2})$$

where  $Y$  represents the fluorescence signal observed after background subtraction,  $Top$  is the fluorescence signal of ANS bound onto GST5118 without any ligand, and  $Bottom$  is the fluorescence signal of ANS at the highest ligand concentration.

**Inhibition Kinetics**—Competition tests between ANS and phenethyl-ITC or CMFDA were performed by measuring GSH transferase and esterase activities. Activity with phenethyl-ITC was monitored in a total volume of 500  $\mu$ l containing 50  $\mu$ M GSH, 50  $\mu$ M ANS, and 10  $\mu$ M to 10 mM phenethyl-ITC. Inhibition tests of esterase activity with CMFDA by ANS or vanillin were performed in a total volume of 200  $\mu$ l with or without 50  $\mu$ M GSH, 50  $\mu$ M ANS, or 300  $\mu$ M vanillin and CMFDA (0  $\mu$ M to 25  $\mu$ M). Similar tests were performed with 25  $\mu$ M CMFDA, 300  $\mu$ M vanillin, and 0–5 mM GSH. The reactions were started by the addition of the purified enzyme, and the catalytic parameters were calculated using the GraphPad<sup>®</sup> software.

**Crystallization**—Two different samples were crystallized, the selenomethionine derivative protein (apoGST5118) and the native protein with bound GSH (holoGST5118). The crystallization experiments were set up using the microbatch under oil (paraffin) method at 277 K. Selenomethionine-apoGST5118 crystals appeared after 4–5 days from droplets containing 2  $\mu$ l of protein solution (15–20 mg/ml protein in TE buffer) and 2  $\mu$ l of precipitating solution (30% PEG 8000, 200 mM sodium acetate, and 100 mM sodium cacodylate, pH 6.5). Crystals of GST5118 in complex with GSH (holoGST5118) were obtained in co-crystallization experiments. The best crystals grew after 2 months in droplets containing a mixture of 485  $\mu$ M protein solution (14 mg/ml) and 6 mM GSH in TE buffer and precipitating solution (30% PEG 8000, 200 mM sodium acetate, and 100 mM HEPES, pH 7.0). Prior to data collection, the crystals were transferred to a cryosolution containing 20% glycerol with mother liquor and flash-cooled to 100 K.

**Data Collection and Processing, Structure Solutions, and Refinements**—ApoGST5118 diffraction data were collected on the synchrotron beamline PX1 at SOLEIL (France) at 0.97918 Å, and holoGST5118 x-ray data were measured on the beam-

**TABLE 1**  
Statistics of x-ray diffraction data collection and model refinement

	ApoGST5118	HoloGST5118
<b>Data collection</b>		
Beam line	PX1, SOLEIL	BM30A/FIP, ESRF
Space group	$P2_12_12_1$	$P2_12_12_1$
No. of dimers in the asymmetric unit	2	2
Cell dimensions $a, b, c$ (Å)	86.05 88.46 157.16	86.57 88.15 156.49
Resolution (Å)	44.23–1.80 (1.89–1.80) <sup>a</sup>	48.48–2.00 (2.11–2.00)
$R_{\text{merge}}$	0.08 (0.619)	0.069 (0.375)
Mean $I/\sigma(I)$	22.2 (3.6)	20.0 (4.7)
Completeness (%)	99.7 (98.1)	100.0 (99.9)
$n$ observations	1,443,573 (112,670)	555,508 (71,179)
Redundancy	12.8 (7.2)	6.8 (6.0)
Wilson $B$ factor (Å <sup>2</sup> )	20.1	21.8
<b>Phasing method</b>		
SAD <sup>b</sup>		
<b>Refinement</b>		
Resolution (Å)	42.54–1.80 (1.84–1.80)	44.90–2.00 (2.02–2.00)
$n$ reflections	111,495 (7,219)	81,509 (2,780)
Cut-off	$F > 0\sigma(F)$	$F > 0\sigma(F)$
$R_{\text{all}}$ (%) <sup>c</sup>	19.3	17.8
$R_{\text{free}}$ (%) <sup>d</sup>	22.7 (30.1)	22.1 (32.5)
Average $B$ -factor (Å <sup>2</sup> )		
Protein atoms	25.1	24.6
Ligand atoms	28.1	30.6
Solvent atoms	29.7	31.2
Ramachandran statistics (%)		
Residues in preferred regions	98.5	98.6
Residues in allowed regions	1.1	1.0
Outlier residues	0.4	0.4
Root mean square deviations		
Bond length (Å)	0.007	0.007
Bond angle (degrees)	1.025	1.034

<sup>a</sup> Values in parentheses are for highest resolution shell.<sup>b</sup> Single-wavelength anomalous dispersion.<sup>c</sup>  $R_{\text{all}}$  was determined from all of the reflections (working set + test set), whereas  $R_{\text{free}}$  corresponds to a subset of reflections (test set).

line BM30A/FIP at the European Synchrotron Radiation Facility (Grenoble, France) at 0.97854 Å. The data sets were processed with XDS (26) and scaled with SCALA (27) from the CCP4 package (28). The molecular replacement method failed to solve the protein structure, probably because the best templates share less than 30% sequence identity with GST5118. The apoGST5118 structure was solved by the single-wavelength anomalous dispersion method using Phenix AutoSol (29). The eight expected selenium atoms were found using SOLVE (30), and resulting single-wavelength anomalous dispersion phases had a figure of merit of 0.38. A nearly complete backbone trace was obtained automatically with RESOLVE (30). ApoGST5118 and holoGST5118 models were refined using PHENIX (29) and manually corrected using COOT (31). The validations of GST5118 crystal structures were automatically performed with MOLPROBITY (32) from PHENIX (29). Final model statistics are shown in Table 1.

ApoGST5118 and holoGST5118 crystals belonged to the space group  $P2_12_12_1$ , and the asymmetric unit consisted of four polypeptide chains. Non-denaturant liquid chromatography and crystal packing analysis using PISA (33) showed that both proteins are homodimeric. The asymmetric unit content corresponds to two dimers where the two protomers are related to each other by a noncrystallographic 2-fold symmetry axis. The electron density maps of holoGST5118 revealed unambiguously the presence of one GSH molecule per monomer. In the apo structure, an electron density peak was found in the active site, identified as a sulfate ion arising most probably from the ammonium fractionation step used during purification. In the holo structure, the corresponding electron density peak was

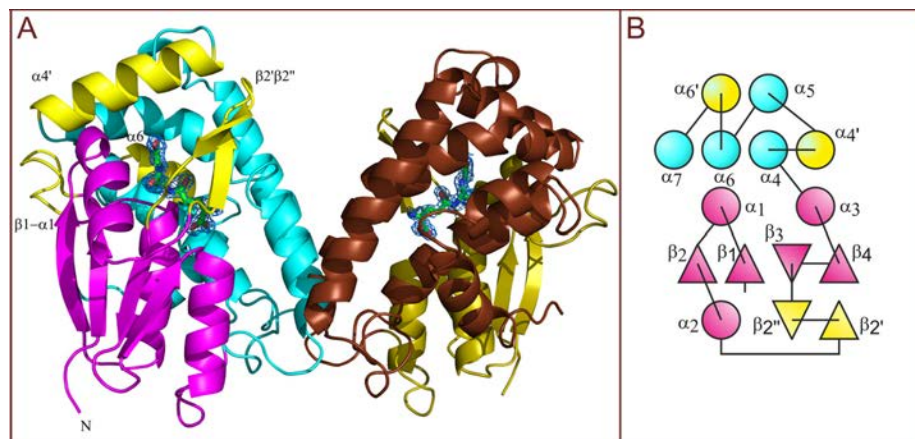
clearly weaker and was interpreted as an acetate ion from the crystallization solution (supplemental Fig. S1). The quality of the final electron density maps is shown in supplemental Fig. S2.

Structural comparisons were done using the DALI structural alignment server (34) and the program Strap (35). Figures of molecular models were prepared with PyMOL. The root mean square deviations between protein structures were calculated with PyMOL or DALI. Secondary structure predictions were carried out with the server Jpred3 (36). The multiple-sequence alignment figure (supplemental Fig. S6) was prepared using Esprict (37).

## RESULTS

**GST5118 X-ray Structures**—Two structures of GST5118 (apo and holo with bound GSH) were solved by using the single-wavelength anomalous diffraction data collected from crystals of the selenomethionylated protein. Each corresponded to the same local packing with two homodimers arranged similarly in the asymmetric units. In each structure, the four monomers adopted quite similar structures. Average root mean square deviations of the six possible superpositions are 0.34 and 0.27 Å for apo and holo structures, respectively. Root mean square deviations increased by at least 0.2 Å when apo and holo monomers were compared. This increase is mainly contributed by the displacement of  $\beta$ -hairpin  $\beta2'\beta2''$  during the apo-holo transition (see below).

The GST5118 monomer adopts the GST canonical fold, which consists of an N-terminal thioredoxin-like motif ( $\beta1\alpha1\beta2\alpha2\beta3\beta4\alpha3$ ) and a C-terminal domain of at least four helices ( $\alpha4\alpha5\alpha6\alpha7$ ). The two most conserved residues in GST



**FIGURE 1. Unique dimeric organization of GST5118.** *A*, view of the original dimer of GST5118. The additional secondary structures, highlighted in yellow in the left monomer, include an elongation of the  $\beta 1$ - $\alpha 1$  loop, a  $\beta$ -hairpin motif in the  $\alpha 2$ - $\beta 3$  loop, and the  $\alpha$ -helices  $\alpha 4'$  and  $\alpha 6'$ . In the left monomer, the N- and C-terminal domains are colored in magenta and cyan, respectively; in the right monomer they are colored in olive and chocolate, respectively. Both monomers have the ability to bind one glutathione molecule, depicted as green sticks, with corresponding  $2F_o - F_c$  electron density (contour level  $1.2 \sigma$ ). *B*, topology diagrams of GST5118 (circles, helices; triangles, strands). The regular secondary structures of GSTs are colored cyan and magenta, whereas original elements are colored yellow.

are present in GST5118: *cis*-Pro-74 in the N-terminal domain, which is assumed to maintain the enzyme in a catalytically competent structure (38), and Asp-205, which is important for the C-terminal domain cohesion through its participation in the N-capping box (39). Nevertheless, the GST5118 structure contains unique features in both N- and C-terminal domains (Fig. 1). The first one is an elongation of about 8 residues of the loop  $\beta 1$ - $\alpha 1$  (11–22) when compared with almost all GSTs. Only GST-like chloride intracellular channel proteins and the recently characterized GST from *Agrobacterium tumefaciens* (40) exhibit this property. In both GST5118 and *A. tumefaciens* GST structures, this long  $\beta 1$ - $\alpha 1$  loop closes one part of the electrophilic substrate pocket (H site). In addition, a  $\beta$ -hairpin motif ( $\beta 2'$  $\beta 2''$ ) between  $\alpha 2$  and  $\beta 3$  is observed for the first time in a GST fold. This motif hinders the formation of the regular GST dimer and partially covers the GSH binding site (G site). The C-terminal domain includes two extra helices:  $\alpha 4'$ , which follows  $\alpha 4$  in an antiparallel way, and  $\alpha 6'$  located between  $\alpha 6$  and  $\alpha 7$ . This second supplementary helix is a characteristic of plant Phi GSTs (41). As for helix  $\alpha 4'$ , to the best of our knowledge, it has been observed only in GRX2 (glutaredoxin 2) (Protein Data Bank entries 3ir4 and 1g7o (42)). This protein is assumed to be the common ancestor of GSTs (43). Furthermore,  $\alpha 4'$  lies in a position similar to the C-terminal helix of the human Alpha and Theta GSTs (supplemental Fig. S3). It has been proposed that a secondary structure in this position covers the H site and reduces its accessibility (44, 45).

GST5118 exhibits a new dimerization mode with an accentuated open V-shape where both subunits are related by a  $C_2$  symmetry. In the canonical GST assembly, the  $\alpha 4$  and  $\alpha 5$  helices of one monomer pack against the N-terminal domain of its partner. In GST5118, the additional  $\beta$ -hairpin prevents this proximity. Helix  $\alpha 4$  of one subunit packs in the groove between helices  $\alpha 4$  and  $\alpha 5$  of the other protomer (Fig. 1 and supplemental Fig. S4). Interestingly, this new oligomeric structure conserves a large cleft in the center of the dimer facing both of the active sites. The size of the buried interface ( $1850 \text{ \AA}^2$ ) is significant and comparable with that observed in the “open” classical

GST dimers like in human Omega GST (46) or in *Drosophila* Sigma GST (47). There are five hydrogen bonds and a hydrophobic lock-and-key motif that hold the two protomers together. The key residue of the lock-and-key motif is Phe-113 from helix  $\alpha 4$ . It fits in the pocket of the other subunit formed by Pro-107 from loop  $\alpha 3$ - $\alpha 4$ ; Gln-114 and Phe-117 from helix  $\alpha 4$ ; Leu-195, Met-196, and Thr-203 from the  $\alpha 5$ - $\alpha 6$  loop; and Ser-205 from helix  $\alpha 6$  (supplemental Fig. S5). Interestingly, this motif is present in Alpha, Mu, and Pi GSTs (48), where the lock is also located between helices  $\alpha 4$  and  $\alpha 5$ . However, in the latter cases, the key is located in loop  $\alpha 2$ - $\beta 3$ , which is replaced by the protruding  $\beta$ -hairpin in GST5118.

In the catalytic center of the holo structure, the GSH molecule fits in the canonical G site, where the anchoring motifs for the  $\gamma$ -glutamyl and cysteinyl moieties are conserved in GSTs from this class. The imidazole group of His-70, from the  $\beta$ -hairpin motif, interacts via a water molecule with the carbonyl group of the  $\gamma$ -glutamyl moiety and with the amide group of the glycyl moiety (Fig. 2). Its carboxylate group is stabilized by the side chains of Tyr-148 and Arg-153 from  $\alpha 4'$ . The thiol sulfur of GSH forms a hydrogen bond with the hydroxyl group of the putative catalytic Ser-22, which is located at the N-terminal end of  $\alpha 1$  in the thioredoxin domain. This Ser-22 occupies the same position as the catalytic residue of the Cys/Ser-GSTs (supplemental Fig. S3). The two residues Trp-21 and Pro-23, which surround the above-mentioned catalytic serine and take part in the H site, are invariant in GSTs from this class (supplemental Fig. S6). Trp-21 is a very well conserved residue in thioredoxins that is involved in enzyme substrate recognition (49), and Pro-23 is a well conserved residue in the active site motif in the class I glutaredoxins (50).

Upon GSH binding, significant conformational changes occur in the GST5118 structure. In the G site, the  $\beta$ -hairpin motif comes closer to the glycyl moiety, allowing its interaction with the side chain of His-70, which adopts a new rotamer (from m-70 to t60). The side chain conformation of Asp-87 also changes to stabilize the N-terminal end of GSH (from t70 to t0). In the H site, Trp-122 and Phe-127 from  $\alpha 4$  undergo confor-

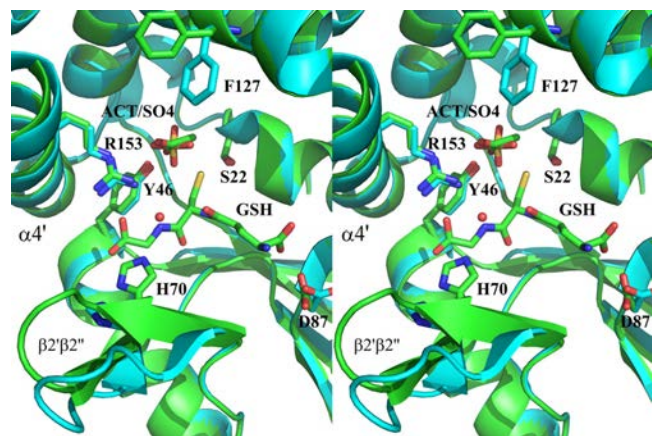


## Ligandin Properties of a New Fungal GST

mational changes that could improve the electrophilic substrate accessibility (Fig. 2).

In the apoGST5118 form, a heavy residual peak of electron density was found in the vicinity of the putative catalytic Ser-22. It was assigned to a sulfate ion, which is stabilized by the side chains of Tyr-46 ( $\alpha 2$ ) and Arg-153 ( $\alpha 4'$ ). In a similar position, a sulfate binding site was reported in the crystal structures of human Theta GST (45), human Zeta GST (51), and yeast elongation factor 1B $\gamma$  (52). In order to stabilize the sulfate anion, the role of Arg-153 ( $\alpha 4'$ ) of GST5118 is played by Arg-239 of human Theta GST located in the C-terminal helix (supplemental Fig. S3). It is important to stress that both residues are located roughly in the same spatial position although not in the same secondary structure. A catalytic role was attributed to the anion binding site in human GST T2-2, which displayed a sulfatase activity toward aralkylsulfates (45).

**GST5118 Activities**—Glutathionylation activity of GST5118 has been tested using CDNB, phenethyl-ITC, 4-nitrophenyl butyrate (PNP-butyrates), and CMFDA as substrates (Table 2). CMFDA is a molecule releasing fluorescence upon activation



**FIGURE 2. Stereo view of the active site of GST5118.** The apo and holo forms are shown *superimposed* in order to highlight the conformational changes upon GSH binding. The anion binding site of GST5118 contains a sulfate ion in the apo form and an acetate ion in the holo form. The carbon atoms of the GSH-liganded and unliganded forms are colored cyan and green, respectively. Other atoms are colored according to their types. For clarity, only selected residues of the active site are shown and labeled.

**TABLE 2**

### Kinetic parameters of GST5118 and GST5118-S22A in enzymatic assays

The apparent  $K_m$  values for all compounds were determined using a concentration range of 0.1–10 mM in the presence of 2 mM GSH for GST5118 and 5 mM for GST5118-S22A. The  $K_m$  value for GSH was determined with 600  $\mu$ M Phenethyl-ITC for GST5118 and 1 mM PNP-butyrates for GST5118-S22A and a concentration range of 0.01–10 mM GSH. The apparent  $K_m$  and  $k_{cat}$  values were calculated by nonlinear regression using the Michaelis-Menten equation ( $r^2 > 0.99$ ). Data are represented as mean  $\pm$  S.D. ( $n \geq 3$ ). ND, not detected. The detection limit was estimated at 0.5 mIU. GOU $\alpha$ O,  $\alpha$ -O-methylumbelliferyl- $\beta$ -hydroxyl-propiovanillone.

Substrate	$K_m$		$k_{cat}$		$k_{cat}/K_m$	
	GST5118	GST5118-S22A	GST5118	GST5118-S22A	GST5118	GST5118-S22A
	$\mu$ M		$\mu$ M		$\text{min}^{-1} \mu\text{M}^{-1}$	
CDNB	ND	653.8 $\pm$ 49.4	ND	1765.0 $\pm$ 115.6	ND	2.69 $\pm$ 0.02
HED	ND	178.6 $\pm$ 7.8	ND	57.3 $\pm$ 6.8	ND	0.32 $\pm$ 0.02
DHA	ND	ND	ND	ND	ND	ND
Phenethyl-ITC	119.0 $\pm$ 14.4	ND	510.9 $\pm$ 16.0	ND	4.3 $\pm$ 0.3	ND
H <sub>2</sub> O <sub>2</sub>	ND	ND	ND	ND	ND	ND
Tertbutyl-OOH	ND	ND	ND	ND	ND	ND
Cu-OOH	ND	3767.0 $\pm$ 437.0	ND	62.6 $\pm$ 1.5	ND	0.011 $\pm$ 0.002
GOU $\alpha$ O	ND	ND	ND	ND	ND	ND
PNP-butyrates	774.3 $\pm$ 88.3	190.7 $\pm$ 24.1	165.0 $\pm$ 5.6	17.6 $\pm$ 0.5	0.21 $\pm$ 0.01	0.092 $\pm$ 0.005
CMFDA without GSH	3.6 $\pm$ 0.6	1.1 $\pm$ 0.1	0.0060 $\pm$ 0.0002	0.27 $\pm$ 0.01	0.0017 $\pm$ 0.0002	0.24 $\pm$ 0.01
CMFDA with GSH	8.0 $\pm$ 0.8	1.5 $\pm$ 0.1	0.17 $\pm$ 0.04	1.20 $\pm$ 0.06	0.021 $\pm$ 0.002	0.80 $\pm$ 0.02
GSH	216.4 $\pm$ 34.9	615.4 $\pm$ 44.2	662.0 $\pm$ 14.2	13.8 $\pm$ 0.6	3.1 $\pm$ 0.4	0.022 $\pm$ 0.002

by esterases (23). GST5118 is active at transferring GSH onto phenethyl-ITC and PNP-butyrates but not CDNB. Surprisingly, GST5118 also exhibits two types of esterase activity on CMFDA, one with and another one without GSH (Table 2). Still, in the absence of GSH, the activity remains weak, with a catalytic efficiency 12-fold lower than in the presence of GSH. By contrast, no thiol transferase activity could be detected with hydroxyethyl disulfide, and the enzyme was also inactive in the reduction of dehydroascorbate. In addition, no peroxidase activity could be detected with either hydrogen peroxide, *tert*-butyl hydroperoxide, or cumene hydroperoxide. Etherase activity was also tested using the synthetic substrate  $\alpha$ -O-methylumbelliferyl- $\beta$ -hydroxylpropiovanillone. This compound is an analog of lignin fragments containing an ether linkage that can be cleaved by Lig proteins (17). Using LigF from *Sphingobium* sp. SYK-6 as control (53), we showed that GST5118 does not possess etherase activity with  $\alpha$ -O-methylumbelliferyl- $\beta$ -hydroxylpropiovanillone. GST5118 does not display sulfatase activity using *para*-nitrophenyl sulfate potassium, although a sulfate binding site has been identified in the structure.

Based on sequence alignments and structural data, we postulated that Ser-22 (GST5118 numbering) could be the catalytic residue responsible for GST5118 activity. Two directed mutations have been performed replacing Ser-22 by an alanine (GST5118-S22A) or a cysteine (GST5118-S22C). Both proteins lost GSH transferase activity using phenethyl-ITC. Surprisingly, GST5118-S22A gained activity with CDNB, hydroxyethyl disulfide, and cumene hydroperoxide (Table 2). On the other hand, GST5118-S22C was inactive with all substrates used in this study (data not shown).

**ANS Binding Site**—ANS is an environment-sensitive fluorescent dye, the fluorescence quantum yield of which increases upon binding to hydrophobic sites of proteins (54). ANS binding onto GST5118 was accompanied by the appearance of a characteristic fluorescence emission spectrum with a maximum at 475 nm, the excitation wavelength being fixed at 385 nm (Fig. 3A). The fluorescence signal, resulting from the interaction between the protein and ANS, was established rapidly in less than 30 s and increased with the molar ratio of ANS to GST5118, revealing a saturation curve. For concentrations

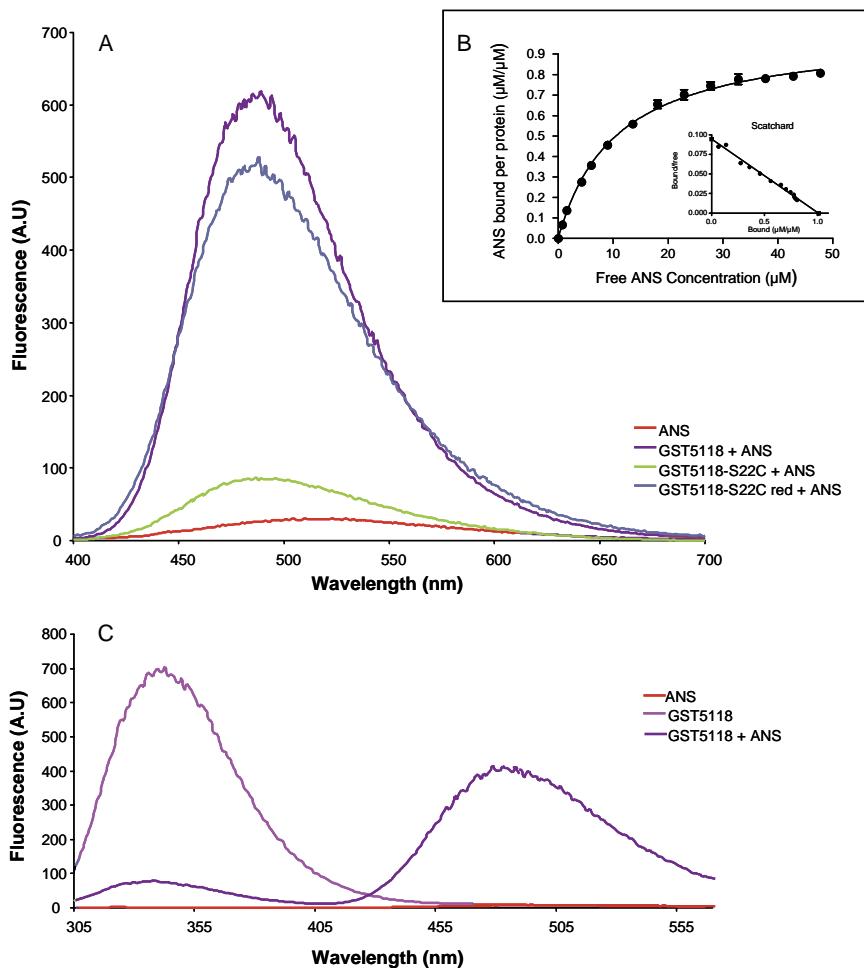


FIGURE 3. **ANS binding onto GST5118.** *A*, comparison of the fluorescence spectra (excitation at 385 nm) of 50  $\mu\text{M}$  ANS bound onto GST5118 and GST5118-S22C DTT-treated (reduced) or not. Emission spectrum of 50  $\mu\text{M}$  ANS alone is reported. *B*, binding saturation of GST5118 with ANS and Scatchard plot. *C*, tryptophan-based fluorescence (excitation at 290 nm, emission at 340 nm) of GST5118 alone and upon the addition of ANS. The emission spectrum of 50  $\mu\text{M}$  ANS alone is reported. Data are represented as mean  $\pm$  S.D. ( $n \pm 3$ ). A.U., absorbance units.

higher than 50  $\mu\text{M}$  ANS, a loss of signal was observed as previously shown for BSA or MurA interaction with ANS (54). A dissociation constant of  $10.65 \pm 0.5 \mu\text{M}$  was obtained by plotting the concentration of ANS bound onto GST5118 against the concentration of free ANS. Data were fitted to Equation 1. The number of ANS binding sites, determined as described under "Experimental Procedures," was  $1.008 \pm 0.015$ , suggesting the binding of only one ANS molecule per monomer (Fig. 3*B*).

In additional experiments, tryptophan-based fluorescence of GST5118 was measured using 290 and 340 nm as excitation and emission wavelengths, respectively, in the presence or absence of ANS (Fig. 3*C*). Fluorescence resonance energy transfer (FRET), revealed by the appearance of a signal at 475 nm, was observed between Trp and ANS, confirming the binding of ANS onto the protein, the energy transfer being highly dependent on the distance between the donor and acceptor molecules (Fig. 3*C*). This result is in accordance with the presence of four Trp residues in the vicinity of the Ser-22 (Trp-21, Trp-26, Trp-122, and Trp-215).

Similar experiments have been performed using the mutated protein GST5118-S22C. Interestingly, the GST5118-S22C

mutant is able to bind ANS only after reduction with DTT (Fig. 3*A*). Besides giving information about the ANS binding site in GST5118, these data demonstrate the robustness of the method. Mass spectrometry experiments performed on GST5118-S22C showed that it covalently binds a GSH adduct. Indeed, purified GST5118-S22C exhibited a molecular mass of 29,033 Da, whereas the previously DTT-treated protein exhibited the expected molecular mass calculated from the primary sequence of 28,728 Da. Additionally, after treatment with GSSG, *S*-(phenylacetophenone)-glutathione, and 2-methyl-*S*-glutathionyl-naphtoquinone, the mass of the reduced protein increased by 305 Da (29,033 Da), which corresponds to a GSH adduct. By contrast, no GSH adduct was detected on the wild type protein, suggesting that the mutated protein is able to bind GSH covalently at the newly added cysteinyl residue, this addition preventing ANS binding onto the protein. ANS binding onto GST5118 was investigated in the presence of various GSH concentrations (up to 10 mM), demonstrating that GSH inhibits ANS binding. The obtained data were fitted to Equation 2, leading to an  $\text{IC}_{50}$  value of  $261.5 \pm 39.8 \mu\text{M}$  (supplemental Fig. S7). This measured value is in accordance with the  $K_m$  for GSH ( $216.4 \pm 34.89 \mu\text{M}$ ) found in the enzymatic assays. Additionally,

## Ligandin Properties of a New Fungal GST

**TABLE 3**

**Competition experiments between ANS and various compounds**

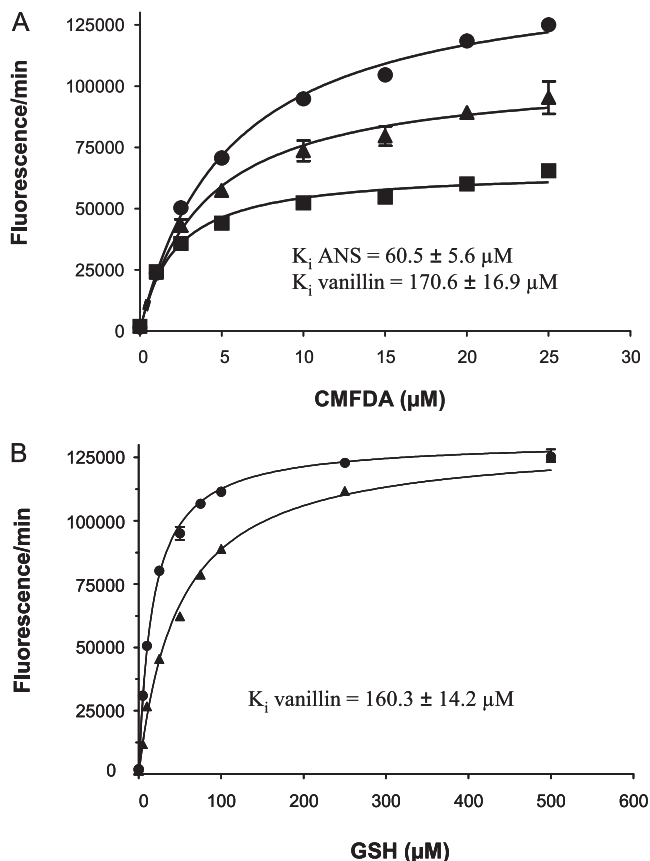
Values given are  $IC_{50}$  fluorescence inhibition obtained by fitting data to Equation 2. Menadione-SG, 2-methyl-S-glutathionyl-naphtoquinone.

GST5118	
$K_d$ ANS	$10.6 \pm 0.5 \mu\text{M}$
No. of fixation sites at $50 \mu\text{M}$ ANS	$1.008 \pm 0.015$
Glutathione	$261.5 \pm 39.8 \mu\text{M}$
<b>Beech compounds</b>	
Coniferaldehyde	$61.3 \pm 1.2 \mu\text{M}$
Vanillin	$216.6 \pm 8.5 \mu\text{M}$
2-Hydroxy-4-methoxybenzophenone	–
4-Chloro-3-nitrobenzoic acid	$1.5 \pm 0.1 \text{mM}$
4'-Hydroxyacetophenone	$2.3 \pm 0.4 \text{mM}$
Gallic acid	–
Vanillic acid	–
Epicatechin	$1.7 \pm 0.1 \text{mM}$
Syringaldehyde	$79.4 \pm 1.6 \mu\text{M}$
Catechin hydrate	$584.3 \pm 98.7 \mu\text{M}$
3-Hydroxy-4-methoxycinnamic acid, predominantly <i>trans</i>	–
<b>Glutathionylated compounds</b>	
2-Bromo-4'-phenylacetophenone	–
Phenylacetophenone-SG	$99.7 \pm 46.5 \mu\text{M}$
Menadione	–
Menadione-SG	$83.7 \pm 38.4 \mu\text{M}$
<b>Sulfated compounds</b>	
1,2-Naphtoquinone-4-sulfonic acid	$334.5 \pm 56.8 \mu\text{M}$
<i>p</i> -Xylene-2-sulfonic acid hydrate	$928.7 \pm 258.8 \mu\text{M}$
<i>p</i> -Xylene	–
4-Aminotoluene-3-sulfonic acid	$1.6 \pm 0.5 \text{mM}$
4-Methylumbelliferyl-sulfate	$223.6 \pm 25.3 \mu\text{M}$
4-Methylumbelliferone	–

excitation transfer between Trp and bound ANS was fully abolished in the presence of 1 mM GSH (supplemental Fig. S8). *S*-(phenylacetophenone)-Glutathione and 2-methyl-*S*-glutathionyl-naphtoquinone were also able to inhibit ANS binding, whereas the same compounds deprived of the GSH adduct could not (Table 3). Taking together, these data suggest strongly that in GST5118, the ANS binding site overlaps at least in part the glutathione binding site (G site).

In the vicinity of catalytic Ser-22, a sulfate/acetate binding site has been detected in the GST5118 structure. Because ANS possesses a sulfonic acid motif, various sulfated compounds were tested for their ability to interact with the ANS binding site in GST5118. ANS binding is inhibited in the presence of 4-aminotoluene-3-sulfonic acid, *p*-xylene-2-sulfonic acid hydrate, 1,2-naphtoquinone-4-sulfonic acid sodium, and also 4-methylumbelliferyl sulfate potassium (Table 3). Moreover, this inhibition seems to be sulfate/sulfonate-dependent because 4-methylumbelliferone and *p*-xylene do not compete with ANS, suggesting that the ANS binding site also overlaps with the sulfate-binding site in GST5118.

Competition experiments between ANS and various substrates have been performed in order to investigate possible overlapping between the substrate binding site and the ANS binding site. As shown above, GST5118 exhibits a relatively high affinity for CMFDA, phenethyl-ITC, and PNP-butyrate with and without GSH (Table 3). ANS binding onto GST5118 remained mainly unaltered in the presence of either phenethyl-ITC or CMFDA (supplemental Fig. S9). Moreover, GST5118 activity against phenethyl-ITC is inhibited in the presence of ANS (supplemental Fig. S10). Concerning CMFDA, only esterase activity with GSH is inhibited by ANS (Fig. 4A). Due to quenching phenomena, PNP-butyrate was not used in this test.



**FIGURE 4. Inhibition of GSH-dependent esterase activity by ANS or vanillin.** Data are represented as mean  $\pm$  S.D. ( $n = 3$ ). *A*, noncompetitive inhibition of esterase activity against CMFDA without inhibitor ( $\bullet$ ) or in the presence of  $50 \mu\text{M}$  ANS ( $\blacksquare$ ) or  $300 \mu\text{M}$  vanillin ( $\blacktriangle$ ).  $K_i$  values were calculated resolving to the non-linear noncompetitive inhibition equation ( $r^2 > 0.99$ ):  $V_{\text{maxinh}} = V_{\text{max}} / (1 + I/K_i)$ ;  $Y = V_{\text{maxinh}} \times X / (K_m + X)$ .  $V_{\text{max}}$  and  $K_m$  represent the maximum enzyme velocity and the Michaelis-Menten constant without inhibitor, whereas  $V_{\text{maxinh}}$  represents the maximum enzyme velocity for one concentration of inhibitor, and  $K_i$  is the inhibition constant. *B*, competitive inhibition of esterase activity GSH-dependent against GSH without inhibitor ( $\bullet$ ) or in the presence of  $300 \mu\text{M}$  vanillin ( $\blacktriangle$ ). The  $K_i$  value was calculated resolving to the non-linear competitive inhibition equation ( $r^2 > 0.99$ ):  $K_{m(\text{obs})} = K_m \times (1 + [I]/K_i)$ ;  $Y = V_{\text{max}} \times X / (K_{m(\text{obs})} + X)$ .  $V_{\text{max}}$  and  $K_m$  represent the maximum enzyme velocity and the Michaelis-Menten constant without inhibitor, whereas  $K_{m(\text{obs})}$  represents the Michaelis-Menten constant in the presence of inhibitor, and  $K_i$  is the inhibition constant.

Taken together, these data suggest that the ANS binding site does not overlap with the substrate binding site because the probe does not compete with hydrophobic substrates as phenethyl-ITC or CMFDA for binding. By contrast, the competition between ANS and GSH (Fig. 4B) and the inhibition of GSH-requiring activities show an overlap between ANS and the G site of GST5118.

**Ligand Screening**—Using ANS as a screening tool, competition experiments have been performed with 11 structurally variable beech compounds (supplemental Fig. S11). Seven of them were recognized by GST5118 (Table 3). Among these products, those exhibiting the better affinity for GST5118 are either aldehydes or compounds harboring electrophilic motifs or atoms such as chlorine or nitrogen. Steric hindrance is not a factor influencing ligand binding because epicatechin and catechin hydrate, two of the largest compounds used, are both bound by GST5118.

To confirm the interaction between these compounds and GST5118, additional experiments were conducted using vanillin and testing the ability of GST5118 to react with CFMDA in the presence and absence of glutathione. Vanillin strongly inhibited the activity in the presence of GSH (Fig. 4) but had no effect in the absence of GSH (data not shown). Kinetic experiments revealed, as expected, a non-competitive inhibition against CFMDA (*i.e.* different binding sites for vanillin and CFMDA) and a competitive inhibition against GSH (*i.e.* overlapping binding sites for vanillin and GSH). In both cases, the  $K_i$  value ( $K_i$  around  $200 \mu\text{M}$ ) is in accordance with the  $\text{IC}_{50}$  value calculated from the experiments performed with ANS ( $216.6 \pm 8.5 \mu\text{M}$ ).

As summarized in Fig. 5, all of these data suggest that GSH, ANS, sulfate/acetate, and ligands such as vanillin bind to GST5118 through overlapping binding sites, which are different from the substrate-binding site.

## DISCUSSION

GST5118 is a newly identified fungal GST with unique properties both in terms of structural organization and biochemistry. GST5118 monomer exhibits two extra secondary structures in addition to the canonical GST framework: 1) the  $\beta$ -hairpin  $\beta 2' \beta 2''$  in the N-terminal domain, which hinders the formation of the regular GST dimer and which may be described as a shield covering the G site once GSH is bound, and 2) the helix  $\alpha 4'$ , which closes the presumed electrophilic substrate binding site and involves a basic side chain in the anion binding site of the enzyme. Based on structural alignments, all five *P. chrysosporium* GSTs from this class should share the same original properties, allowing us to define a new structural GST class with a novel dimerization mode. This class holds the unique property of having conserved all the secondary structures of monomeric glutaredoxin 2, which is the presumed common ancestor of the cytosolic GSTs (43). All other structures solved so far lost at least helix  $\alpha 4'$ . The catalytic Ser-22 is in hydrogen-bonding contact with the thiol sulfur atom of the cysteinyl moiety of GSH and thus could participate in the stabilization of the thiolate anion. The GST5118-S22A mutant retains the ability to conjugate GSH onto electrophilic substrates and exhibits weak peroxidase and thiol transferase activities. This suggests that another residue could stabilize the thiolate form of GSH and thus lower its  $\text{p}K_a$ . In the vicinity of GSH, Asn-24 which is conserved in isoforms of this class (supplemental Fig. S6) could fulfill this role with few conformational changes. In accordance, this Asn residue is conserved in Ure2p class and has been shown to be crucial for peroxidase activity of *Saccharomyces cerevisiae* Ure2p (55). This observation could explain why GST5118-S22A shows such an activity pattern.

A sulfate binding pocket was identified in the N-terminal domain of GST5118 structure, close to the GSH binding site. This could explain the binding of ANS, which possesses a sulfonic acid moiety, at the G site and the competition between ANS and GSH. The unique binding mode of ANS onto GST5118 allowed us to use it as a ligand screening tool. Glutathionylated and sulfur-containing compounds as well as products resulting from wood degradation compete with ANS, suggesting that all of them bind at the G site. This is in accordance

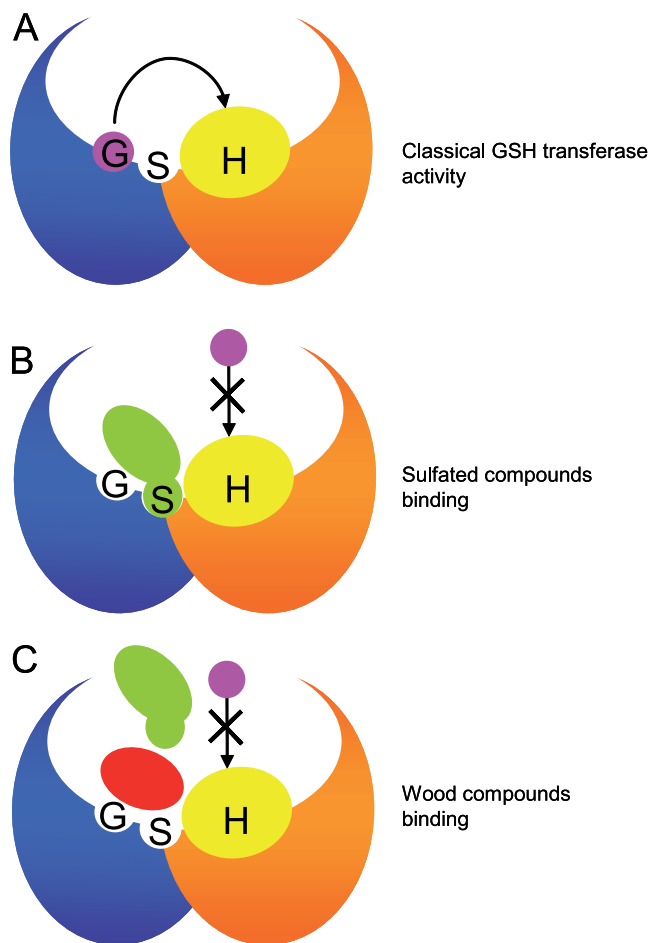


FIGURE 5. Schematic representation of compounds binding onto GST5118 monomer. The N-terminal domain containing the GSH binding site (G) as well as the ANS and sulfated compound binding site (S) is depicted in blue (on the left), whereas the C-terminal domain containing the substrate binding site (H) is depicted in orange (on the right). Binding of both GSH (pink) and substrate (yellow) in their respective sites (A) enables the classical GSH transferase activity of the enzyme. Binding of ANS and sulfated compounds in the S site (B) or binding of non-substrate wood compounds (red) in the N-terminal domain (C) inhibits the GSH binding but does not affect substrate binding and alters the GSH transferase activity of the enzyme.

with the structural data concerning glutathionylated and sulfated compounds; however, it was unexpected for wood compounds because the recognition of hydrophobic substrates has never been described to occur at the G site.

The consequence of binding wood compounds at the G site is the inability of GST5118 to accept and therefore transfer GSH anymore. This enzyme could thus have two functions, possibly depending on the intracellular concentration of GSH and wood compounds: a classical GSH transferase activity, especially with phenethyl-ITC as substrate, and a ligandin property toward wood extractive compounds. The affinity for ligand binding is likely to be driven by very fine interactions. An example of this fine ligand specificity is the pair vanillin/vanillic acid, in which the first compound is well recognized by GST5118, whereas the latter cannot displace ANS. The only difference between these two molecules is the addition of a single atom of oxygen, transforming the aldehyde function into an acidic one. Vanillin is

## Ligandin Properties of a New Fungal GST

known to be one of the key intermediates found during wood decay, and it was shown to be taken up by fungal cells (56). However, vanillin is not a nutrient substrate but rather acts as a chemical stress on fungal cells (57). It has also been suggested that it can be responsible for activation of fungal metabolic pathways for phenolic compounds (58). This ligandin property of GST5118 could thus be helpful to protect the cell against highly reactive xenobiotics or secondary metabolites. Similar functions were assigned to plant Phi and Tau GSTs that participate in the transport of highly reactive products to the vacuole without catalyzing their conjugation and thus sequestering them away from critical intracellular targets (59, 60).

In conclusion, the structural and functional properties of GST5118 allow us to define a new GST class that we name GSTFuA (for fungal specific class A). Accordingly, GST5118 from *P. chrysosporium* can thus be renamed GSTFuA1.

*Acknowledgments*—We thank the staffs of the BM30A beamline at the European Synchrotron Radiation Facility (Grenoble, France) and the Proxima 1 beamline at SOLEIL (Gif-sur-Yvette, France) for kind assistance during data collection.

### REFERENCES

- Hayes, J. D., Flanagan, J. U., and Jowsey, I. R. (2005) Glutathione transferases. *Annu. Rev. Pharmacol. Toxicol.* **45**, 51–88
- Mannervik, B. (2012) Five decades with glutathione and the GSTome. *J. Biol. Chem.* **287**, 6072–6083
- Oakley, A. J., Lo Bello, M., Nuccetelli, M., Mazzetti, A. P., and Parker, M. W. (1999) The ligandin (non-substrate) binding site of human Pi class glutathione transferase is located in the electrophile binding site (H-site). *J. Mol. Biol.* **291**, 913–926
- Lederer, B., and Böger, P. (2003) Binding and protection of porphyrins by glutathione S-transferases of *Zea mays* L. *Biochim. Biophys. Acta* **1621**, 226–233
- Dixon, D. P., Laphorn, A., Madesis, P., Mudd, E. A., Day, A., and Edwards, R. (2008) Binding and glutathione conjugation of porphyrinogens by plant glutathione transferases. *J. Biol. Chem.* **283**, 20268–20276
- Bilang, J., Macdonald, H., King, P. J., and Sturm, A. (1993) A soluble auxin-binding protein from *Hyoscyamus muticus* is a glutathione S-transferase. *Plant Physiol.* **102**, 29–34
- Conn, S., Curtin, C., Bézier, A., Franco, C., and Zhang, W. (2008) Purification, molecular cloning, and characterization of glutathione S-transferases (GSTs) from pigmented *Vitis vinifera* L. cell suspension cultures as putative anthocyanin transport proteins. *J. Exp. Bot.* **59**, 3621–3634
- Morel, M., Ngadin, A. A., Droux, M., Jacquot, J. P., and Gelhaye, E. (2009) The fungal glutathione S-transferase system. Evidence of new classes in the wood-degrading basidiomycete *Phanerochaete chrysosporium*. *Cell Mol. Life Sci.* **66**, 3711–3725
- Thuillier, A., Ngadin, A. A., Thion, C., Billard, P., Jacquot, J. P., Gelhaye, E., and Morel, M. (2011) Functional diversification of fungal glutathione transferases from the ure2p class. *Int. J. Evol. Biol.* **2011**, 938308
- McGoldrick, S., O'Sullivan, S. M., and Sheehan, D. (2005) Glutathione transferase-like proteins encoded in genomes of yeasts and fungi. Insights into evolution of a multifunctional protein superfamily. *FEMS Microbiol. Lett.* **242**, 1–12
- Lian, H. Y., Jiang, Y., Zhang, H., Jones, G. W., and Perrett, S. (2006) The yeast prion protein Ure2. Structure, function, and folding. *Biochim. Biophys. Acta* **1764**, 535–545
- Garcerá, A., Barreto, L., Piedrafitra, L., Tamarit, J., and Herrero, E. (2006) *Saccharomyces cerevisiae* cells have three Omega class glutathione S-transferases acting as 1-Cys thiol transferases. *Biochem. J.* **398**, 187–196
- Choi, J. H., Lou, W., and Vancura, A. (1998) A novel membrane-bound glutathione S-transferase functions in the stationary phase of the yeast *Saccharomyces cerevisiae*. *J. Biol. Chem.* **273**, 29915–29922
- Fraser, J. A., Davis, M. A., and Hynes, M. J. (2002) A gene from *Aspergillus nidulans* with similarity to URE2 of *Saccharomyces cerevisiae* encodes a glutathione S-transferase which contributes to heavy metal and xenobiotic resistance. *Appl. Environ. Microbiol.* **68**, 2802–2808
- Burns, C., Geraghty, R., Neville, C., Murphy, A., Kavanagh, K., and Doyle, S. (2005) Identification, cloning, and functional expression of three glutathione transferase genes from *Aspergillus fumigatus*. *Fungal Genet. Biol.* **42**, 319–327
- Meux, E., Prosper, P., Ngadin, A., Didierjean, C., Morel, M., Dumarçay, S., Lamant, T., Jacquot, J. P., Favier, F., and Gelhaye, E. (2011) Glutathione transferases of *Phanerochaete chrysosporium*. S-Glutathionyl-p-hydroquinone reductase belongs to a new structural class. *J. Biol. Chem.* **286**, 9162–9173
- Masai, E., Ichimura, A., Sato, Y., Miyauchi, K., Katayama, Y., and Fukuda, M. (2003) Roles of the enantioselective glutathione S-transferases in cleavage of  $\beta$ -aryl ether. *J. Bacteriol.* **185**, 1768–1775
- Lee, S., Monnappa, A. K., and Mitchell, R. J. (2012) Biological activities of lignin hydrolysate-related compounds. *BMB Rep.* **45**, 265–274
- Lee, C. Y., Sharma, A., Cheong, J. E., and Nelson, J. L. (2009) Synthesis and antioxidant properties of dendritic polyphenols. *Bioorg. Med. Chem. Lett.* **19**, 6326–6330
- Schenk, P. M., Baumann, S., Mattes, R., and Steinbiss, H. H. (1995) Improved high level expression system for eukaryotic genes in *Escherichia coli* using T7 RNA polymerase and rare Arg tRNAs. *BioTechniques* **19**, 196–198, 200
- D'Ambrosio, K., Kauffmann, B., Rouhier, N., Benedetti, E., Jacquot, J. P., Aubry, A., and Corbier, C. (2003) Crystallization and preliminary x-ray studies of the glutaredoxin from poplar in complex with glutathione. *Acta Crystallogr. D Biol. Crystallogr.* **59**, 1043–1045
- Couturier, J., Koh, C. S., Zaffagnini, M., Winger, A. M., Gualberto, J. M., Corbier, C., Decottignies, P., Jacquot, J. P., Lemaire, S. D., Didierjean, C., and Rouhier, N. (2009) Structure-function relationship of the chloroplastic glutaredoxin S12 with an atypical WCSYS active site. *J. Biol. Chem.* **284**, 9299–9310
- Zhang, J., Shibata, A., Ito, M., Shuto, S., Ito, Y., Mannervik, B., Abe, H., and Morgenstern, R. (2011) Synthesis and characterization of a series of highly fluorogenic substrates for glutathione transferases, a general strategy. *J. Am. Chem. Soc.* **133**, 14109–14119
- Kim, D. E., Kim, K. H., Bae, Y. J., Lee, J. H., Jang, Y. H., and Nam, S. W. (2005) Purification and characterization of the recombinant arylsulfatase cloned from *Pseudoalteromonas carrageenovora*. *Protein Expr. Purif.* **39**, 107–115
- Koh, C. S., Navrot, N., Didierjean, C., Rouhier, N., Hirasawa, M., Knaff, D. B., Wingsle, G., Samian, R., Jacquot, J. P., Corbier, C., and Gelhaye, E. (2008) An atypical catalytic mechanism involving three cysteines of thioredoxin. *J. Biol. Chem.* **283**, 23062–23072
- Kabsch, W. (2010) XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132
- Evans, P. (2006) Scaling and assessment of data quality. *Acta Crystallogr. D Biol. Crystallogr.* **62**, 72–82
- Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Pottert, E. A., Powell, H. R., Read, R. J., Vagin, A., and Wilson, K. S. (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 235–242
- Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221
- Terwilliger, T. (2004) SOLVE and RESOLVE. Automated structure solution, density modification, and model building. *J. Synchrotron Radiat.* **11**, 49–52
- Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501
- Chen, V. B., Arendall, W. B., 3rd, Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., and Richardson, D. C.

- (2010) MolProbity. All-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 12–21
33. Krissinel, E., and Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774–797
  34. Holm, L., and Rosenström, P. (2010) Dali server. Conservation mapping in 3D. *Nucleic Acids Res.* **38**, W545–W549
  35. Gille, C. (2006) Structural interpretation of mutations and SNPs using STRAP-NT. *Protein Sci.* **15**, 208–210
  36. Cole, C., Barber, J. D., and Barton, G. J. (2008) The Jpred 3 secondary structure prediction server. *Nucleic Acids Res.* **36**, W197–W201
  37. Gouet, P., Robert, X., and Courcelle, E. (2003) ESPript/ENDscript. Extracting and rendering sequence and 3D information from atomic structures of proteins. *Nucleic Acids Res.* **31**, 3320–3323
  38. Allocati, N., Casalone, E., Masulli, M., Ceccarelli, I., Carletti, E., Parker, M. W., and Di Ilio, C. (1999) Functional analysis of the evolutionarily conserved proline 53 residue in *Proteus mirabilis* glutathione transferase B1-1. *FEBS Lett.* **445**, 347–350
  39. Cocco, R., Stenberg, G., Dragani, B., Rossi Principe, D., Paludi, D., Manervik, B., and Aceto, A. (2001) 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.* **276**, 32177–32183
  40. Skopelitou, K., Dhavala, P., Papageorgiou, A. C., and Labrou, N. E. (2012) A glutathione transferase from *Agrobacterium tumefaciens* reveals a novel class of bacterial GST superfamily. *PLoS ONE* **7**, e34263
  41. Reinemer, P., Prade, L., Hof, P., Neufelind, T., Huber, R., Zettl, R., Palme, K., Schell, J., Koelln, I., Bartunik, H. D., and Bieseler, B. (1996) Three-dimensional structure of glutathione *S*-transferase from *Arabidopsis thaliana* at 2.2 Å resolution. Structural characterization of herbicide-conjugating plant glutathione *S*-transferases and a novel active site architecture. *J. Mol. Biol.* **255**, 289–309
  42. Xia, B., Vlamis-Gardikas, A., Holmgren, A., Wright, P. E., and Dyson, H. J. (2001) Solution structure of *Escherichia coli* glutaredoxin-2 shows similarity to mammalian glutathione *S*-transferases. *J. Mol. Biol.* **310**, 907–918
  43. Frova, C. (2006) Glutathione transferases in the genomics era. New insights and perspectives. *Biomol. Eng.* **23**, 149–169
  44. Sinning, I., Kleywegt, G. J., Cowan, S. W., Reinemer, P., Dirr, H. W., Huber, R., Gilliland, G. L., Armstrong, R. N., Ji, X., and Board, P. G. (1993) Structure determination and refinement of human Alpha class glutathione transferase A1-1 and a comparison with the Mu and Pi class enzymes. *J. Mol. Biol.* **232**, 192–212
  45. Rossjohn, J., McKinstry, W. J., Oakley, A. J., Verger, D., Flanagan, J., Chelvanayagam, G., Tan, K. L., Board, P. G., and Parker, M. W. (1998) Human Theta class glutathione transferase. The crystal structure reveals a sulfate-binding pocket within a buried active site. *Structure* **6**, 309–322
  46. Board, P. G., Coggan, M., Chelvanayagam, G., Easteal, S., Jermiin, L. S., Schulte, G. K., Danley, D. E., Hoth, L. R., Griffor, M. C., Kamath, A. V., Rosner, M. H., Chrunk, B. A., Perregaux, D. E., Gabel, C. A., Geoghegan, K. F., and Pandit, J. (2000) Identification, characterization, and crystal structure of the Omega class glutathione transferases. *J. Biol. Chem.* **275**, 24798–24806
  47. Agianian, B., Tucker, P. A., Schouten, A., Leonard, K., Bullard, B., and Gros, P. (2003) Structure of a *Drosophila* Sigma class glutathione *S*-transferase reveals a novel active site topography suited for lipid peroxidation products. *J. Mol. Biol.* **326**, 151–165
  48. Dirr, H., Reinemer, P., and Huber, R. (1994) X-ray crystal structures of cytosolic glutathione *S*-transferases. Implications for protein architecture, substrate recognition, and catalytic function. *Eur. J. Biochem.* **220**, 645–661
  49. Menchise, V., Corbier, C., Didierjean, C., Jacquot, J. P., Benedetti, E., Saviano, M., and Aubry, A. (2000) Crystal structure of the W35A mutant thioredoxin h from *Chlamydomonas reinhardtii*. The substitution of the conserved active site Trp leads to modifications in the environment of the two catalytic cysteines. *Biopolymers* **56**, 1–7
  50. Lillig, C. H., Berndt, C., and Holmgren, A. (2008) Glutaredoxin systems. *Biochim. Biophys. Acta* **1780**, 1304–1317
  51. Polekhina, G., Board, P. G., Blackburn, A. C., and Parker, M. W. (2001) Crystal structure of maleylacetoacetate isomerase/glutathione transferase Zeta reveals the molecular basis for its remarkable catalytic promiscuity. *Biochemistry* **40**, 1567–1576
  52. Jeppesen, M. G., Ortiz, P., Shepard, W., Kinzy, T. G., Nyborg, J., and Andersen, G. R. (2003) The crystal structure of the glutathione *S*-transferase-like domain of elongation factor 1B $\gamma$  from *Saccharomyces cerevisiae*. *J. Biol. Chem.* **278**, 47190–47198
  53. Masai, E., Katayama, Y., Kubota, S., Kawai, S., Yamasaki, M., and Morohoshi, N. (1993) A bacterial enzyme degrading the model lignin compound  $\beta$ -etherase is a member of the glutathione *S*-transferase superfamily. *FEBS Lett.* **323**, 135–140
  54. Schonbrunn, E., Eschenburg, S., Luger, K., Kabsch, W., and Amrhein, N. (2000) Structural basis for the interaction of the fluorescence probe 8-anilino-1-naphthalene sulfonate (ANS) with the antibiotic target MurA. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 6345–6349
  55. Zhang, Z. R., Bai, M., Wang, X. Y., Zhou, J. M., and Perrett, S. (2008) “Restoration” of glutathione transferase activity by single-site mutation of the yeast prion protein Ure2. *J. Mol. Biol.* **384**, 641–651
  56. Shimizu, M., Kobayashi, Y., Tanaka, H., and Wariishi, H. (2005) Transport mechanism for vanillin uptake through fungal plasma membrane. *Appl. Microbiol. Biotechnol.* **68**, 673–679
  57. Shimizu, M., Yuda, N., Nakamura, T., Tanaka, H., and Wariishi, H. (2005) Metabolic regulation at the tricarboxylic acid and glyoxylate cycles of the lignin-degrading basidiomycete *Phanerochaete chrysosporium* against exogenous addition of vanillin. *Proteomics* **5**, 3919–3931
  58. Nakamura, T., Ichinose, H., and Wariishi, H. (2012) Flavin-containing monooxygenases from *Phanerochaete chrysosporium* responsible for fungal metabolism of phenolic compounds. *Biodegradation* **23**, 343–350
  59. Marrs, K. A., Alfenito, M. R., Lloyd, A. M., and Walbot, V. (1995) A glutathione *S*-transferase involved in vacuolar transfer encoded by the maize gene *Bronze-2*. *Nature* **375**, 397–400
  60. Kitamura, S., Shikazono, N., and Tanaka, A. (2004) TRANSPARENT TESTA 19 is involved in the accumulation of both anthocyanins and proanthocyanidins in *Arabidopsis*. *Plant J.* **37**, 104–114

## SUPPLEMENTAL DATA

### **Supplemental Figure S1. Stereo views of difference maps ( $3\sigma$ contour level, mFo-DFc maps) in the region of the catalytic active site of apoGST5118 (top) and holoGST5118 (bottom).**

The positive peak was assigned to a sulfate ion in the apo structure (top) and to an acetate ion in the holo structure (bottom). The coordinates of the structures are shown as sticks

### **Supplemental Figure S2. Stereo views of final SIGMAA-weighted $2mFo-DFc$ maps ( $1.2\sigma$ contour level) in the region of the catalytic active site of apoGST5118 (top) and holoGST5118 (bottom).**

The coordinates of the structures are shown as sticks

### **Supplemental Figure S3. Stereo views of the superpositions of GST5118 onto glutaredoxin 2 from *Salmonella typhimurium* (top) and theta class GST from human (bottom).**

Top : GST5118 and StGrx2 (pdb entry 3IR4) are represented in cartoon ribbon style and colored in magenta and yellow, respectively. The stereo view highlights the good superimposition between the GST5118 and GRX2 structures and in particular the presence of the helix  $\alpha 4'$  in both structures. The number of atoms used in the superposition was 198 CA atoms and the resulting rmsd was 2.9 Å. In both structures, the glutathione and the catalytic residue are shown as sticks.

Bottom : GST5118 and hGST T2-2 (pdb entry 3LJR) are represented in cartoon ribbon style and colored in magenta and yellow, respectively. The stereo view highlights the superimposition of the helix  $\alpha 4'$  of GST5118 and the C-terminal helix of hGST T2-2. Both helices involve an arginine (shown as sticks) to stabilize an anion in the active site. The number of atoms used in the superposition was 192 CA atoms and the resulting rmsd was 3.2 Å. In both structures, a sulfate anion and the catalytic serine are shown as sticks.

### **Supplemental Figure S4. Views of the GST5118 dimer and of the canonical GST dimer.**

GST5118 (top) and Phi GST from *A. thaliana* (bottom, pdb entry 1GNW) are represented in cartoon ribbon style and their N-ter and C-ter domains are colored magenta and green, respectively. In the regular GST dimer such as in AtGSTF, the helices  $\alpha 4$  and  $\alpha 5$  of one monomer pack against the N-terminal domain of the other monomer. In GST5118 the  $\beta 2'\beta 2''$  motif hinders this proximity, the helix  $\alpha 4$  of one subunit packs in the groove between helices  $\alpha 4$  and  $\alpha 5$  of the other protomer.

### **Supplemental Figure S5. Stereo view of the lock and key motif in the GST5118 dimer.**

The monomers A and B are represented in cartoon ribbon style and colored cyan and green, respectively. The key (Phe-113), and the residues forming the lock are shown as sticks and are labeled. The transparency surface highlights the pocket hosting the key.

### **Supplemental Figure S6. Sequence alignment based on the structural superimposition of GST5118 and representatives of different GST (or GST-like) classes found by DALI to be the closest structural neighbors of GST5118.**

AtGSTF : a Phi class GST from *Arabidopsis thaliana* (PDB ID 1GNW ; UniProKB P46422) ; wGSTT : a tau class GST from bread wheat (*Aegilops tauschii*) (PDB ID 1GWC ; UniProKB O04941) ; StGRX2 : GRX2 from *Salmonella typhimurium* (PDB ID 3IR4 ; UniProKB Q7CQR3).

The sequence alignment was generated with the programs STRAP and ESPript. GST5118 and StGRX2 secondary structure elements are shown above and below the sequence alignment, respectively.

#### **Supplemental Figure S7. Competition experiment between ANS and GSH binding.**

ANS fluorescence signal decreases upon increasing GSH concentration. Experiment was performed with 50  $\mu\text{M}$  ANS in triplicate and results were least square fitted to the equation 2 described in material and methods to obtain IC50 value.

#### **Supplemental Figure S8. Tryptophan based fluorescence experiments illustrating ANS displacement upon GSH addition.**

Upon tryptophan excitation at 290 nm, FRET between 50  $\mu\text{M}$  ANS and 3  $\mu\text{M}$  GST5118 (purple), is inhibited with addition of 1 mM GSH (orange). Emission spectra of 50  $\mu\text{M}$  ANS with 1 mM GSH is in yellow while GST5118-S22C is in brown.

#### **Supplemental Figure S9. Tryptophan based fluorescence experiments illustrating ANS binding upon addition of substrates.**

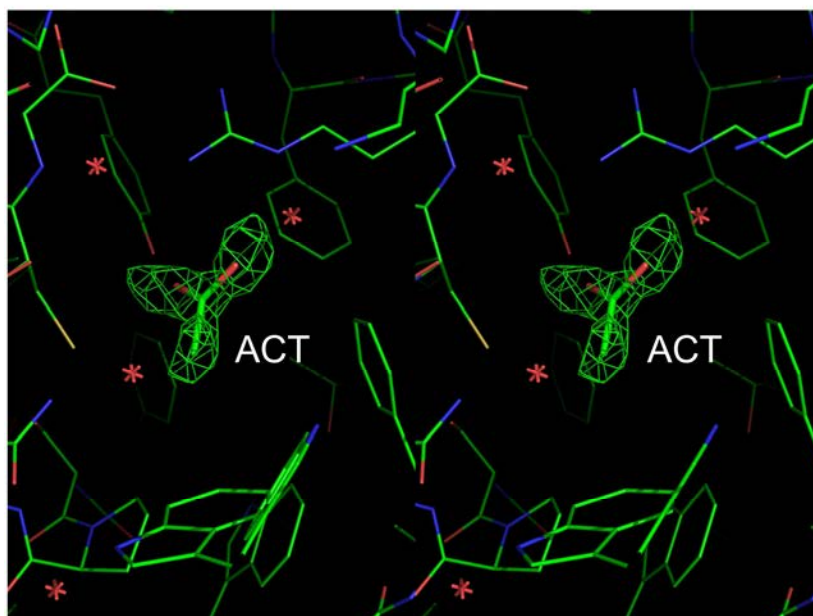
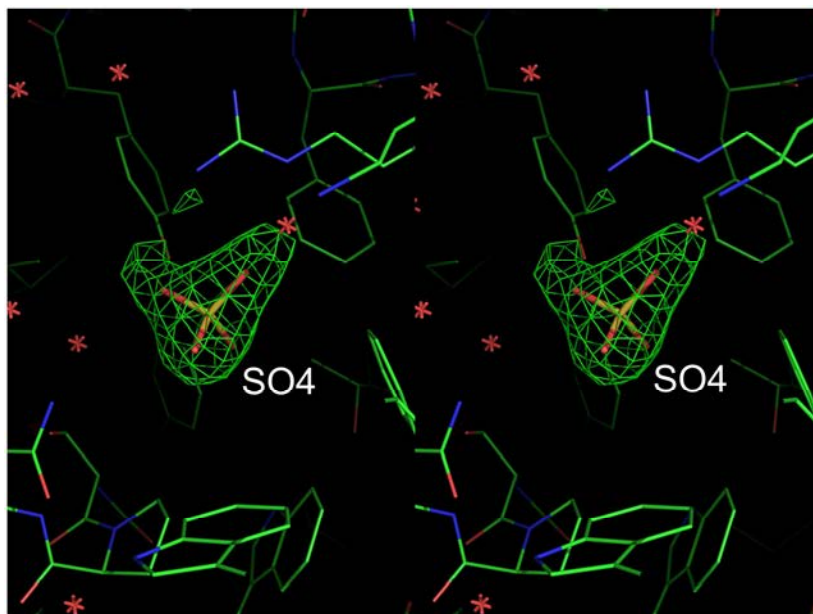
Upon tryptophan excitation at 290 nm, FRET between 50  $\mu\text{M}$  ANS and 3  $\mu\text{M}$  GST5118 (purple) is not altered in presence of 5 mM phenetyl-ITC (black) or 50  $\mu\text{M}$  CMFDA (cyan). Emission spectra of 5 mM phenetyl-ITC or 50  $\mu\text{M}$  CMFDA with 50  $\mu\text{M}$  ANS are in grey and dark green respectively. Emission spectrum of GST5118 alone is in pink.

#### **Supplemental Figure S10. Inhibition of GSH transferase activity against ITC by ANS.**

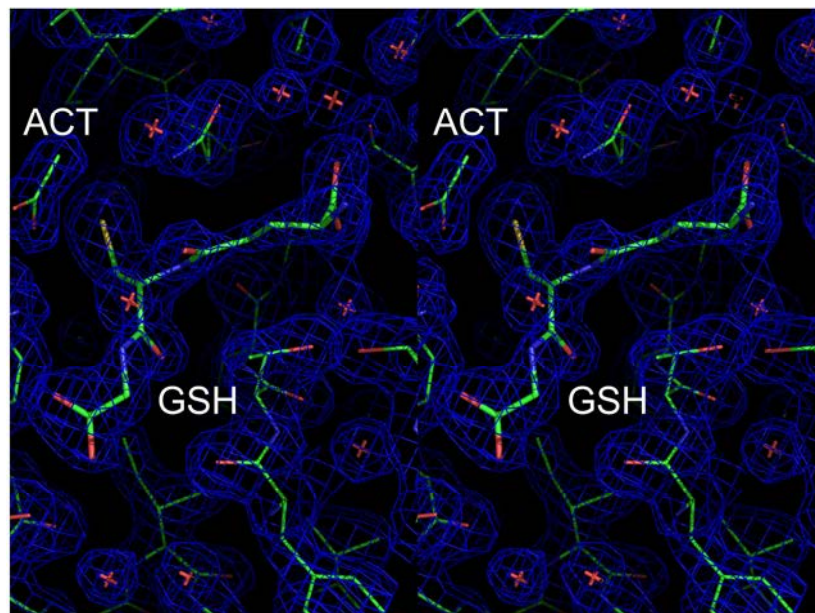
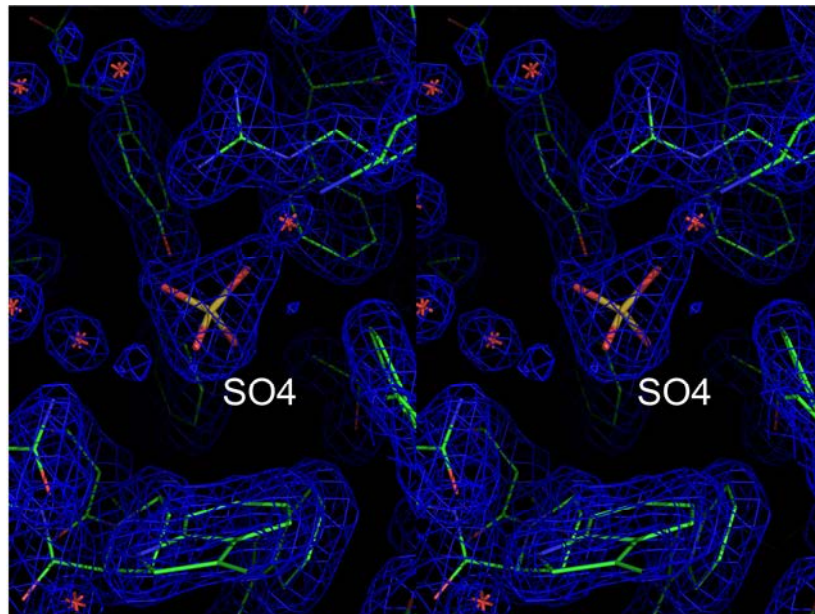
GSH transferase activity was monitored at 274 nm without (•) or with (triangles) 50  $\mu\text{M}$  ANS. Data is represented as mean  $\pm$  S.D. ( $n \pm 3$ ).

#### **Supplemental Figure S11. Compounds used in this study.**

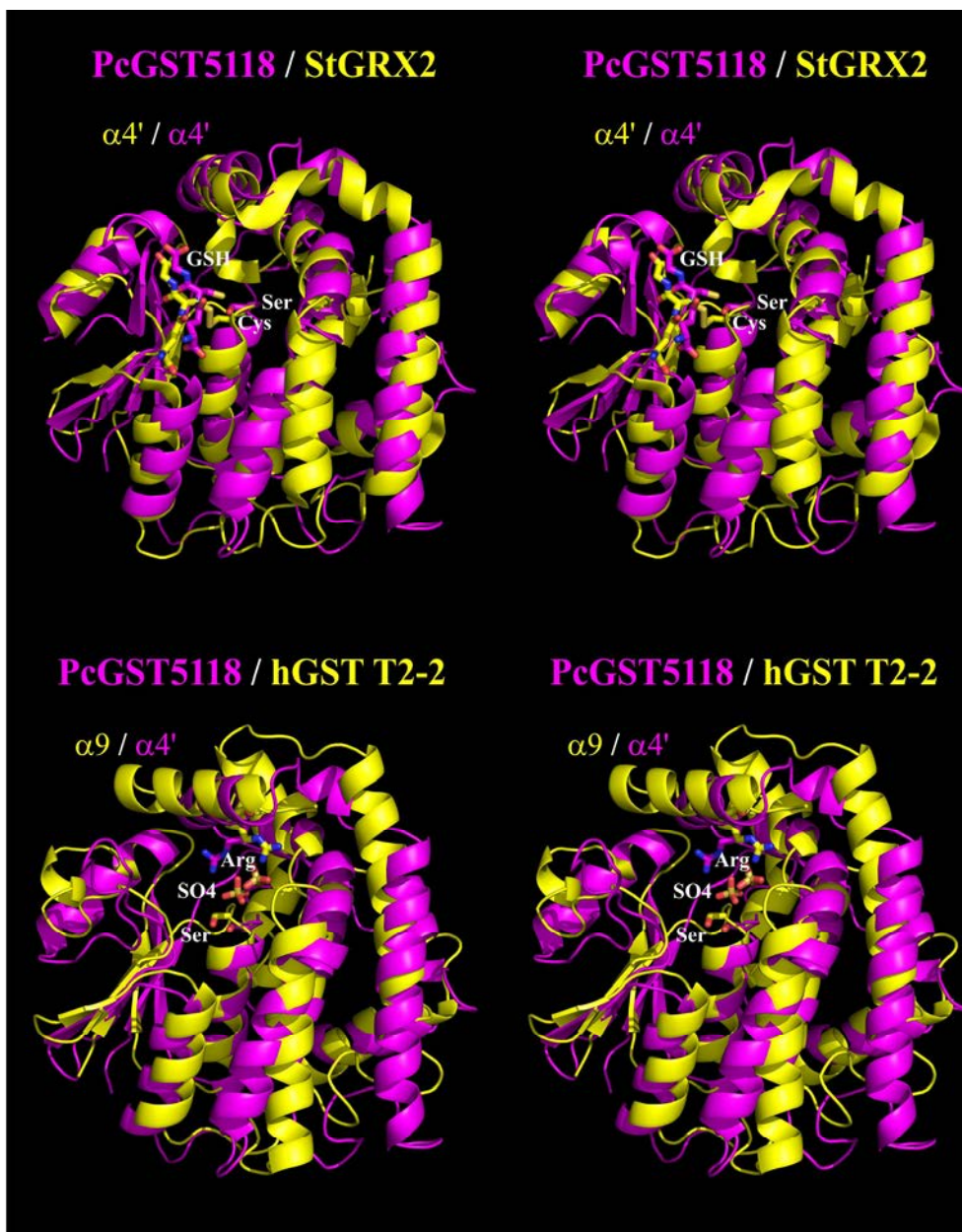




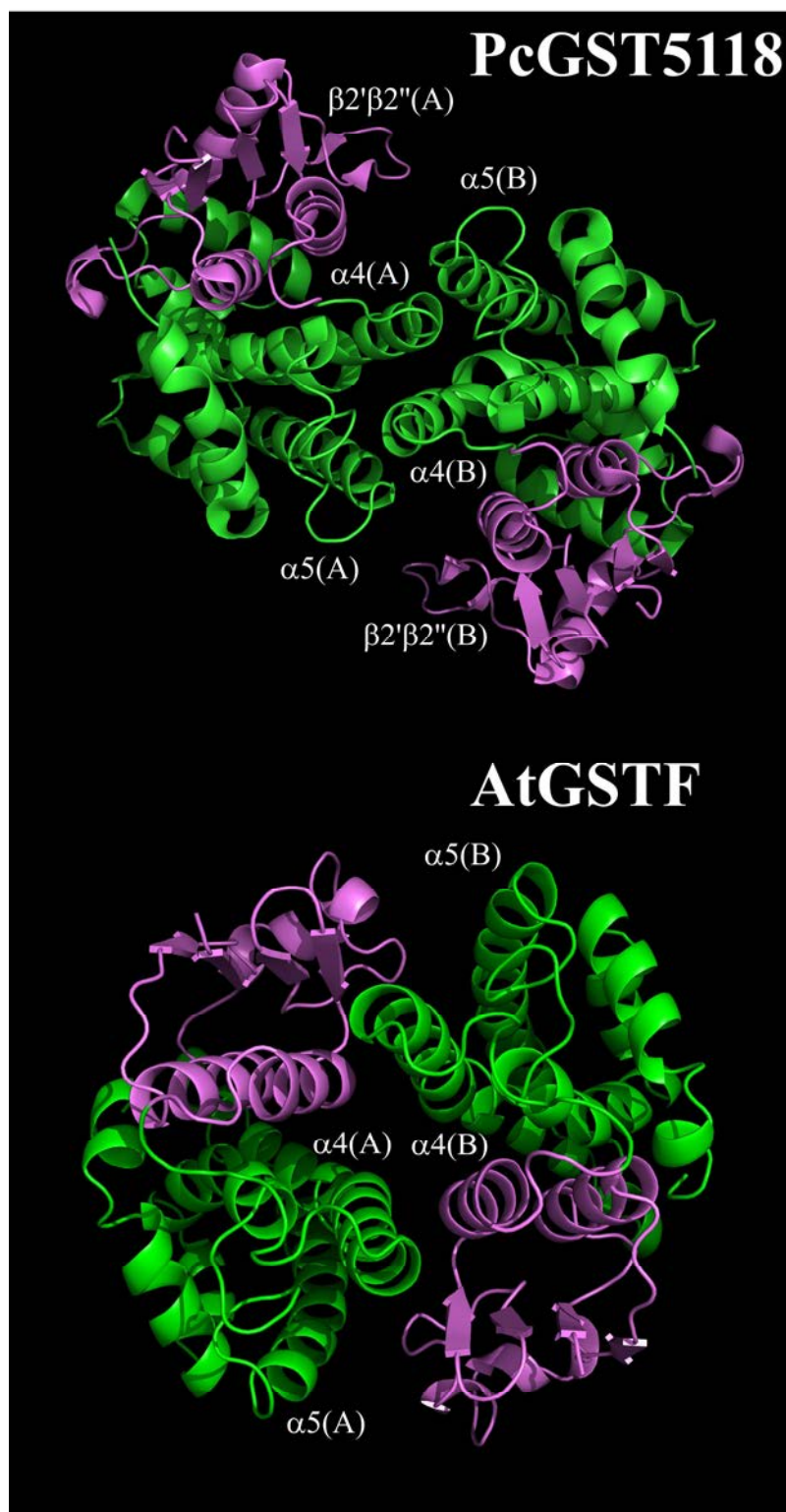
**Supplemental Figure S1**



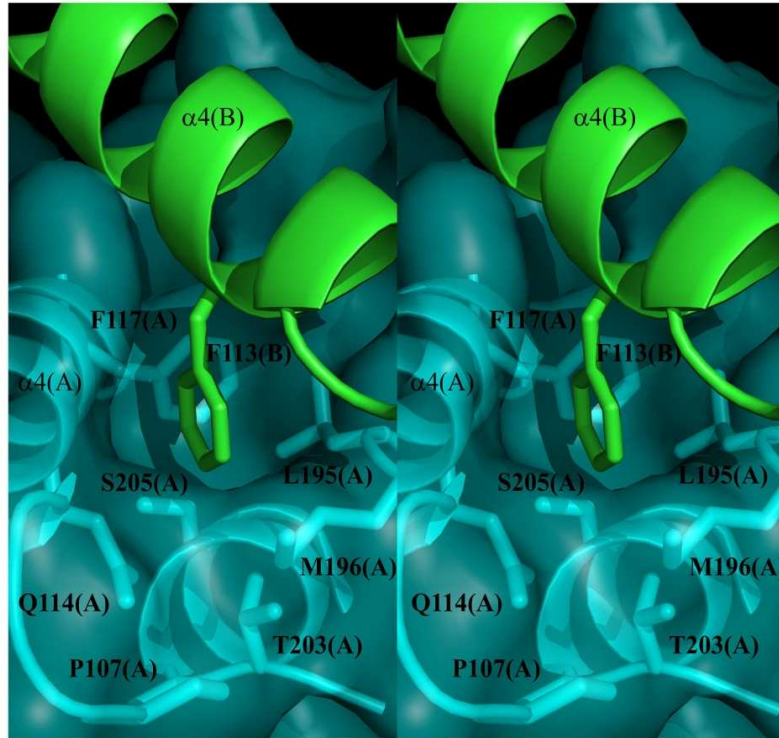
Supplemental Figure S2



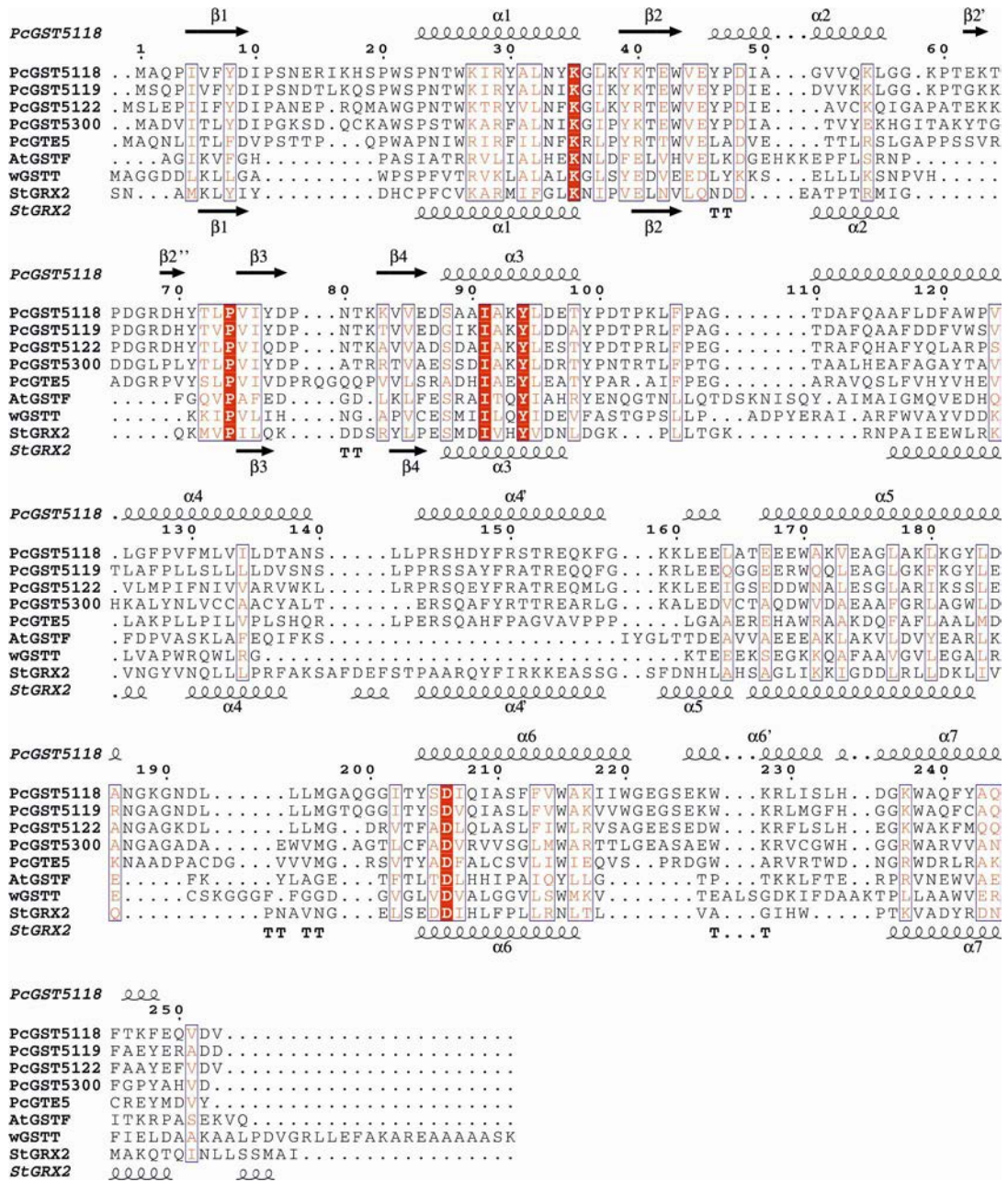
Supplemental Figure S3



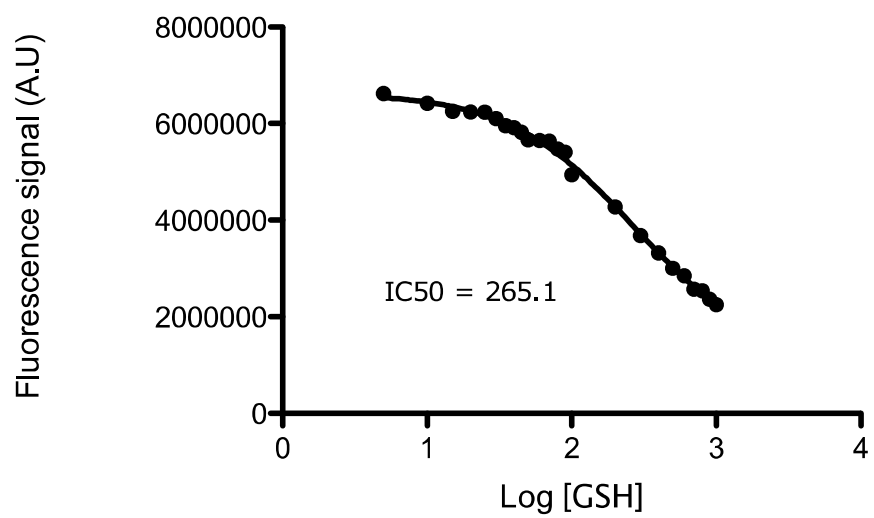
Supplemental Figure S4



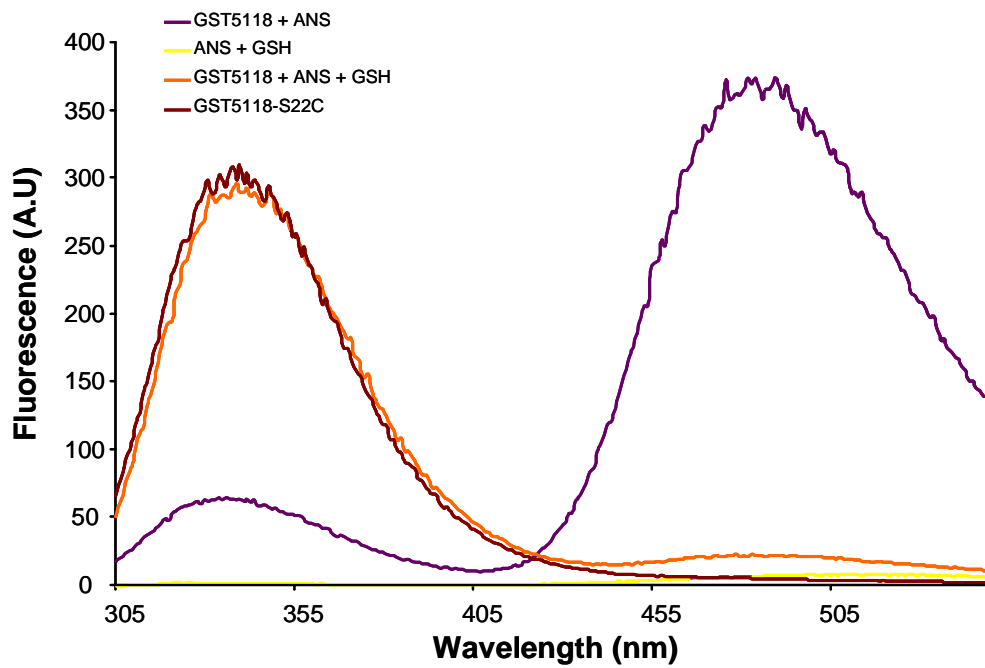
Supplemental Figure S5



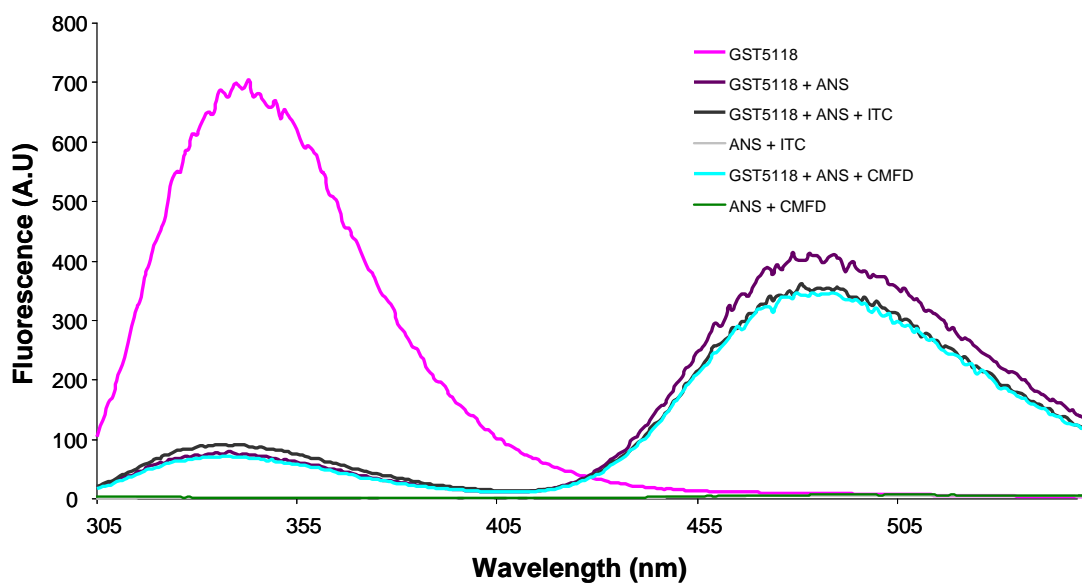
Supplemental Figure S6



**Supplemental Figure S7**

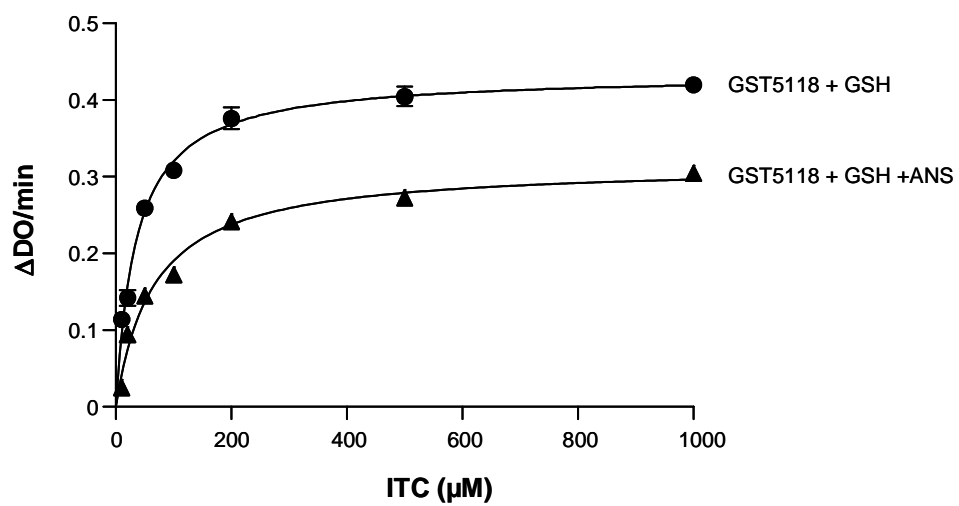


Supplemental Figure S8



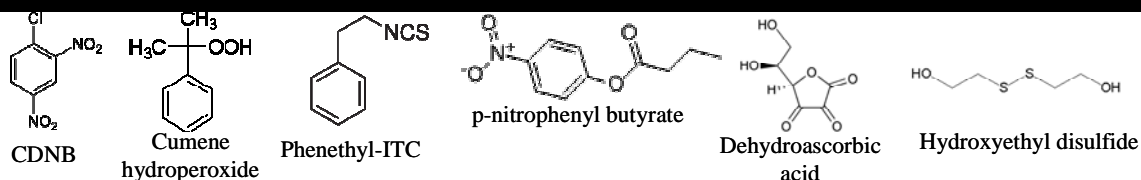
Supplemental Figure S9



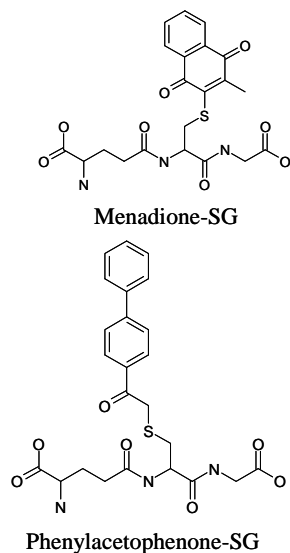


Supplemental Figure S10

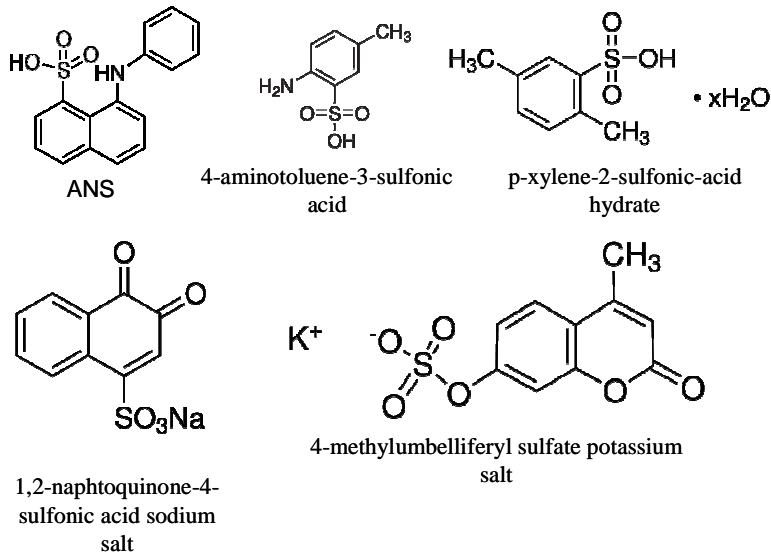
« classical » compounds



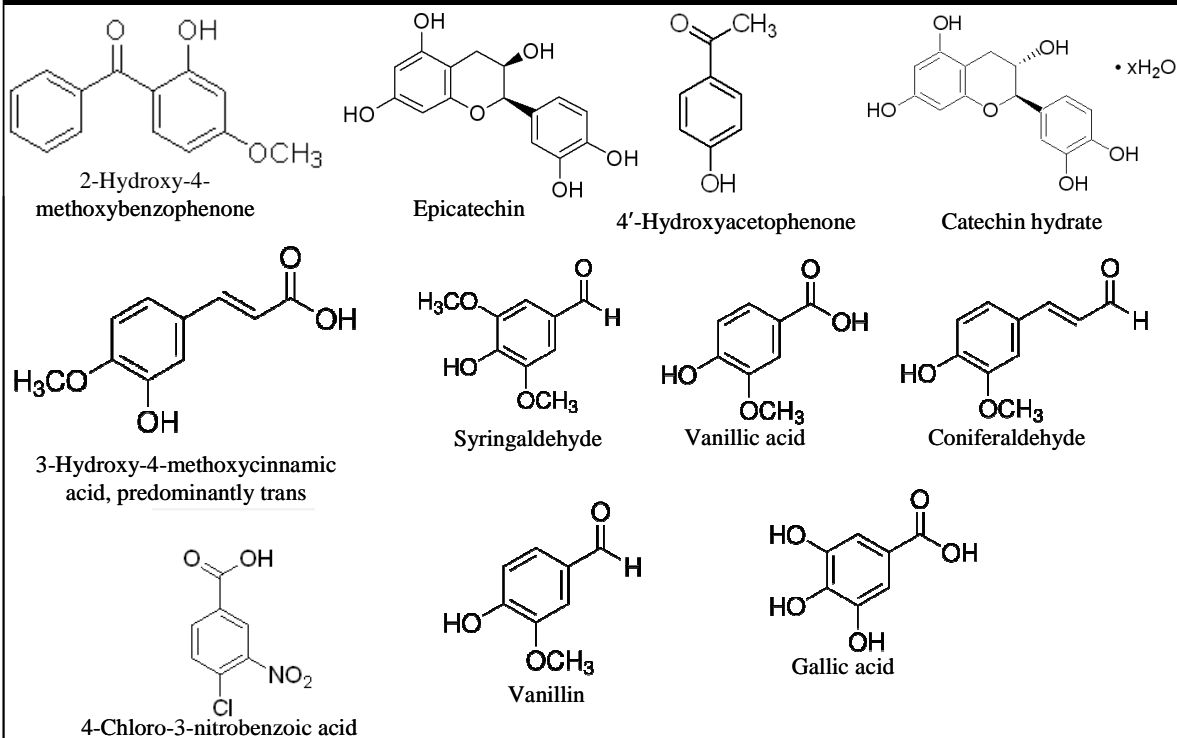
Glutathionylated compounds



Sulfur containing compounds



Beech extract compounds



Supplemental Figure S11

### Supplemental Table 1: Primers used in this study.

The *Nco*I and *Bam*HI cloning restriction sites are underlined in the primers. The mutagenic codons are in bold.

	Sequence
GST5118 for	5' <u>CCCCCATGGCTCAGCCCATCGTGTC</u> 3'
GST5118 rev	5' <u>CCCCGGATCCCTATACATCAACCTGCTC</u> 3'
GST5118-S22A for	5' AAGCACTCACCGTGGGCTCCCAATACCTGGAAA 3'
GST5118-S22A rev	5' TTTCCAGGTATTGGGAGCCACGGTGAGTGCTT 3'
GST5118-S22C for	5' AAGCACTCACCGTGGTGTCCCAATACCTGGAAA 3'
GST5118-S22C rev	5' TTTCCAGGTATTGGGACACCACGGTGAGTGCTT 3'

## **B.III.2 PcGSTFuA2 et PcGSTFuA3**

### **B.III.2.a Introduction**

L'article précédent présentait la caractérisation biochimique et structurale de la glutathion transférase (GST) PcGSTFuA1 de *Phanerochaete chrysosporium* (GSTFuA pour Fungal specific class A). Cette dernière appartient à une classe de GST spécifique aux champignons qui est particulièrement étendue dans les organismes impliqués dans la dégradation du bois. Même si la PcGSTFuA1 possède un repliement global de type GST comme attendu, celle-ci présente des particularités structurales uniques. De façon intéressante la mutation de la sérine catalytique putative en alanine ne rend pas l'enzyme inactive mais change sa spécificité de substrat et lui permet d'acquérir de nouvelles activités. Une analyse par fluorescence a révélé que le site G pourrait participer à la fixation de molécules issues de la dégradation du bois. Ce nouvel article compare huit GSTFuA, quatre isoformes de *P. chrysosporium* et quatre isoformes de *Coprinus cinereus*. Après une analyse génomique des séquences, les propriétés enzymatiques des huit isoformes sont présentées ainsi que leur capacité à fixer des ligands par fluorescence. Les structures cristallographiques des GSTFuA2 et GSTFuA3 de *P. chrysosporium* sont décrites en relation avec les résultats biochimiques. Enfin la famille des GSTFuA est discutée dans un contexte évolutif.

### **B.III.2.b Résultats et discussion**

L'analyse génomique des séquences de GSTFuA montre que cette classe de GST ne semble être reliée à la taxonomie des champignons mais plutôt à leur environnement et leur mode de vie. Ainsi, cette classe est plutôt présente dans les champignons interagissant avec le bois. A l'exception d'une GSTFuA (GSTFuA5 de *P. chrysosporium*) qui est présente dans pratiquement tous les organismes, les autres forment des cluster organisme par organisme, suggérant une diversification récente de cette nouvelle classe. La plupart des GSTFuA possèdent une sérine à la position du résidu catalytique des GST à sérine ou à cystéine. De plus la plupart des résidus impliqués dans le site anionique mis en évidence dans la structure cristallographique de la GSTFuA1 de *P. chrysosporium* sont conservés dans les GSTFuA.

La structure cristallographique de PcGSTFuA3 a été obtenue à haute résolution (2,1 Å) alors que celle de la PcGSTFuA2 est modeste (3,1 Å). Les PcGSTFuA2 et 3 présentent les mêmes particularités structurales que la PcGSTFuA1 avec des différences localisées en particulier dans le site H de l'enzyme. La grande hélice  $\alpha_4$  de la PcGSTFuA1 se présente en deux hélices  $\alpha_{4A}$  et  $\alpha_{4B}$  dans la PcGSTFuA2 car cette dernière présente une insertion d'un résidu dans cette région.

La longue boucle  $\beta 1-\alpha 1$  n'adopte pas la même conformation dans les PcGSTFuA1 et 3 probablement car la similarité de séquence est faible dans cette zone. Même si le glutathion n'est pas présent dans le site G des structures PcGSTFuA2 et 3, les deux isoformes sont plus proches de la forme holo de la PcGSTFuA1 (en présence de GSH) que de la forme apo. En effet, dans les deux cas le site G est occupé par des molécules de solvant provenant du milieu de cristallisation.

Les analyses biochimiques ont été réalisées à partir quatre GSTFuA de *P. chrysosporium* (PcGSTFuA1 à 4) et quatre GSTFuA de *C. cinereus* (CcGSTFUA2461, CcGSTFUA6800, CcGSTFUA6801 et CcGSTFUA6820). A l'exception de la CcGSTFUA682, toutes présentent au moins une activité avec les substrats testés même si une grande variabilité de spécificité de substrat est observée. Toutefois, l'activité thioltransférase n'a pas été détectée et seule la PcGSTFuA3 présente une activité peroxydase. Les études avec des sondes fluorescentes montrent que celles-ci interfèrent plutôt avec le site G (poche de fixation du glutathion) que le site H. La mise en compétition des sondes avec des composés issus de la dégradation révèle à nouveau une grande variabilité d'une GSTFuA à l'autre, en particulier pour celles de *C. cinereus*.

L'ubiquité de la classe GSTFuA dans les basidiomycètes saprophytes suggère que cette famille de protéine possède un rôle physiologique dans ces organismes. Les GSTFuA présentent une faible diversité structurale alors que cette diversité est importante en termes de spécificité de substrat et de spécificité de ligand. En conclusion, l'ensemble de ces résultats conduisent à supposer que la versatilité de la famille des GSTFuA dans les champignons saprophytes reflète l'adaptation de ces organismes face à la pression écologique induite par la variabilité des composés organiques issus de la dégradation du bois.

### **B.III.2.c Publication GSTFuA2-3**

Voir en page suivante le texte adressé à PLOS One pour publication le 26 Juillet 2013, et accepté sous réserve de corrections le 05 Septembre 2013.

## Diversification of fungal specific class A glutathione transferases in saprotrophic fungi

Yann Mathieu<sup>1,2,5§</sup>, Pascalita Prosper<sup>3,4§</sup>, Frédérique Favier<sup>3,4</sup>, Luc Harvengt<sup>5</sup>, Claude Didierjean<sup>3,4</sup>, Jean-Pierre Jacquot<sup>1,2</sup>, Mélanie Morel-Rouhier<sup>1,2</sup>, Eric Gelhaye<sup>1,2\*</sup>.

<sup>§</sup>The authors have equally participated to the work.

1 : Université de Lorraine, IAM, UMR 1136, IFR 110 EFABA, Vandoeuvre-les-Nancy, F-54506, France

2 : INRA, IAM, UMR 1136, Champenoux, 54280, France

3 : Université de Lorraine, CRM2, UMR 7036, Vandoeuvre-les-Nancy, F-54506, France

4 : CNRS, CRM2, UMR 7036, Vandoeuvre-les-Nancy, F-54506, France

5 : Laboratoire de biotechnologie, Pôle Biotechnologie et Sylviculture Avancée, FCBA, Campus Forêt-Bois de Pierroton, 33610 Cestas, France

\*Address correspondence to : Gelhaye Eric, Université de Lorraine, UMR 1136 “Interactions Arbres-Microorganismes”, Faculté des Sciences et Technologies BP 70239, 54506 Vandoeuvre-lès-Nancy Cedex, France; Tel + 33 3 83 68 42 28; E-mail : Eric.Gelhaye@univ-lorraine.fr

**Keywords : glutathione transferases ; ligandin ; saprotrophic fungi ; GSTFuA**

### Abstract

Glutathione transferases (GSTs) form a superfamily of multifunctional proteins with essential roles in cellular detoxification processes. The distribution of fungal-specific class A GSTs was investigated in saprotrophic fungi revealing a recent diversification within this class. Biochemical characterization of eight GSTFuA isoforms from *Phanerochaete chrysosporium* and *Coprinus cinereus* demonstrated functional diversity in saprotrophic fungi. The three-dimensional structures of three *P. chrysosporium* isoforms feature structural differences explaining the functional diversity of these enzymes. Competition experiments between fluorescent probes, and various molecules, showed that these GSTs function as ligandins with various wood extractives at a L-site overlapping the glutathione binding pocket. By combining genomic data with structural and biochemical determinations, we propose that this class of GST has evolved in response to environmental constraints induced by wood chemistry.

### Introduction

Glutathione transferases (GSTs; EC 2.5.1.18) constitute a large superfamily of multifunctional enzymes that are involved in phase II of xenobiotic detoxification by catalyzing the conjugation of glutathione (GSH) to a variety of electrophilic compounds [1,2]. Most of the GSTs are dimeric proteins, each monomer harbouring two domains : a thioredoxin N-terminal domain containing the conserved GSH binding site (G site) and a more variable C-terminal  $\alpha$ -helical domain (H site) usually involved in the binding of the GSH acceptor compound [3]. Besides their catalytic activities, some GSTs can also exhibit ligandin properties, involving a so-called L-site. The latter property has been defined as the capacity to bind non-substrate ligands contributing to intracellular sequestration and transport of xenobiotics or hormones [4]. In plants, GSTs could be involved in transport of hydrophobic compounds such as pigments [5]. More recently, the human glutathione transferase omega 1-1 was also shown to possess a L-site binding S-(4-nitrophenacyl)

1 glutathione at the dimer interface and was suggested to be the binding location of  
2 uncompetitive inhibitors such as tocopherol [6].

3 The release of fungal genomes allowed to unravel a specific evolution of cytosolic GSTs in  
4 these organisms in correlation with their way of life [2],[7]. Indeed, saprotrophic fungi such  
5 as the wood-decayer *P. chrysosporium* or the litter decomposer *C. cinereus* exhibit a high  
6 number of GST encoding genes in comparison to symbiotic fungi or biotrophic pathogens.  
7 The fungal specific class GSTFuA is particularly concerned by this expansion. For instance,  
8 *C. cinereus* possesses a high number of GSTFuA with 14 isoforms representing nearly the  
9 half of the total GSTs found in this organism (32 GST encoding sequences), whereas 5  
10 isoforms of GSTFuA, on a total of 25 GSTs, are present in *P. chrysosporium*. In a recent  
11 study, we have characterized for the first time a member of this fungal specific class in *P.*  
12 *chrysosporium* (PcGSTFuA1). PcGSTFuA1 displays unique structural and biochemical  
13 features, exhibiting overlapping G and L-sites [8].

14 The aim of this study was to extend the characterization of the GSTFuA class using  
15 comparative genomic, biochemical approaches performed on eight GSTFuA proteins (four  
16 from *P. chrysosporium* and four from *C. cinereus*) coupled with structural studies of  
17 three isoforms from *P. chrysosporium*.

## 19 **Material and methods**

20 - *Material* - Hydroxyethylidisulfide (HED) was from Pierce. 5-chloromethylfluorescein  
21 diacetate (CMFDA) was from Invitrogen. S-(phenylacetophenone)-glutathione (PAP-SG) and  
22 2-methyl-S-glutathionyl-naphthoquinone (menadione-SG) were synthesized as previously  
23 described [9]. All other reagents were from Sigma-Aldrich.

24  
25 - *Cloning of fungal GSTFUAs* - The open reading frame sequences were amplified from *P.*  
26 *chrysosporium* and *C. cinereus* cDNA libraries using forward and reverse primers (Table 1),  
27 and cloned into the NcoI and BamHI restriction sites (underlined in the primers) of pET-3d  
28 (Novagen). The amplified sequences encoded proteins in which an alanine has been inserted  
29 after the initiator methionine to improve protein production.

30  
31 - *Expression and purification of the recombinant proteins* - For protein production, the  
32 *Escherichia coli* BL21(DE3) strain, containing the pSBET plasmid, was co-transformed with  
33 the recombinant plasmids [10]. Cultures were progressively amplified up to 2 L in LB  
34 medium supplemented with ampicillin and kanamycin at 37 °C. Protein expression was  
35 induced in the exponential phase by adding 100 µM isopropyl β-D-thiogalactopyranoside for  
36 4 h at 37 °C. The cultures were then centrifuged for 15 min at 4400 × *g*. The pellets were  
37 resuspended in 30 mL of TE NaCl (30 mM Tris-HCl, pH 8.0, 1 mM EDTA, 200 mM NaCl)  
38 buffer. Cell lysis was performed on ice by sonication (3 × 1 min with intervals of 1 min), and  
39 the soluble and insoluble fractions were separated by centrifugation for 30 min at 27,000 × *g*  
40 at 4°C. The soluble part was then fractionated with ammonium sulphate in two steps, and the  
41 protein fraction precipitating between 40 and 80 % of the saturation contained the  
42 recombinant protein, as estimated by 15 % SDS-PAGE. The protein was purified by size  
43 exclusion chromatography after loading on an ACA44 (5 × 75 cm) column equilibrated in TE  
44 NaCl buffer. The fractions containing the protein were pooled, dialyzed by ultrafiltration to  
45 remove NaCl, and loaded onto a DEAE-cellulose column (Sigma) in TE (30 mM Tris-HCl,  
46 pH 8.0, 1 mM EDTA) buffer. The proteins were eluted using a 0 – 0.4 M NaCl gradient.  
47 Finally, the fractions of interest were pooled, dialyzed, and concentrated by ultrafiltration  
48 under nitrogen pressure (YM10 membrane; Amicon). Purity was checked by SDS-PAGE.  
49 Protein concentrations were determined spectrophotometrically using a molar extinction  
50 coefficient at 280 nm of 68870 M<sup>-1</sup>.cm<sup>-1</sup> for PcGSTFuA1, 67380 M<sup>-1</sup>.cm<sup>-1</sup> for PcGSTFuA2,

1 58900 M<sup>-1</sup>.cm<sup>-1</sup> for PcGSTFuA3, 75860 M<sup>-1</sup>.cm<sup>-1</sup> for PcGSTFuA4, 68410 M<sup>-1</sup>.cm<sup>-1</sup> for  
2 CcGSTFuA2461, 67380 M<sup>-1</sup>.cm<sup>-1</sup> for CcGSTFuA6800, 69900 M<sup>-1</sup>.cm<sup>-1</sup> for CcGSTFuA6801  
3 and 66350 M<sup>-1</sup>.cm<sup>-1</sup> for CcGSTFuA 6820.

4  
5 - *Sequence analysis* - All sequences were retrieved from the Joint Genome Institute (JGI)  
6 database (<http://genome.jgi-psf.org/programs/fungi/index.jsf>) from the fungal genomic  
7 database MycoCosm. The sequences have been obtained with Blastp (BLOSUM matrix  
8 applying default parameters) using all *P. chrysosporium* GSTFuA sequences as template.  
9 Fungi showed in the phylogenetic analysis have been chosen according to their saprotrophic  
10 properties. The sequence of PcGSTFuA4 has been modified compared to the one available on  
11 JGI and the new sequence has been deposited to Genbank under the KC192375 identity  
12 number. Sequence alignments have been done using ClustalW and phylogenetic analysis were  
13 conducted using Muscle and MEGA5 (Neighbour-joining) programs [11]. The pairwise  
14 deletion option was active to handle alignment gaps and the Poisson correction model was  
15 used for distance computation. Bootstrap tests were conducted using 1000 replicates.

16  
17 - *Activity measurements* - The activities of purified proteins in thiol transferase activity with  
18 hydroxyethyl disulfide (HED) assay or for reduction of dehydroascorbate (DHA) were  
19 performed as described by Couturier and coworkers [12]. The GSH transferase activity was  
20 assessed with phenethyl isothiocyanate (phenethyl-ITC) prepared in 2% (v/v) acetonitrile, 1-  
21 chloro-2,4-dinitrobenzene (CDNB) and 4-nitrophenyl butyrate (PNP-butyrate) prepared in  
22 DMSO. For these three substrates the reactions were monitored respectively at 274 nm, 340  
23 nm and 412 nm following the increase in absorbance arising from the formation of the S-  
24 glutathionylated adduct. The reactions with CDNB and PNP-butyrate were performed in 100  
25 mM phosphate buffer, pH 7.5, in presence of GSH (5 mM) while the reaction with phenethyl-  
26 ITC was performed at pH 6.5 with an identical GSH concentration. Peroxidase activities were  
27 monitored as follows: 1 mM peroxide (hydrogen peroxide, tertbutyl hydroperoxide or cumene  
28 hydroperoxide) in 30 mM Tris-HCl, pH 8.0, was incubated in presence of 2 mM GSH, 200  
29 μM NADPH, 0.5 IU glutathione reductase. The activity was followed by monitoring the  
30 decrease in absorbance arising from NADPH oxidation in this coupled enzyme assay system  
31 showing the formation of oxidized glutathione (GSSG). All the reactions were started by  
32 addition of the purified enzyme (500 nM) and monitored with a Cary 50 UV-Visible  
33 spectrophotometer (VARIAN). To determine the catalytic properties of the enzyme, steady  
34 state assays were performed using variable substrate concentrations (from 10 μM to 10 mM)  
35 and the catalytic parameters were calculated using the GraphPad® software.

36  
37 - *8-anilino-1-naphthalene sulfonic acid (ANS) binding and ligand screening* – All binding and  
38 competition experiments were performed in TE buffer pH 8.0. ANS binding onto purified  
39 proteins was investigated by monitoring fluorescence upon addition of various ANS  
40 concentrations, depending on the enzyme, to 3 μM protein using a fluorescence  
41 spectrophotometer Cary Eclipse (VARIAN). The excitation wavelength was set at 385 nm  
42 and the emission spectra recorded from 400 to 700 nm. Additional tryptophan-based  
43 fluorescence experiments were performed in presence or absence of ANS. After excitation at  
44 290 nm, emission spectra were recorded from 305 to 550 nm. ANS binding at the H-site was  
45 investigated in the same conditions with 50 μM CMFDA.

46 The correlation between ANS concentration and fluorescence yield was obtained by  
47 incubating, 100 μM of enzyme with 1 μM ANS, giving the maximal value for 1 μM ANS  
48 bound to the protein. Fluorescence from samples containing ANS alone was subtracted as a  
49 background. Proteins were then incubated at a concentration of 3 μM with increasing ANS  
50 concentrations ranging from 0 to 1 mM. To determine the dissociation constant of ANS the



1 following equation (Equation 1) was applied  $y = B_{max} * [Free\ ANS]/(K_d + [Free\ ANS])$   
2 where  $y$  is the specific binding, i.e. the concentration of ANS bound per protein and  $B_{max}$  the  
3 maximum number of specific binding sites.

4 Competition experiments were performed in a total volume of 250  $\mu$ L containing 3  $\mu$ M  
5 enzyme, variable ANS concentration depending on the enzyme, and various putative ligands  
6 (from 1  $\mu$ M to 10 mM). Fluorescence emission from protein-free samples containing ANS  
7 and potential ligand were subtracted as background. A decrease in fluorescence, indicating  
8 competition between ligand and ANS, was measured after 10 min of incubation with  
9 excitation at 385 nm and emission at 475 nm on a VICTOR™ X5 plate reader (PerkinElmer).  
10 IC50 values were obtained by fitting data to the equation (Equation 2)  $Y = Bottom + (Top -$   
11  $Bottom)/(1 + 10^{(Log[Ligand] - Log\ IC50)})$  where  $Y$  is the fluorescence signal observed after  
12 background subtraction,  $Top$  is the fluorescence signal of ANS without ligand and  $Bottom$  the  
13 fluorescence signal of ANS at the highest ligand concentration.

14  
15 - *Esterase activity: fluorescence based assays* - Esterase activity was measured in microplates  
16 using 5-chloromethylfluorescein diacetate (CMFDA) or 4-Methylumbelliferyl acetate (MUA)  
17 as substrates, the cleavage of which releases the fluorescent compounds fluorescein and 4-  
18 methylumbelliferone respectively. Experiments were performed on a VICTOR™ X5 plate  
19 reader (PerkinElmer) in 50 mM phosphate buffer pH 8.0 for CMFDA and 30 mM Tris-HCl  
20 buffer pH 8 for MUA. Both assays were carried out in a total volume of 200  $\mu$ L. The  
21 reactions were started by addition of the purified enzyme and fluorescence was measured  
22 every minute for 1 hour with excitation at 485 nm and emission at 535 nm for CMFDA. For  
23 MUA, excitation and emission were at 355 nm and 460 nm respectively. Catalytic parameters  
24 were calculated using the GraphPad® software from steady state experiments performed  
25 without or with 5 mM GSH using substrates ranging from 0.25 to 50  $\mu$ M for CMFDA and 10  
26 to 800  $\mu$ M for MUA.

27  
28 - *Inhibition Kinetics* – Competition assays between CcGSTFuAs and wood extractives were  
29 performed by measuring esterase activity on MUA without glutathione. Activities were  
30 monitored in a total volume of 200  $\mu$ L containing 0 to 500  $\mu$ M MUA and 0 to 10 mM of  
31 wood extractive. The reactions were started by the addition of purified enzyme (2  $\mu$ M), and  
32 the catalytic parameters were calculated using the GraphPad® software. For noncompetitive  
33 inhibition,  $K_i$  values were calculated resolving the non-linear noncompetitive inhibition  
34 equation:  $V_{maxInh} = V_{max}/(1 + I/K_i)$ ;  $Y = V_{maxInh} \times X/(K_m + X)$ .  $V_{max}$  and  $K_m$  represent the  
35 maximum enzyme velocity and the Michaelis-Menten constant without inhibitor, whereas  
36  $V_{maxInh}$  represents the maximum enzyme velocity for one concentration of inhibitor, and  $K_i$  is  
37 the inhibition constant. For competitive inhibition, the  $K_i$  value was calculated resolving the  
38 non-linear competitive inhibition equation:  $K_{m(obs)} = K_m \times (1 + [I]/K_i)$ ;  $Y = V_{max} \times X/(K_{m(obs)} +$   
39  $X)$ .  $V_{max}$  and  $K_m$  represent the maximum enzyme velocity and the Michaelis-Menten constant  
40 without inhibitor, whereas  $K_{m(obs)}$  represents the Michaelis-Menten constant in the presence of  
41 inhibitor, and  $K_i$  is the inhibition constant.

42  
43 - *Crystallization, Data Collection, Structure Determination and Refinement* - Crystallization  
44 experiments of PcGSTFuA2 and PcGSTFuA3 were performed by the microbatch under oil  
45 (paraffin) method at 4°C and 20°C, respectively. Drops were prepared by mixing equal  
46 volumes (2  $\mu$ l) of protein and precipitating agent solutions. TE buffered solutions of  
47 PcGSTFuA2 and PcGSTFuA3 had protein concentrations of 20 mg.mL<sup>-1</sup> and 28.3 mg.mL<sup>-1</sup>,  
48 respectively. Suitable crystals for x-ray diffraction were obtained without optimization using  
49 Hampton crystallization screen 1. PcGSTFuA2 crystallized in condition 9 (0.2 M ammonium  
50 acetate, 0,1 M sodium citrate tribasic dehydrate, pH 5,6 and 30% w/v polyethylene glycol

1 4,000) and PCGSTFuA3 crystals were obtained in condition 34 (0.1 M sodium acetate  
2 trihydrate, pH 4.6, 2.0 M sodium formate solution). X-ray diffraction experiments were  
3 performed at 100 K with crystals soaked briefly in crystallization solution supplemented with  
4 20% (v/v) glycerol prior to flash cooling.

5 X-ray data sets were collected on beamline PROXIMA-1 at the French national Synchrotron  
6 facility SOLEIL (Gif-sur-Yvette, France). The data were indexed and processed using XDS  
7 [13], scaled and merged with SCALA from the CCP4 program package [14,15]. Both  
8 structures were solved by molecular replacement with MOLREP [16] using the coordinates of  
9 PcGSTFUA1 in complex with GSH (Protein Data Bank code 4G19). The models were  
10 refined using PHENIX [17] interspersed with manual inspection and corrections using COOT  
11 [18]. The validation of both crystal structures was performed with MOLPROBITY [19]. All  
12 figures were prepared with PyMOL and crystallographic statistics are available in Table 2.  
13 Atomic coordinates and structure factors for PcGSTFuA2 and PcGSTFuA3 have been  
14 deposited in the Protein Data Bank (PDBID: 4MLV and 4MLW, respectively).

## 15 **Results**

16 - *Genomic analysis of GSTFUA coding sequences* - The fungal genomic portal access  
17 MycoCosm on JGI is currently providing an increasing dataset of fungal genome sequences.  
18 GSTFuA sequences have been collected from the JGI database in all available genomes of  
19 Chytridiomycota, Mucoromycotina, Saccharomycotina, Pezizomycotina, Pucciniomycotina,  
20 Ustilagomycotina and Agaricomycotina. The number of detected GSTFuA sequences is  
21 highly variable between and inside fungal phyla and is not related to the number of predicted  
22 gene models in the genome (Supplemental Table S1). Among Ascomycota, no GSTFuA-  
23 related sequence was identified in Saccharomycotina genomes, while 0 to 6 sequences were  
24 found in Pezizomycotina genomes. Similarly, among Basidiomycota, GSTFuA sequences  
25 were only detected in Agaricomycotina genomes (from 0 to 17 sequences). The distribution of  
26 this GST class is thus not correlated to fungal taxonomy but rather to the fungal way of life  
27 since, for instance, wood interacting fungi exhibit 2 to 17 GSTFuA sequences according to  
28 the species, whereas pathogens, such as *Puccinia graminis*, do not display any GSTFuA  
29 sequences. The phylogenetic analysis presented in Figure 1 focuses on Agaricomycotina  
30 GSTFuAs, including also some bacterial etherases (LigE and LigF from *Sphingobium* sp.  
31 SYK-6) since these proteins involved in lignin breakdown display homology with GSTFuAs.  
32 Globally, the sequences cluster according to the organism. However, a specific subclass could  
33 be identified which includes PcGSTFuA5 and its homologues. One isoform belonging to this  
34 subclass has been identified in each studied genome, except in *Gloeophyllum trabeum* and  
35 *Punctularia strigosozonata*. The putative proteins of this subclass are predicted to be located  
36 in mitochondria, suggesting that they could have a conserved function in this organelle.  
37 Figure 1 highlights the amino acids surrounding the putative catalytic residue. Most of the  
38 sequences exhibit a serine, which is known to be involved in the classical glutathione  
39 transferase activity [20]. However, some changes could be detected in PcGSTFuA3 (serine to  
40 glycine), CcGSTFuA6820 (serine to asparagine), and in the proteins belonging to the  
41 PcGSTFuA5 subclass (serine to alanine) suggesting differential catalytic properties (Figure  
42 2).  
43

44 In previous work a sulphate binding pocket has been identified in the crystal structure  
45 of PcGSTFuA1. The sulphate ion was located near the catalytic residue Ser-22 and was  
46 stabilized by the side chains of Tyr-46 and Arg-153. Globally among GSTFuAs investigated  
47 in this study, these residues are well conserved, except for the CcGSTFuA2461 tyrosine,  
48 which is substituted by a phenylalanine (Figure 2).  
49

1 - *Protein production* - To investigate functionally this GST class, four isoforms of *P.*  
2 *chryso sporium* were produced in *E. coli* and purified. Unfortunately PcGSTFuA5 could not  
3 be produced whatever the culture conditions used (solid or liquid cultures, static or agitated  
4 conditions, rich (malt) or minimal medium, wood or anthracene as substrate). In addition, five  
5 isoforms from *C. cinereus* were selected over the 14 encoded in the genome, to represent the  
6 general diversity of GSTFuA found in this organism. Unfortunately, even if the homologue of  
7 PcGSTFuA5 (CcGSTFuA6672) was successfully amplified and expressed in *E. coli*, the  
8 resulting protein remained insoluble despite various tested protocols, preventing biochemical  
9 characterization of any isoform belonging to this subclass.

10  
11 - *Structural diversity* – Among the 5 existing isoforms in *P. chryso sporium*, two three-  
12 dimensional structures were resolved (PcGSTFuA2 and PcGSTFuA3). PcGSTFuA2  
13 crystallized in the space group  $C222_1$  with three dimers in the asymmetric unit. Global NCS  
14 restraints were imposed during refinement because diffraction limit of the crystals was 3.2 Å.  
15 The structure of PcGSTFuA3 was solved at high resolution (2.1 Å) from crystals having the  
16  $P6_122$  space group with one polypeptide chain in the asymmetric unit. As assumed in our  
17 recent report [8] PcGSTFuA2, PcGSTFuA3 and PcGSTFuA1 share the same structure, which  
18 contains distinctive features when compared with canonical GST structure [4] (Figure 3) : (i)  
19 a new mode of dimerization, (ii) an original  $\beta$ -hairpin ( $\beta 2' \beta 2''$ ) that inhibits the formation of  
20 the regular GST dimer and acts as a lid over the G-site and (iii) a supplementary helix ( $\alpha 4'$ ),  
21 between helices  $\alpha 4$  and  $\alpha 5$ , that closes the presumed H-site.

22 In our recent study [8], we described the crystal structures of PcGSTFuA1 with and  
23 without GSH named apo and holo forms, respectively. The analysis of the models revealed  
24 significant conformational changes of the enzyme upon GSH binding. On the one hand the  $\beta$ -  
25 hairpin ( $\beta 2' \beta 2''$ ) closes the G-site and a number of residues modify their side-chain  
26 conformations to interact with glutathione. On the other hand two residues open slightly the  
27 H-site by changing their lateral-chain conformations. Structural superpositions of apo/holo  
28 PcGSTFuA1, PcGSTFuA2 and PcGSTFuA3 unambiguously showed that PcGSTFuA2 and  
29 PcGSTFuA3 are in the holo conformation in the conditions used. However, in both structures,  
30 no peak could be assigned to glutathione in the electron density maps. The observed holo  
31 conformations can be explained by the presence of crystallization agent molecules in the G  
32 sites, which mimic some of the interactions formed by bound GSH. A citrate ion and three  
33 formate molecules were assigned in the PcGSTFuA2 and PcGSTFuA3 structures,  
34 respectively (Figure 4).

35 The distance plot, which compares superposed structures of holo PcGSTFuA1 and  
36 PcGSTFuA2 (Supplemental Figure S1), reveals no big differences. The largest is localized in  
37 the center of the helix  $\alpha 4$  where a residue is inserted in PcGSTFuA2 sequence (Figures 3 and  
38 4B). This difference results in a disruption in the continuity of helix  $\alpha 4$ . This region has to be  
39 considered as two helices in PcGSTFuA2 ( $\alpha 4A$  and  $\alpha 4B$ , Figure 3). It is worth noting that the  
40 disruption region is near the presumed H-site. This structural diversity could explain  
41 functional variabilities between PcGSTFuA1 and PcGSTFuA2. The distance plot, which  
42 compares holo PcGSTFuA1 and PcGSTFuA3, revealed a significant difference in the  $\beta 1$ - $\alpha 1$   
43 loop. It is linked to the low sequence identity between PcGSTFuA1 and PcGSTFuA3 in this  
44 region and to a one-residue shorter loop in PcGSTFuA3 (Figures 3 and 4C). The  $\beta 1$ - $\alpha 1$  loop  
45 in PcGSTFuA3 is longer than that of almost all GSTs and reduces the solvent accessibility of  
46 the H-site as observed for  $\beta 1$ - $\alpha 1$  loop in PcGSTFuA1 [8].

47 The comparison of the quaternary structures of PcGSTFuA1 and PcGSTFuA2  
48 revealed little difference. The V-shaped PcGSTFuA3 dimer is roughly 5° smaller than that  
49 observed in PcGSTFuA1. Among the two hydrophilic patches involved in dimer stabilization,  
50 only one consists of residues conserved in both isoforms. The consequence is an increase in

1 the number of intermolecular hydrogen bonds from 5 (in PcGSTFuA1 dimer) to 12 (in  
2 PcGSTFuA3) leading to a more stable dimer. The residues involved in the hydrophobic lock-  
3 and-key motif are well conserved in PcGSTFuA1, PcGSTFuA2 and PcGSTFuA3.

4 In the active site, Ser-22 (PcGSTFuA1 numbering), which catalyses the GSH addition  
5 on ITC in PcGSTFuA1, is present in PcGSTFuA2 but absent in PcGSTFuA3 (Figure 2).  
6 However, Asn-24 (PcGSTFuA1 numbering), which has been assumed to be involved in the  
7 peroxidase activity of PcGSTFuA1, is conserved in the three isoforms. This residue exhibits  
8 the same conformation in all three structures (Figure 4). Comparison of the active sites  
9 revealed that the residues involved in the GSH and in the sulphate binding sites in  
10 PcGSTFuA1 are well conserved and could play the same role in PcGSTFuA2 and  
11 PcGSTFuA3. In the structure of PcGSTFuA2, a citrate molecule mimics the binding of GSH  
12 glycyl moiety. The citrate ion stabilization involves the main chains of Tyr-71 and Val-73 and  
13 the side-chains of Tyr-46, His-70 and Arg-154 (Figure 4B). In PcGSTFuA3, three formate  
14 ions mimic the binding of the two GSH carboxylate groups and of the GSH thiolate group. A  
15 fourth formate molecule occupies the position of the sulphate ion in PcGSTFuA1 (Figure 4C).  
16 The structural variability is rather localized in the presumed electrophilic substrate binding  
17 site even though most of the residues are hydrophobic in the three isoforms. The local  
18 structure diversity is highlighted in Figure 4.

19  
20 - *Diversity of GSTFuAs activities* – The catalytic pattern of the different enzymes was studied  
21 using various substrates. Glutathione transferase activity was tested using 1-chloro-2,4-  
22 dinitrobenzene (CDNB), phenethyl isothiocyanate (phenethyl-ITC) and 4-nitrophenyl  
23 butyrate (PNP-butyrate). PcGSTFuAs, although clustering together in our phylogenetic  
24 analysis, exhibited various catalytic patterns (Table 3). PcGSTFuA1 displayed various  
25 activities in contrast to PcGSTFuA2 (both proteins having 74% of identity), which remained  
26 inactive against all tested substrates. Moreover, PcGSTFuA1 and PcGSTFuA3 were active as  
27 GSH transferase with phenethyl-isothiocyanate and PNP-butyrate.

28 This catalytic variability was also observed when comparing the *C. cinereus* isoforms.  
29 CcGSTFuA2461, CcGSTFuA6800 and CcGSTFuA6801 were active as GSH transferase  
30 against phenethyl-ITC, but only CcGSTFuA2461 and CcGSTFuA6801 were active with the  
31 universal GST substrate CDNB. In addition, CcGSTFuA6801 was able to cleave the ester  
32 linkage of PNP-butyrate.

33 None of the tested GSTFuAs exhibited thiol transferase activity using HED and DHA as  
34 substrates. PcGSTFuA3 was the only tested protein active against cumene hydroperoxide,  
35 whereas CcGSTFuA6820 remained inactive whatever the tested substrate.

36 CMFDA is a molecule releasing fluorescence upon activation by esterases [21]. As shown  
37 earlier [8], PcGSTFuA1 displayed activity against CMFDA, in presence or in absence of  
38 GSH. This feature is not shared by the other tested PcGSTFuAs as they exhibited an esterase  
39 activity only in presence of GSH (Table 3). Activity of CcGSTFuAs against CMFDA has  
40 been also tested demonstrating that only CcGSTFuA6800 and CcGSTFuA6801 were able to  
41 cleave this substrate either in absence or presence of GSH.

42  
43 - *Fluorescent probes to screen the ligandin property of GSTFuAs* – 8-anilino-1-naphthalene  
44 sulfonic acid (ANS) is an environment sensitive fluorescent dye used to characterize  
45 hydrophobic sites of proteins [22]. This probe has been successfully used to analyze the  
46 ligandin properties of PcGSTFuA1 [8]. Using a similar approach, we have tested ANS  
47 binding onto the different recombinant proteins, following the appearance of a characteristic  
48 fluorescence emission signal at 475 nm (excitation at 375 nm). Globally, all GSTFuAs were  
49 able to bind ANS, but some properties, i.e. number and position of the binding sites, varied on  
50 the different isoform (Figure 5). The number of ANS binding sites, determined by fitting data

1 of saturation experiments to Equation 1 detailed in material and methods section, is  
2 approximately equal to 1 for PcGSTFuAs, whereas CcGSTFuAs exhibited values below 0.7  
3 suggesting a lower affinity of these *C. cinereus* isoforms for the probe.

4 The ANS binding site onto PcGSTFuA1 has been shown to overlap the glutathione binding  
5 site [8]. Similarly, GSH inhibits ANS binding onto PcGSTFuAs and CcGSTFuA6801. The  
6 IC50 values determined after fitting the obtained data to Equation 2 (see materials and  
7 methods) are reported in Figure 5. These measured values are in accordance with the Km of  
8 GSH found in the enzymatic assays for PcGSTFUA1/3 and CcGSTFUA6801. In addition,  
9 PAP-SG and menadione-SG were able to inhibit ANS binding only for PcGSTFuAs, while  
10 the same compounds deprived of the GSH adduct could not. These results demonstrate that  
11 ANS binding overlaps at least partially the glutathione binding site onto PcGSTFuAs. In  
12 CcGSTFuAs this peculiar feature is not conserved anymore since, except for  
13 CcGSTFuA6801, GSH or glutathionylated compounds did not inhibit ANS binding.

14 To determine the ANS binding site onto PcGSTFuAs, we used the fluorescence properties of  
15 tryptophan (Figure 6). Tryptophan fluorescence of PcGSTFuAs was measured using 290 nm  
16 and 340 nm as excitation and emission wavelengths respectively in presence or absence of  
17 ANS. Fluorescence resonance energy transfer (FRET), revealed by the appearance of a signal  
18 at 475 nm, was observed between Trp and ANS confirming the binding of ANS onto the  
19 proteins, the energy transfer being highly dependent on the distance between the donor and  
20 acceptor molecules. Taking advantage of the FRET phenomena in presence of ANS,  
21 competition experiments have been performed to investigate putative overlap between the H-  
22 site and ANS binding site. From kinetic data, all PcGSTFuAs exhibit a high affinity for  
23 CMFDA with GSH (Table 3). ANS binding onto PcGSTFuAs remained unaltered in presence  
24 of CMFDA demonstrating that the ANS binding site does not interfere with CFMDA binding  
25 site (Figure 6). Taken together these results suggest that specific ANS binding at the G-site, as  
26 described previously for PcGSTFuA1, is a well-conserved feature in this class for  
27 *P. chrysosporium*. In contrast, for CcGSTFuAs, this feature is not so well conserved as shown  
28 by the lack of competition for protein binding between GSH or glutathionylated compounds  
29 and ANS in these isoforms, except for CcGSTFuA6801.

30 In order to investigate potential ligandin properties of GSTFuAs (i.e. determination of L-site),  
31 we used 4-Methylumbelliferyl acetate (MUA) as a probe, which requires the cleavage of one  
32 ester bond before its detection by fluorescence. Interestingly, all GSTFuAs were able to  
33 cleave MUA even in absence of glutathione, however with a very small catalytic efficiency  
34 (Table 4). Interestingly, the only enzyme inactive in the tested enzymatic assays  
35 (CcGSTFUA6820) turned out to be active against this compound without GSH,  
36 demonstrating that it still possesses the ability to recognize hydrophobic compounds.  
37 Addition of glutathione with MUA led to two distinct behaviours: (i) a loss or a strong  
38 alteration of the ability to cleave the ester bond of MUA (CcGSTFuA2461, CcGSTFuA6820,  
39 CcGSTFuA6800, PcGSTFuA2, PcGSTFuA3 and PcGSTFuA4) (Table 4) (ii) a strong  
40 increase in catalytic efficiency upon GSH addition without alteration of the apparent affinity  
41 towards MUA (CcGSTFuA6801, PcGSTFuA1).

42 These results demonstrate that except in CcGSTFuA6801 and PcGSTFuA1, MUA and  
43 glutathione binding sites overlap at least partially onto the tested proteins. In conclusion,  
44 experiments using either ANS or MUA demonstrate that the overlap of G-site and L-site is a  
45 conserved feature among the tested GSTFuAs.

46  
47 - *Diversity of GSTFuAs Ligandin properties* - Using ANS or MUA as screening tools,  
48 competition experiments have been performed for PcGSTFuAs and CcGSTFuAs with 11  
49 structurally variable beech compounds. Similarly to what was observed in enzymatic tests,  
50 GSTFuAs exhibited various ligandin profiles (Table 5). CcGSTFuA6820 was able to bind 8

1 compounds; PcGSTFuA1 and CcGSTFuA6801 were able to bind 7 compounds;  
2 PcGSTFuA2/3 and CcGSTFuA2461 were able to bind 4 while PcGSTFuA4 was able to bind  
3 3 of them. Among these products, those recognized by all PcGSTFuAs are mostly small  
4 aldehydes while CcGSTFuAs mainly recognized larger hydrophobic compounds or phenolic  
5 acids. Steric hindrance is a factor influencing ligand binding since epicatechin and catechin  
6 hydrate, two of the largest compounds used, are both bound by PcGSTFuA1,  
7 CcGSTFuA6801 and CcGSTFuA6820 but not by other GSTFuAs.

8 Taken together, our results suggest that the ligandin function linked to the G-site is present in  
9 enzymes of the GSTFuA class which can bind a variety of ligands especially in *C. cinereus*.

## 10 11 **Discussion**

12 Using comparative genomics, we show here that the GSTFuA class is widely  
13 distributed within Agaricomycotina. The expansion of this GST class seems to be linked to  
14 the fungal ability to cope with recalcitrant organic matter as wood or litter, independently of  
15 the degradation strategy (litter degraders, white or brown rotters). Among saprotrophic  
16 basidiomycetes, GSTFuA sequences cluster according to taxonomy suggesting a recent  
17 diversification of this fungal class as already observed for the fungal Ure2p class [23,24].  
18 Nevertheless, the ubiquity of GSTFuAs among saprotrophic basidiomycetes could suggest  
19 putative important physiological functions in these organisms.

20 Our biochemical and structural studies have been performed with eight isoforms  
21 produced from two different agaricomycetes: *P. chrysosporium* and *C. cinereus*. The obtained  
22 data demonstrate a large catalytic versatility of GSTFuAs associated with little structural  
23 diversity. GSTFuAs exhibited indeed different catalytic profiles, some of them being able to  
24 act as GSH transferases against various substrates (phenethyl-ITC, CDNB and PNP-butyrates),  
25 peroxidases and also esterases. However none presented thiol transferase activity.  
26 Surprisingly, three isoforms were inactive against the classical GST substrates, with  
27 nevertheless a weak esterase activity against CMFDA for two of them. These results suggest  
28 that evolution of multigenic families, and especially GSTs, results in a specialization of the  
29 various isoforms. Indeed, GSTs are often referred to as promiscuous enzymes with low  
30 specificity. Here our data suggest one clear primary function (conjugation of GSH with  
31 isothiocyanates) together with some other lower level activities (peroxidase or esterase),  
32 which can be regarded as a vestige of the broad specificity of ancestors or an emerging  
33 function due to neofunctionalization events.

34 Sequence alignment and structural data indicate that three potential residues could be  
35 involved in catalysis: Ser22, Asn24 and Tyr46 (PcGSTFuA1 numbering). Ser22 is involved  
36 in glutathionylation activity of GSTFuA1 [8], Asn24 seems to be crucial for the peroxidase  
37 activity of *Saccharomyces cerevisiae* Ure2p [25], while Tyr46 stabilizes the sulphate ion in  
38 PcGSTFuA1 [8]. None of these residues are strictly conserved within the GSTFuAs  
39 investigated in this study, in agreement with the different catalytic profiles observed. The  
40 mutant PcGSTFuA1 S22A generated in a previous study revealed the complex relationship  
41 between this catalytic diversity and primary sequence variability. Indeed, the resulting  
42 enzyme, which gained thiol transferase and peroxidase activities exhibited a catalytic profile  
43 relatively similar to PcGSTFuA3 that possesses a glycyl residue at the same position.

44 The inability of some of these enzymes to act efficiently as transferases raises the  
45 question of their function in the fungal cell. One particular feature of PcGSTFuA1 is the  
46 presence of a ligand binding site overlapping the G-site. This putative L-site is responsible for  
47 the binding of toxic wood compounds [8]. Here we show by using fluorescent probes that this  
48 feature is conserved in the eight studied isoforms. As for enzymatic activities, GSTFuAs  
49 exhibited versatility in this ligandin property. Variations of this ligandin property were  
50 highlighted by competition experiments between ANS or MUA and beech compounds. In

1 these competition assays, all GSTFuAs investigated possessed their own ligandin profile,  
2 including those that were inactive in enzymatic assays. Previously identified L-sites in GSTs  
3 are located in the dimer interface for the mu-class of the parasitic worm *Schistosoma*  
4 *japonicum* [26], the squid sigma-class [27], and the human omega class [6], while it overlaps  
5 the H-site for the phi-class of *Arabidopsis thaliana* [28] or the human mu-class [29]. Our  
6 competition experiments with fluorescent probes demonstrated that this L-site, in the  
7 GSTFuA class, is located in the vicinity of the G-site without overlapping the H-site, and is  
8 able to recognize a large variety of compounds possessing different chemical functions.

9 GSTs being detoxification enzymes, their diversity reflects the variable nature of the  
10 environment in which organisms need to survive. This statement is especially true for  
11 saprotrophic fungi since wood is a complex substrate. Indeed, wood chemical and anatomical  
12 traits influence its decomposition by fungi, leading to numerous compounds originating from  
13 lignin [30]. This variability in chemical composition and this heterogeneity imply that fungi  
14 must develop a way to respond to the myriad of potentially toxic wood compounds resulting  
15 from the variability observed from one tree to another. For this purpose, GSTFuAs have  
16 apparently evolved from needs related more to ligand recognition than conjugation of the  
17 compounds. Indeed, some isoforms were unable to conjugate GSH onto the tested  
18 hydrophobic compounds but possessed a ligandin property at least for these tested  
19 compounds. This variability observed among GSTFuAs might have been driven by survival  
20 strategies to provide a superior metabolic system to cope with plant-related compounds  
21 including lignin and/or its derivatives. This phenomena has already been noticed for  
22 cytochromes monooxygenases (P450) of *P. chrysosporium* [31].

23 Another example of driven evolution concerns the P450s from plants and insects.  
24 Plants have evolved P450s implicated in biosynthesis pathways of secondary metabolites as  
25 defence against insects while insects have evolved P450s to counter the toxic effects of these  
26 plants metabolites. Both evolutionary pathways involve continual gene duplications and  
27 mutations resulting in alteration of substrate range and catalytic properties [32].

28 In conclusion, we postulate that the versatility and promiscuity of the extended  
29 GSTFuA class in saprotrophic fungi reflects the adaptation of these organisms to ecological  
30 pressure induced by wood chemical variability.

31  
32 *Acknowledgments – The authors would like to thank the staff of the the Proxima 1 beamline at*  
33 *SOLEIL (Gif-sur-Yvette, France) for their kind assistance during data collection.*

## 34 35 36 **REFERENCES**

- 37  
38 1. Hayes JD, Flanagan JU, Jowsey IR (2005) Glutathione transferases. *Annu Rev Pharmacol*  
39 *Toxicol* 45: 51-88.
- 40 2. McGoldrick S, O'Sullivan SM, Sheehan D (2005) Glutathione transferase-like proteins  
41 encoded in genomes of yeasts and fungi: insights into evolution of a multifunctional  
42 protein superfamily. *FEMS Microbiol Lett* 242: 1-12.
- 43 3. Mannervik B (2012) Five decades with glutathione and the GSTome. *J Biol Chem* 287:  
44 6072-6083.
- 45 4. Oakley A (2011) Glutathione transferases: a structural perspective. *Drug Metab Rev* 43:  
46 138-151.
- 47 5. Conn S, Curtin C, Bezier A, Franco C, Zhang W (2008) Purification, molecular cloning,  
48 and characterization of glutathione S-transferases (GSTs) from pigmented *Vitis*  
49 *vinifera* L. cell suspension cultures as putative anthocyanin transport proteins. *J Exp*  
50 *Bot* 59: 3621-3634.

- 1 6. Brock J, Board PG, Oakley AJ (2013) Structural insights into omega-class glutathione  
2 transferases: a snapshot of enzyme reduction and identification of a non-catalytic  
3 ligandin site. *PLoS One* 8: e60324.
- 4 7. Morel M, Meux E, Mathieu Y, Thuillier A, Chibani K, et al. (2013) Xenomic networks  
5 variability and adaptation traits in wood decaying fungi. *Microb Biotechnol* 6: 248-  
6 263.
- 7 8. Mathieu Y, Prosper P, Buee M, Dumarcay S, Favier F, et al. (2012) Characterization of a  
8 *Phanerochaete chrysosporium* glutathione transferase reveals a novel structural and  
9 functional class with ligandin properties. *J Biol Chem* 287: 39001-39011.
- 10 9. Meux E, Prosper P, Ngadin A, Didierjean C, Morel M, et al. (2011) Glutathione  
11 transferases of *Phanerochaete chrysosporium*: S-glutathionyl-p-hydroquinone  
12 reductase belongs to a new structural class. *J Biol Chem* 286: 9162-9173.
- 13 10. Schenk PM, Baumann S, Mattes R, Steinbiss HH (1995) Improved high-level expression  
14 system for eukaryotic genes in *Escherichia coli* using T7 RNA polymerase and rare  
15 Arg tRNAs. *Biotechniques* 19: 196-198, 200.
- 16 11. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, et al. (2011) MEGA5: molecular  
17 evolutionary genetics analysis using maximum likelihood, evolutionary distance, and  
18 maximum parsimony methods. *Mol Biol Evol* 28: 2731-2739.
- 19 12. Couturier J, Koh CS, Zaffagnini M, Winger AM, Gualberto JM, et al. (2009) Structure-  
20 function relationship of the chloroplastic glutaredoxin S12 with an atypical WCSYS  
21 active site. *J Biol Chem* 284: 9299-9310.
- 22 13. Kabsch W (2010) XDS. *Acta Crystallographica Section D: Biological Crystallography* 66:  
23 125-132.
- 24 14. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, et al. (2011) Overview of the  
25 CCP4 suite and current developments. *Acta Crystallographica Section D: Biological*  
26 *Crystallography* 67: 235-242.
- 27 15. Evans P (2006) Scaling and assessment of data quality. *Acta Crystallographica Section D:*  
28 *Biological Crystallography* 62: 72-82.
- 29 16. Vagin A, Teplyakov A (2010) Molecular replacement with MOLREP. *Acta*  
30 *Crystallographica Section D: Biological Crystallography* 66: 22-25.
- 31 17. Adams PD, Afonine PV, Bunkóczi G, Chen VB, Davis IW, et al. (2010) PHENIX: A  
32 comprehensive Python-based system for macromolecular structure solution. *Acta*  
33 *Crystallographica Section D: Biological Crystallography* 66: 213-221.
- 34 18. Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot.  
35 *Acta Crystallographica Section D-Biological Crystallography* 66: 486-501.
- 36 19. Chen VB, Arendall Iii WB, Headd JJ, Keedy DA, Immormino RM, et al. (2010)  
37 MolProbity: All-atom structure validation for macromolecular crystallography. *Acta*  
38 *Crystallographica Section D: Biological Crystallography* 66: 12-21.
- 39 20. Cummins I, Dixon DP, Freitag-Pohl S, Skipsey M, Edwards R (2011) Multiple roles for  
40 plant glutathione transferases in xenobiotic detoxification. *Drug Metab Rev* 43: 266-  
41 280.
- 42 21. Zhang J, Shibata A, Ito M, Shuto S, Ito Y, et al. (2011) Synthesis and Characterization of  
43 a Series of Highly Fluorogenic Substrates for Glutathione Transferases, a General  
44 Strategy. *Journal of the American Chemical Society* 133: 14109-14119.
- 45 22. Schonbrunn E, Eschenburg S, Luger K, Kabsch W, Amrhein N (2000) Structural basis for  
46 the interaction of the fluorescence probe 8-anilino-1-naphthalene sulfonate (ANS)  
47 with the antibiotic target MurA. *Proc Natl Acad Sci U S A* 97: 6345-6349.
- 48 23. Thuillier A, Ngadin AA, Thion C, Billard P, Jacquot JP, et al. (2011) Functional  
49 diversification of fungal glutathione transferases from the ure2p class. *Int J Evol Biol*  
50 2011: 938308.



- 1 24. Thuillier A, Roret T, Favier F, Gelhaye E, Jacquot JP, et al. (2013) Atypical features of a  
2 Ure2p Glutathione transferase from *Phanerochaete chrysosporium*. *FEBS Lett* 387:  
3 2125-2130.
- 4 25. Zhang ZR, Bai M, Wang XY, Zhou JM, Perrett S (2008) "Restoration" of glutathione  
5 transferase activity by single-site mutation of the yeast prion protein Ure2. *J Mol Biol*  
6 384: 641-651.
- 7 26. McTigue MA, Williams DR, Tainer JA (1995) Crystal structures of a schistosomal drug  
8 and vaccine target: glutathione S-transferase from *Schistosoma japonica* and its  
9 complex with the leading antischistosomal drug praziquantel. *J Mol Biol* 246: 21-27.
- 10 27. Ji X, von Rosenvinge EC, Johnson WW, Armstrong RN, Gilliland GL (1996) Location of  
11 a potential transport binding site in a sigma class glutathione transferase by x-ray  
12 crystallography. *Proc Natl Acad Sci U S A* 93: 8208-8213.
- 13 28. Reinemer P, Prade L, Hof P, Neufeind T, Huber R, et al. (1996) Three-dimensional  
14 structure of glutathione S-transferase from *Arabidopsis thaliana* at 2.2 Å resolution:  
15 structural characterization of herbicide-conjugating plant glutathione S-transferases  
16 and a novel active site architecture. *J Mol Biol* 255: 289-309.
- 17 29. Oakley AJ, Lo Bello M, Nuccetelli M, Mazzetti AP, Parker MW (1999) The ligandin  
18 (non-substrate) binding site of human Pi class glutathione transferase is located in the  
19 electrophile binding site (H-site). *J Mol Biol* 291: 913-926.
- 20 30. Cornwell WK, Cornelissen JHC, Allison SD, Bauhus J, Eggleton P, et al. (2009) Plant  
21 traits and wood fates across the globe: rotted, burned, or consumed? *Global Change*  
22 *Biology* 15: 2431-2449.
- 23 31. Hirose S, Tazaki M, Hiratsuka N, Yanai S, Kabumoto H, et al. (2011) Insight into  
24 functional diversity of cytochrome P450 in the white-rot basidiomycete *Phanerochaete*  
25 *chrysosporium*: involvement of versatile monooxygenase. *Biochem Biophys Res*  
26 *Commun* 407: 118-123.
- 27 32. Schuler MA (2011) P450s in plant-insect interactions. *Biochim Biophys Acta* 1814: 36-  
28 45.
- 29
- 30

1 **FOOTNOTES**

2

3 \* This work was supported by Convention industrielle de formation par la recherche (CIFRE)  
4 and the French National Research Agency through the (ANR-09-BLAN-0012) grant and the  
5 Laboratory of Excellence ARBRE (ANR-12- LABXARBRE-01) grant, the Ministère de  
6 l'Enseignement Supérieur, de la Recherche et de la Technologie, the Institut National de la  
7 Recherche Agronomique (INRA), and the Centre National de la Recherche Scientifique  
8 (CNRS).

9

10 The abbreviations used are: ANS, 8-anilino-1-naphtalene sulfonic acid; CDNB, 1-chloro-2,4-  
11 dinitrobenzene; CMFDA, 5-chloromethylfluorescein diacetate; DHA, dihydroascorbate;  
12 DMSO, dimethyl sulfoxide, GSH, reduced glutathione; GSSG, oxidized glutathione; GST,  
13 glutathione-S-transferase; GSTFuA, glutathione-S-transferase Fungal specific class A; HED,  
14 hydroxyethyl disulfide; JGI, Joint Genome Institute; MUA, 4-methylumbelliferyl acetate;  
15 NADPH, reduced nicotinamide adenine dinucleotide phosphate; phenethyl-ITC, phenethyl  
16 isothiocyanate; PAP-SG, S-(phenylacetophenone)-glutathione; PNP-butyrate, paranitrophenyl  
17 butyrate

## FIGURE LEGENDS

### FIGURE 1. Phylogenetic distribution of GSTFuA sequences in saprotrophic fungi and their bacterial Lig homologues from *Sphingobium* sp SYK-6.

Only bootstrap values over 70 are reported and indicated at nodes. Bold bars are compression of sequences from the same organism and boxes indicates putative catalytic residues in bold. PcGSTFuAs and CcGSTFuAs are highlighted in blue and purple respectively while the PcGSTFuA5 group of putative mitochondrial GSTFuAs is highlighted in red. LigEFG are highlighted in green. *Agaricus bisporus* (Agabi varbu) group1 contains: Prot ID 79084, 116016, 74698, 114110, 116585, 108910, 115923, 115917, 122742, 78843, 60492, 115916, 45066, 78836. *Coprinopsis cinerea* (Copci) group1 contains: Prot ID 1283, 2461, 6800, 6801, 4672, 6796, 6820, 6821. *Schizophyllum commune* (Schco) group 1 contains: Prot ID 111982, 59314, 85860. Schco group 2 contains: Prot ID 57691, 114676, 12387, 236992, 81614. *Pleurotus ostreatus* (Pleos) group 1 contains: Prot ID 1074017, 1074021, 1061513, 1103298. *Gloeophyllum trabeum* (Glotr) group 1 contains: Prot ID 109364, 74336, 61485, 76864, 76907. *Coniophora puteana* (Conpu) group 1 contains: Prot ID 98863, 169024, 159836, 113027, 132784. *Serpula lacrymans* (Serla) group 1 contains: Prot ID 94516, 162189, 185784, 112566, 186005, 107446, 780076, 106153, 111690, 185150, 185168, 115219, 188903. *Punctularia strigosozonata* (Punst) group 1 contains: Prot ID 97823, 114205, 50105, 120997, 34993. *Heterobasidion annosum* (Hetan) group 1 contains: 104697, 64692, 426787. *Stereum hirsutum* (Stehi) group 1 contains: Prot ID 161804, 130786, 97768. (Stehi) group 2 contains: Prot ID 125105, 103989, 161719, 171888, 142045. *Fomitopsis pinicola* (Fompi) group 1 contains: Prot ID 127980, 160680, 123507, 80060, 60354, 92199, 92198. *Trametes versicolor* (Trave) group 1 contains: Prot ID 47690, 24779, 64328, 148196. *Dichomitus squalens* (Dicsq) group 1 contains: prot ID 138582, 82843, 102613. *Ceriporiopsis subvermispora* (Cersu) group 1 contains: prot ID 163524, 73205, 83044. *Postia placenta* (Pospl) group 1 contains: Prot ID 87703, 50241, 88670, 50291, 43801, 95829, 91470, 108840.

FIGURE 2. Sequence alignment of GSTFuAs investigated in this study. Alignment was performed with ClustalW, and potential catalytic residues are highlighted in red.

FIGURE 3. Views of the superpositions of PcGSTFuA1-2 monomers (left) and PcGSTFuA1-3 monomers (right). PcGSTFuA1 cartoon is colored aquamarine while PcGSTFuA2-3 cartoons are colored salmon. The 'fold diversities' in PcGSTFuA1-3 isoforms are highlighted in yellow : (left) the disruption of the helix  $\alpha_4$  in PcGSTFuA2 ; (right) the loop  $\beta_1$ - $\alpha_1$ . Ligands are shown in stick mode. Glutathione and acetate molecules bound in PcGSTFuA1 are colored aquamarine. Citrate molecule (left) and formate molecules (right) molecules bound in PcGSTFuA2 and 3, respectively, are colored yellow.

FIGURE 4. Stereo views of the active sites of PcGSTFuA1 (A), PcGSTFuA2 (B) and PcGSTFuA3 (C). The carbon atoms of the ligands are colored orange : (A) a glutathione molecule and an acetate ion ; (B) a citrate molecule ; (C) four formate molecules. Most of the carbon atoms are colored green. The carbon atoms, which are colored magenta, highlight structural diversities in PcGSTFuA2 (B) and in PcGSTFuA3 (C) in relation to PcGSTFuA1. Other atoms are colored according to their types.

FIGURE 5. ANS binding and competition experiments onto PcGSTFuAs and CcGSTFuAs. (A)  $K_d$  values were determined by plotting concentration of ANS bound onto GSTFuAs against concentration of free ANS. Values given for glutathione, menadione-SG and phenylacetophenone-SG the values given are IC50 obtained by fitting data to equation 2 described in Material and Methods. (B) Number of binding sites for PcGSTFuAs and CcGSTFuA determined by fitting data to Equation 1.

FIGURE 6. Tryptophan based fluorescence experiments illustrating ANS binding site upon addition of substrate onto (A) PcGSTFuA2, (B) PcGSTFuA3, (C) PcGSTFuA4. Upon tryptophan excitation at 290 nm, FRET between 100/75/250  $\mu$ M ANS and 3  $\mu$ M PcGSTFuA2/3/4 is characterized

1 by the apparition of a signal at 475 nm (red) and is not altered in presence of 50  $\mu$ M CMFDA (green).  
2 Emission spectra of ANS alone and PcGSTFuAs alone are colored in purple and blue respectively.

3

4 **SUPPLEMENTAL FIGURE S1. Distance between equivalent  $C\alpha$  positions after superpositions**  
5 **of apo/holo PcGSTFuA1, PcGSTFuA2 (top) and of apo/holo PcGSTFuA1 and PcGSTFuA3**  
6 **(bottom).** Pair-wise RMS deviation of corresponding  $C\alpha$  atoms after superposition of apo  
7 PcGSTFuA1 and PcGSTFuA2 is 0.84 Å and after superposition of holo PcGSTFuA1 and  
8 PcGSTFuA2 is 0.61 Å. Pair-wise RMS deviation of corresponding  $C\alpha$  atoms after superposition of  
9 apo PcGSTFuA1 and PcGSTFuA3 is 1.04 Å and after superposition of holo PcGSTFuA1 and  
10 PcGSTFuA3 is 0.88 Å.

1 **TABLES**

2

3 **Table 1. Primers used in this study.**

4

5 The *Nco*I and *Bam*HI cloning restriction sites are underlined in the primers.

6

	Sequence
PeGSTFUA1 for	5' <u>CCCCC</u> CATGGCTCAGCCATCGTGTC 3'
PeGSTFUA1 rev	5' CCCC <u>GGATCC</u> CTATACATCAACCTGCTC 3'
PeGSTFUA2 for	5' <u>CCCC</u> CATGGCTTCCCAGCCATTGTCTTC 3'
PeGSTFUA2 rev	5' CCCC <u>GGATCC</u> TTAGTCATCTGCCCCGCTC 3'
PeGSTFUA3 for	5' <u>CCCC</u> CATGGCTTCCCTCGAGCCCATCATC 3'
PeGSTFUA3 rev	5' CCCC <u>GGATCC</u> CTAGACGTCTACGAACTC 3'
PeGSTFUA4 for	5' <u>CCCC</u> CATGGCCGATGTTATCACCTGTACG 3'
PeGSTFUA4 rev	5' CCCC <u>GGATCC</u> CTACAGGTCTACATGCGC 3'
CcGSTFUA2461 for	5' <u>CCCCC</u> CATGGCAATAACCTTCTACGACCTA 3'
CcGSTFUA2461 rev	5' <u>CCCCCGGATCC</u> TACAATTTACTTCCCGT 3'
CcGSTFUA6800 for	5' <u>CCCCCC</u> CATGGCAATTATACTATACGACCTC 3'
CcGSTFUA6800 rev	5' <u>CCCCCGGATCC</u> TTACACAACAGTAGTATA 3'
CcGSTFUA6801 for	5' <u>CCCCCC</u> CATGGCAATCACCTTCTACGATATC 3'
CcGSTFUA6801 rev	5' <u>CCCCCGGATCC</u> TTATGCGACAGTGTGATA 3'
CcGSTFUA6820 for	5' <u>CCCCCC</u> CATGGCAATCACCTTCTACGACTCC 3'
CcGSTFUA6820 rev	5' <u>CCCCCGGATCC</u> TAATGAACAGCCTGATA 3'

7

8

9

10

11

12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30

31

1 **Table 2. Statistics of X-ray diffraction data collection and model refinement**  
 2

	PcGSTFuA2	PcGSTFuA3
<b>Data collection</b>		
Beam line	PX1, SOLEIL	PX1, SOLEIL
Space group	$C222_1$	$P2_12_12_1$
Number of dimers in the ASU <sup>a</sup>	3	0.5
Cell dimensions a, b, c (Å)	100.41 202.47 192.76	87.00 87.00 225.70
Resolution (Å)	48.96-3.20 (3.37-3.20) <sup>b</sup>	45.16-2.10 (2.21-2.10)
Rmerge	0.12 (0.58)	0.064 (0.49)
Mean I/σ (I)	20.7 (4.9)	25.1 (4.0)
Completeness (%)	100.0 (100.0)	99.6 (97.4)
n observations	433,165 (54,032)	441,971 (42,972)
Redundancy	13.2 (11.4)	14.5 (10.2)
Wilson B factor (Å <sup>2</sup> )	78.2	35.9
<b>Phasing method <sup>c</sup></b>		
	<b>MR</b>	<b>MR</b>
<b>Refinement</b>		
Resolution (Å)	48.96-3.20 (3.30-3.20)	45.16-2.10 (2.17-2.10)
n reflections	32,677 (2,679)	30,346 (2,576)
Cutoff	$F > 1.34\sigma(F)$	$F > 1.35\sigma(F)$
Rall (%) <sup>d</sup>	19.2	18.8
Rfree (%) <sup>d</sup>	25.0 (29.6)	21.6 (31.1)
Average B-factor (Å <sup>2</sup> )		
Protein atoms	72.6	38.2
Ligand atoms	70.9	56.6
Solvent atoms	- <sup>e</sup>	42.2
Ramachandran statistics (%)		
Residues in preferred regions	96.3	98.0
Residues in allowed regions	3.7	1.6
Outlier residues	0.0	0.4
R.m.s. <sup>e</sup> deviations		
Bond length (Å)	0.017	0.007
Bond angle (°)	1.536	1.085

3  
 4  
 5 <sup>a</sup> ASU : Asymmetric unit.

6 <sup>b</sup> Values in parentheses are for highest resolution shell.

7 <sup>c</sup> MR : Molecular Replacement.

8 <sup>d</sup> Rall was determined from all the reflections (working set + test set), whereas Rfree corresponds to a subset of  
 9 reflections (test set).

10 <sup>e</sup> No water molecule was added in the PsGSTFuA2 3.2 Å model.

11 <sup>f</sup> R.m.s. : Root mean square.  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29

**Table 3. Kinetic parameters of PcGSTFuAs and CcGSTFuAs in enzymatic assays.**

The apparent  $K_m$  values for all compounds were determined using a concentration range of 0.1-10 mM in the presence of 5 mM GSH. The  $K_m$  value for GSH was determined with 1 mM Phenethyl-ITC for PcGSTFuA1, PcGSTFuA3, CcGSTFuA2461, CcGSTFuA6801 and 50  $\mu$ M CMFDA for CcGSTFuA6800 with a concentration range of 0.01 to 10 mM GSH. The apparent  $K_m$  and  $k_{cat}$  values were calculated by nonlinear regression using the Michaelis-Menten equation ( $r^2 > 0.99$ ). Data are represented as mean  $\pm$  S.D. ( $n \pm 3$ ). ND: Not Detected. The detection limit was estimated at 0.5 mUI.

Km ( $\mu$ M)	PcGSTFuA1	PcGSTFuA2	PcGSTFuA3	PcGSTFuA4	CcGSTFuA2461	CcGSTFuA6800	CcGSTFuA6801	CcGSTFuA6820
CDNB	ND	ND	ND	ND	4205 $\pm$ 557	ND	1172 $\pm$ 84.96	ND
Tertbutyl	ND	ND	ND	ND	ND	ND	ND	ND
Cumen	ND	ND	2037 $\pm$ 556	ND	ND	ND	ND	ND
H2O2	ND	ND	ND	ND	ND	ND	ND	ND
HED	ND	ND	ND	ND	ND	ND	ND	ND
DHA	ND	ND	ND	ND	ND	ND	ND	ND
ITC	119.0 $\pm$ 14.4	ND	53.16 $\pm$ 5.37	ND	191.2 $\pm$ 17.8	2431 $\pm$ 254	41.56 $\pm$ 3.73	ND
PNP-butyrate	774.3 $\pm$ 88.3	ND	612.6 $\pm$ 81.1	ND	ND	ND	177.7 $\pm$ 8.9	ND
CMFDA (-GSH)	3.6 $\pm$ 0.6	ND	ND	ND	ND	0.153 $\pm$ 0.023	0.945 $\pm$ 0.074	ND
CMFDA (+GSH)	8.0 $\pm$ 0.8	1.07 $\pm$ 0.16	5.1 $\pm$ 0.5	1.442 $\pm$ 0.07	ND	5.05 $\pm$ 1.29	3.13 $\pm$ 0.73	ND
GSH	216.4 $\pm$ 34.9	ND	21.1 $\pm$ 3.6	ND	435.0 $\pm$ 29.4	474.9 $\pm$ 43.8	167.3 $\pm$ 10.5	ND

kcat ( $\text{min}^{-1}$ )	PcGSTFuA1	PcGSTFuA2	PcGSTFuA3	PcGSTFuA4	CcGSTFuA2461	CcGSTFuA6800	CcGSTFuA6801	CcGSTFuA6820
CDNB	ND	ND	ND	ND	443.2 $\pm$ 39.6	ND	251.9 $\pm$ 8.1	ND
Tertbutyl	ND	ND	ND	ND	ND	ND	ND	ND
Cumen	ND	ND	76.70 $\pm$ 9.78	ND	ND	ND	ND	ND
H2O2	ND	ND	ND	ND	ND	ND	ND	ND
HED	ND	ND	ND	ND	ND	ND	ND	ND
DHA	ND	ND	ND	ND	ND	ND	ND	ND
ITC	510.9 $\pm$ 16.0	ND	130.9 $\pm$ 2.9	ND	368.1 $\pm$ 10.3	469.2 $\pm$ 29.9	1065.0 $\pm$ 20.2	ND
PNP-butyrate	165.0 $\pm$ 5.6	ND	246.5 $\pm$ 10.2	ND	ND	ND	115.7 $\pm$ 1.4	ND
CMFDA (-GSH)	0.006 $\pm$ 0.0002	ND	ND	ND	ND	0.123 $\pm$ 0.003	0.343 $\pm$ 0.005	ND
CMFDA (+GSH)	0.17 $\pm$ 0.04	0.033 $\pm$ 0.001	0.65 $\pm$ 0.02	0.002 $\pm$ 0.0001	ND	0.39 $\pm$ 0.05	0.85 $\pm$ 0.08	ND
GSH	662.0 $\pm$ 14.2	ND	52.33 $\pm$ 1.68	ND	46.01 $\pm$ 2.45	37.15 $\pm$ 1.6	31.95 $\pm$ 0.53	ND

kcat/ $K_m$ ( $\text{min}^{-1}/\mu\text{M}^{-1}$ )	PcGSTFuA1	PcGSTFuA2	PcGSTFuA3	PcGSTFuA4	CcGSTFuA2461	CcGSTFuA6800	CcGSTFuA6801	CcGSTFuA6820
CDNB	ND	ND	ND	ND	0.11 $\pm$ 0.07	ND	0.21 $\pm$ 0.09	ND
Tertbutyl	ND	ND	ND	ND	ND	ND	ND	ND
Cumen	ND	ND	0.037 $\pm$ 0.017	ND	ND	ND	ND	ND
H2O2	ND	ND	ND	ND	ND	ND	ND	ND
HED	ND	ND	ND	ND	ND	ND	ND	ND
DHA	ND	ND	ND	ND	ND	ND	ND	ND
ITC	4.3 $\pm$ 0.3	ND	2.46 $\pm$ 0.54	ND	1.92 $\pm$ 0.58	0.19 $\pm$ 0.11	25.62 $\pm$ 5.41	ND
PNP-butyrate	0.21 $\pm$ 0.01	ND	0.40 $\pm$ 0.12	ND	ND	ND	0.65 $\pm$ 0.16	ND
CMFDA (-GSH)	0.0017 $\pm$ 0.0002	ND	ND	ND	ND	0.82 $\pm$ 0.13	0.36 $\pm$ 0.06	ND
CMFDA (+GSH)	0.021 $\pm$ 0.002	0.031 $\pm$ 0.006	0.13 $\pm$ 0.04	1.39*10 <sup>-3</sup> $\pm$ 0.14*10 <sup>-3</sup>	ND	0.077 $\pm$ 0.038	0.27 $\pm$ 0.11	ND
GSH	3.1 $\pm$ 0.4	ND	2.48 $\pm$ 0.47	ND	0.11 $\pm$ 0.08	0.078 $\pm$ 0.036	0.19 $\pm$ 0.05	ND

10

1 **Table 4. Kinetic parameters of PcGSTFuAs and CcGSTFuAs towards MUA in absence**  
 2 **or presence of GSH.**

3  
 4 The apparent  $K_m$  values were determined using a concentration range of 100-800  $\mu\text{M}$  in  
 5 absence (-GSH) or presence of 5 mM GSH (+ GSH). The apparent  $K_m$  and  $k_{\text{cat}}$  values were  
 6 calculated by nonlinear regression using the Michaelis-Menten equation ( $r^2 > 0.99$ ). Data are  
 7 represented as mean  $\pm$  S.D. ( $n \pm 3$ ). ND: Not Detected. The detection limit was estimated at  
 8 0.5 mUI.  
 9

$K_m$ ( $\mu\text{M}$ )	PcGSTFUA1	PcGSTFUA2	PcGSTFUA3	PcGSTFUA4
MUA (-GSH)	$262.8 \pm 33.7$	$23.5 \pm 6.2$	$33.89 \pm 4.78$	$51.56 \pm 7.09$
MUA (+GSH)	$206.7 \pm 23.3$	$216.1 \pm 45.2$	$372.1 \pm 53.9$	N.D

$k_{\text{cat}}$ ( $\text{min}^{-1}$ )	PcGSTFUA1	PcGSTFUA2	PcGSTFUA3	PcGSTFUA4
MUA (-GSH)	$0.095 \pm 0.004$	$0.302 \pm 0.006$	$0.182 \pm 0.003$	$0.208 \pm 0.004$
MUA (+GSH)	$57.83 \pm 2.90$	$1.27 \pm 0.14$	$29.65 \pm 2.36$	N.D

$k_{\text{cat}}/K_m$ ( $\text{min}^{-1}/\mu\text{M}^{-1}$ )	PcGSTFUA1	PcGSTFUA2	PcGSTFUA3	PcGSTFUA4
MUA (-GSH)	$0.36 \cdot 10^{-3} \pm 0.12 \cdot 10^{-3}$	$0.0128 \pm 0.0001$	$5.37 \cdot 10^{-3} \pm 0.62 \cdot 10^{-3}$	$4.03 \cdot 10^{-3} \pm 0.56 \cdot 10^{-3}$
MUA (+GSH)	$0.28 \pm 0.12$	$5.87 \cdot 10^{-3} \pm 3.1 \cdot 10^{-3}$	$0.079 \pm 0.044$	N.D

$K_m$ ( $\mu\text{M}$ )	CcGSTFuA2461	CcGSTFuA6800	CcGSTFuA6801	CcGSTFuA6820
MUA (-GSH)	$38.7 \pm 7.1$	$25.39 \pm 4.37$	$182.3 \pm 29.4$	$16.9 \pm 2.1$
MUA (+GSH)	N.D	$111.7 \pm 11.1$	$189.7 \pm 7.3$	N.D

$k_{\text{cat}}$ ( $\text{min}^{-1}$ )	CcGSTFuA2461	CcGSTFuA6800	CcGSTFuA6801	CcGSTFuA6820
MUA (-GSH)	$0.026 \pm 0.001$	$0.065 \pm 0.002$	$0.070 \pm 0.004$	$0.078 \pm 0.002$
MUA (+GSH)	N.D	$5.10 \pm 0.18$	$7.79 \pm 0.13$	N.D

$k_{\text{cat}}/K_m$ ( $\text{min}^{-1}/\mu\text{M}^{-1}$ )	CcGSTFuA2461	CcGSTFuA6800	CcGSTFuA6801	CcGSTFuA6820
MUA (-GSH)	$0.67 \cdot 10^{-3} \pm 0.14 \cdot 10^{-3}$	$2.56 \cdot 10^{-3} \pm 0.46 \cdot 10^{-3}$	$0.38 \cdot 10^{-3} \pm 0.13 \cdot 10^{-3}$	$4.61 \cdot 10^{-3} \pm 0.9 \cdot 10^{-3}$
MUA (+GSH)	N.D	$0.046 \pm 0.016$	$0.041 \pm 0.017$	N.D

10



**Table 5. Competition experiments between ANS or MUA and wood compounds.**

For ANS competition, the values given are IC<sub>50</sub> obtained by fitting data to equation 2 described in Material and Methods.

For MUA competition, the values given are K<sub>i</sub> obtained by fitting data to noncompetitive or competitive inhibition of esterase activity in the presence of 0 to 10 mM inhibitor.

The apparent K<sub>i</sub> values were calculated by nonlinear regression using the competitive or noncompetitive inhibitions equations described in Material and Methods ( $r^2 > 0.99$ ). Data are represented as mean  $\pm$  S.D. ( $n \pm 3$ ). ND: Not determined since no competition was observed with GSH for ANS and MUA.

	PcGSTFuA1	PcGSTFuA2	PcGSTFuA3	PcGSTFuA4
Fluorescent probe	ANS	ANS	ANS	ANS
Coniferaldehyde	62.3 $\pm$ 1.2 $\mu$ M	115.6 $\pm$ 44.1 $\mu$ M	68.4 $\pm$ 25.3 $\mu$ M	146.6 $\pm$ 5.1 $\mu$ M
Vanillin	217.8 $\pm$ 8.5 $\mu$ M	84.7 $\pm$ 5.2 $\mu$ M	134.0 $\pm$ 7.6 $\mu$ M	308.9 $\pm$ 29.3 $\mu$ M
Methoxybenzophenone	-	-	-	-
4-chloro-3-nitrobenzoic acid	1.5 $\pm$ 0.1 mM	1.3 $\pm$ 0.2 mM	234.7 $\pm$ 56.7 $\mu$ M	-
4'-hydroxyacetophenone	2.2 $\pm$ 0.4 mM	-	-	-
Gallic acid	-	-	-	-
Vanillic acid	-	-	-	-
Epicatechin	1.6 $\pm$ 0.1 mM	-	-	-
Syringaldehyde	62.5 $\pm$ 1.6 $\mu$ M	89.3 $\pm$ 2.8 $\mu$ M	37.2 $\pm$ 3.8 $\mu$ M	198.5 $\pm$ 25.2 $\mu$ M
Catechin hydrate	575.3 $\pm$ 98.7 $\mu$ M	-	-	-
Metoxycinnamic	-	-	-	-

	CcGSTFuA2461	CcGSTFuA6800	CcGSTFu6801	CcGSTFuA6820
Fluorescent probe	MUA	ANS/MUA	ANS	MUA
Coniferaldehyde	-	ND	327.4 $\pm$ 50.6 $\mu$ M	768.8 $\pm$ 98.4 $\mu$ M (non-competitive)
Vanillin	-	ND	423.6 $\pm$ 107.3 $\mu$ M	-
Methoxybenzophenone	-	ND	-	-
4-chloro-3-nitrobenzoic acid	287.4 $\pm$ 29.7 $\mu$ M (competitive)	ND	2.3 $\pm$ 1.3 mM	1.3 $\pm$ 0.03 mM (non-competitive)
4'-hydroxyacetophenone	352.1 $\pm$ 25.1 $\mu$ M (competitive)	ND	2.9 $\pm$ 1 mM	1.2 $\pm$ 0.03 mM (non-competitive)
Gallic acid	113.9 $\pm$ 8.9 $\mu$ M (competitive)	ND	-	566.4 $\pm$ 16.5 $\mu$ M (non-competitive)
Vanillic acid	1 $\pm$ 0.1 mM (competitive)	ND	-	-
Epicatechin	-	ND	3.9 $\pm$ 2.3 mM	1.4 $\pm$ 0.04 mM (non-competitive)
Syringaldehyde	-	ND	275.8 $\pm$ 97.8 $\mu$ M	594.2 $\pm$ 37.6 $\mu$ M (non-competitive)
Catechin hydrate	-	ND	3.7 $\pm$ 1.7 mM	564.2 $\pm$ 12 $\mu$ M (non-compet)
etoxycinnamic	-	ND	-	1.5 $\pm$ 0.04 mM (non-competitive)

12

Figure 1  
 Click here to download high resolution image

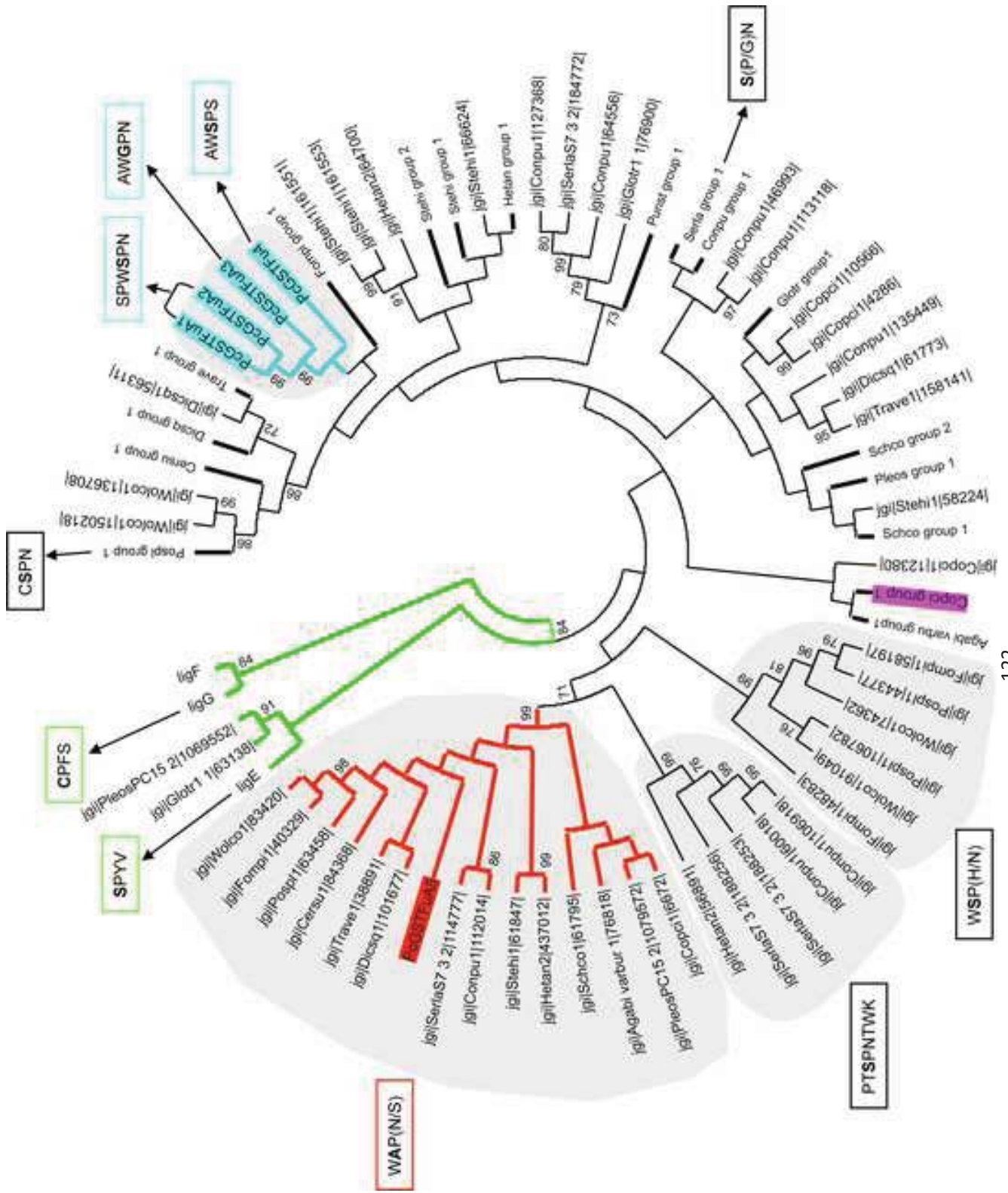


Figure2  
 Click here to download high resolution image

CcGSTFuA682D	- - - - M I T L Y D S - R L T H L Q A - - A D P F T W K A R F V L N Y K N L P Y K T E W T P U H D I A S V Y H K H N L Q	53
CcGSTFuA6801	- - - - M I T F Y D I P S T T S L N S - - W C P P D G W K T R Y T L N F P K G L P P Y K T E W V E P P D I A A L Y Q K N N V P	54
CcGSTFuA6800	- - - - M I I L Y D L P S N L P T Q A - - W C P G D T W K A R F C L N Y K K G L P P Y K T E W V E P P D V P S I Y K K H G I S	54
CcGSTFuA2461	- - - - M I T F Y D I P S I L P E N D - - W C P P D T W K T R L C L O N Y K K G L D Y K T E W V E P P D I A K V C Q C L G I A	54
PcGSTFuA1	- M A Q P I V F Y D I P S N E R I K H S P W E P P D T W K T R Y A L N Y K K G L K Y K T E W V E P P D I A G V V Q C L G G K	59
PcGSTFuA2	- M S Q P I V F Y D I P S N D T L K Q S P W E P P D T W K T R Y A L N Y K K G L K Y K T E W V E P P D I E D V V K C L G G K	59
PcGSTFuA3	M S L E P I I F Y D I P A N E - P R Q M A W C P P D T W K T R Y V L N F K K G L K Y R T E W V E P P D I E A V C K Q I G A P	59
PcGSTFuA4	- M A D V L T L Y D I P G K S - D Q C K A W E P L T W K A R F A L N I K G I P Y K T E W V E P P D I A T V Y E K H G I T	58
CcGSTFuA682D	P V C S R T D P S T G K E E P Y Y T I P V I F D S T G - - - K A I A D S L E I A K Y L D E T Y P D T P K	103
CcGSTFuA6801	A P A K Q S D G S - - - P Y Y S L P L I H D S K T N - - - T F V A D S L E I A K Y L D K T Y P D T P K	99
CcGSTFuA6800	P V C S F G E G G - - - K Y Y S L P V I H D T S T G - - - T T I A E S L E I A K Y L D A T Y P D A P S	99
CcGSTFuA2461	A T A T W R D G S - - - P Y F S V P A I V D V Q E S E S D T G F T M K R T A L A E S F E I A K Y L D A A Y P D T K R	109
PcGSTFuA1	P T E K T P D G R - - - D H Y T L P V I Y D P N T K - - - - K V V E D S A A I A K Y L D E T Y P D T P K	104
PcGSTFuA2	P T G K K P D G R - - - D H Y T V P V I Y D P N T K - - - - T V V E D G I K I A K Y L D D A Y P D T P R	104
PcGSTFuA3	A T E K K P D G R - - - D H Y T L P V I Q D P N T K - - - - A V V A D S D A I A K Y L E S T Y P D T P R	104
PcGSTFuA4	A K Y T G D D G L - - - P L Y T L P V I Y D P A T R - - - - R T V A E S S D I A K Y L D R T Y P N T	103
CcGSTFuA682D	V V P E T G E D G V D A K Q Q V D T F M G V M Y G A L G P L V L P V F Q Q S Y Q H V E E S K P F F A E A R T K D F	163
CcGSTFuA6801	V L P D G E E - - - G L A K T T Q F L G E L G G T M F P L F P I T L H S V G K I L N E V S K D H F K V A R A K N L V S	155
CcGSTFuA6800	V I P E G V E G S A E A L S D Q A E V F R - - - K L V D Y P L V L P I T L H S V G K I L N P P S K A H V V Q T R A Q I L G L	159
CcGSTFuA2461	L F P A G T D - - - A F Q A A F L D F A W P V L G - F P V F M L V I L D T A N S L P R S H D Y F R S T R E Q K F G K	166
PcGSTFuA1	L F P A G T D - - - A F Q A A F D D F V W S V T L A F P L S L L L D V S N S L P P R S S A Y F R A T R E Q Q F G K	160
PcGSTFuA2	L F P A G T D - - - A F Q A A F Y Q L A R P S V L - M P I F N I V V A R V W K L L R P R S Q E V F R A T R E Q M L G K	159
PcGSTFuA3	L F P E G T R - - - A F Q H A F Y Q L A R P S V L - M P I F N I V V A R V W K L L R P R S Q E V F R A T R E Q M L G K	159
PcGSTFuA4	L F P T G T A - - - A L H E A F A G A Y T A V H K - - - A L Y N L V C C A A C Y A L T E R S Q A F Y R T T R E A R L G K	157
CcGSTFuA682D	W F G N P K S L L D L L P L K X E D E D K G W E A A E K T F A G L D A R V P - - G R G K E V G W Y L G G E S P T F V D	221
CcGSTFuA6801	V F N N I K S I D E I Q L S E K E V E E R - W E Q V R A G F Q Q G L D S K F G - - G E G K E L A W F A G - D K I T F A D F	211
CcGSTFuA6800	F G - - - V G T I E E F K F S E E D V N K G W D H L K R S I F D A L D K K F G - - G E D G K F G W Y L G - A R I S F A D F	213
CcGSTFuA2461	K E - - - S L D E V A V K T D E E R E K I W G D L R E G L K A I D R Y Y K - - E T D E A G P W M L G - K T V S F A D F	219
PcGSTFuA1	K L E - - - - - - - - - E L A T E E E W A K V E A G L A X L K G Y L D A N G K G N D L L M G A Q G G I T Y S D I	207
PcGSTFuA2	R L E - - - - - - - - - E Q G G E E R W Q Q L E A G L G K F K G Y L E R N G A G N D L L M G T Q G G I T Y S D V	208
PcGSTFuA3	K L E - - - - - - - - - E I G S E O D W N A L E S G L A R I K S S L E A N G A G K D L L M G D R - - V T F A D L	205
PcGSTFuA4	A L E - - - - - - - - - D V C T A Q D W V D A E A A L A G W L D A N G A G A D A E W V M G A G T L C F A D V	205
CcGSTFuA682D	V L G G F L I Y L R S I Y G E E S E G W K R R I V G W N E G K W G A F L D K L S A Y Q A V H -	266
CcGSTFuA6801	A L A G M L L W I K L V W G E E S A E W K R V S G W N N G K W A T M L K K L E A Y H T V A -	256
CcGSTFuA6800	V L G G S L F L K T A W G E D S E E W A K V K N W H N G K W A A L N A L D S Y T V V -	258
CcGSTFuA2461	V V A F L V A L K R V F G E E E W A K V K S L W D D G R W G H L L Q K A T A E L T G S K L	265
PcGSTFuA1	Q I A S F F V W A K I I W G E S E K W K R L I S L H D G K W A Q F Y A Q T T K F E Q V D V	253
PcGSTFuA2	Q I A S L F V W A K V V W G E G S E K W K R L M G F H G G K W A Q F C A Q F A E Y E R A D D	254
PcGSTFuA3	Q L A S L F I W L R V S A G E E S E D W K R F L S L H E G K W A K F M Q Q F A A Y E F V D V	251
PcGSTFuA4	R V V S G L W A R T T L G E A S A E W K R R V C G W H G G R W A R V V A N F G P Y A H V D -	250

Figure3  
[Click here to download high resolution image](#)

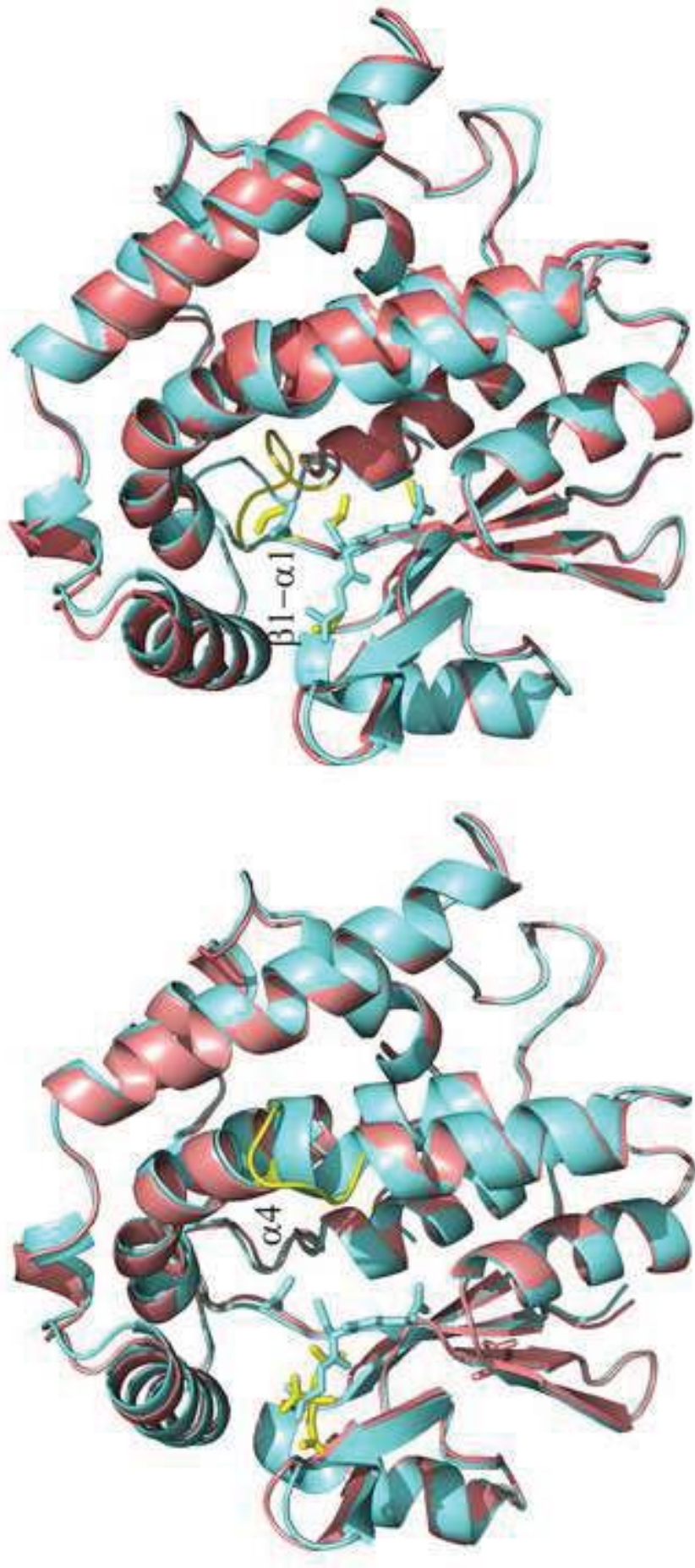


Figure4  
[Click here to download high resolution image](#)

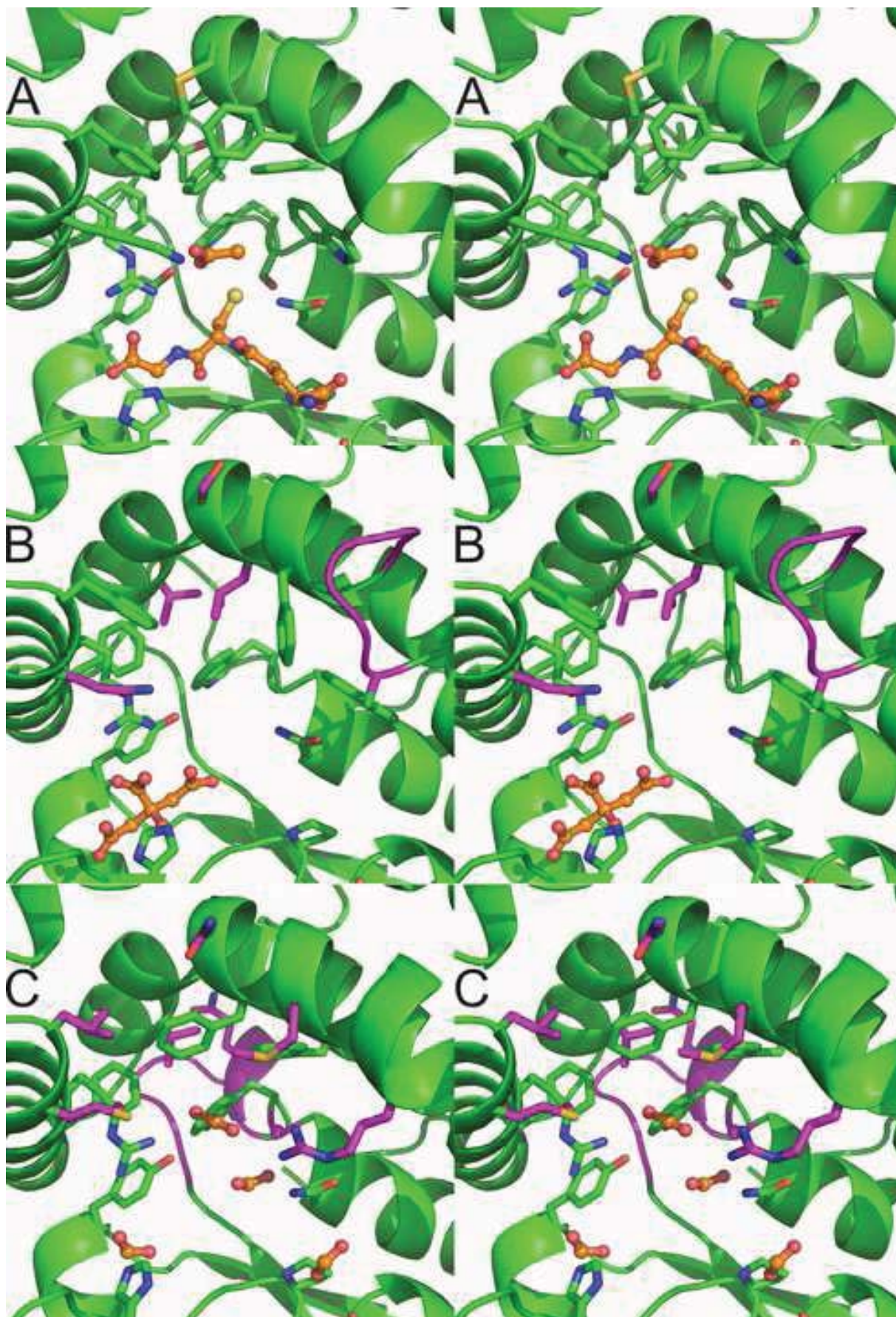


Figure5  
[Click here to download high resolution image](#)

	PCGSTFuA1	PCGSTFuA2	PCGSTFuA3	PCGSTFuA4	PCGSTFuA2461	CCGSTFuA6800	CCGSTFuA6801	CCGSTFuA6820
Kd at ANS saturation ( $\mu\text{M}$ )	10.6 $\pm$ 0.5	7.23 $\pm$ 0.58	13.75 $\pm$ 1.56	144.5 $\pm$ 36.2	32.1 $\pm$ 3.1	17.03 $\pm$ 2.24	9.19 $\pm$ 0.31	5.24 $\pm$ 0.33
ANS saturation ( $\mu\text{M}$ )	50	100	75	250	100	100	75	75
Glutathione ( $\mu\text{M}$ )	272.2 $\pm$ 31.5	25.7 $\pm$ 3.6	62.1 $\pm$ 5.8	39.2 $\pm$ 4.8	-	-	199.9 $\pm$ 45.6	-
Menadione	-	-	-	-	-	-	-	-
IC 50 Menadione-SG ( $\mu\text{M}$ )	87.3 $\pm$ 38.4	150.7 $\pm$ 35.3	147.7 $\pm$ 49.8	52.72 $\pm$ 19.2	-	-	-	-
2-Bromo-4'-phenylacetophenone	-	-	-	-	-	-	-	-
Phenylacetophenone-SG	99.7 $\pm$ 46.5	286.4 $\pm$ 29.3	27.7 $\pm$ 3.6	142.2 $\pm$ 41.6	-	-	-	-

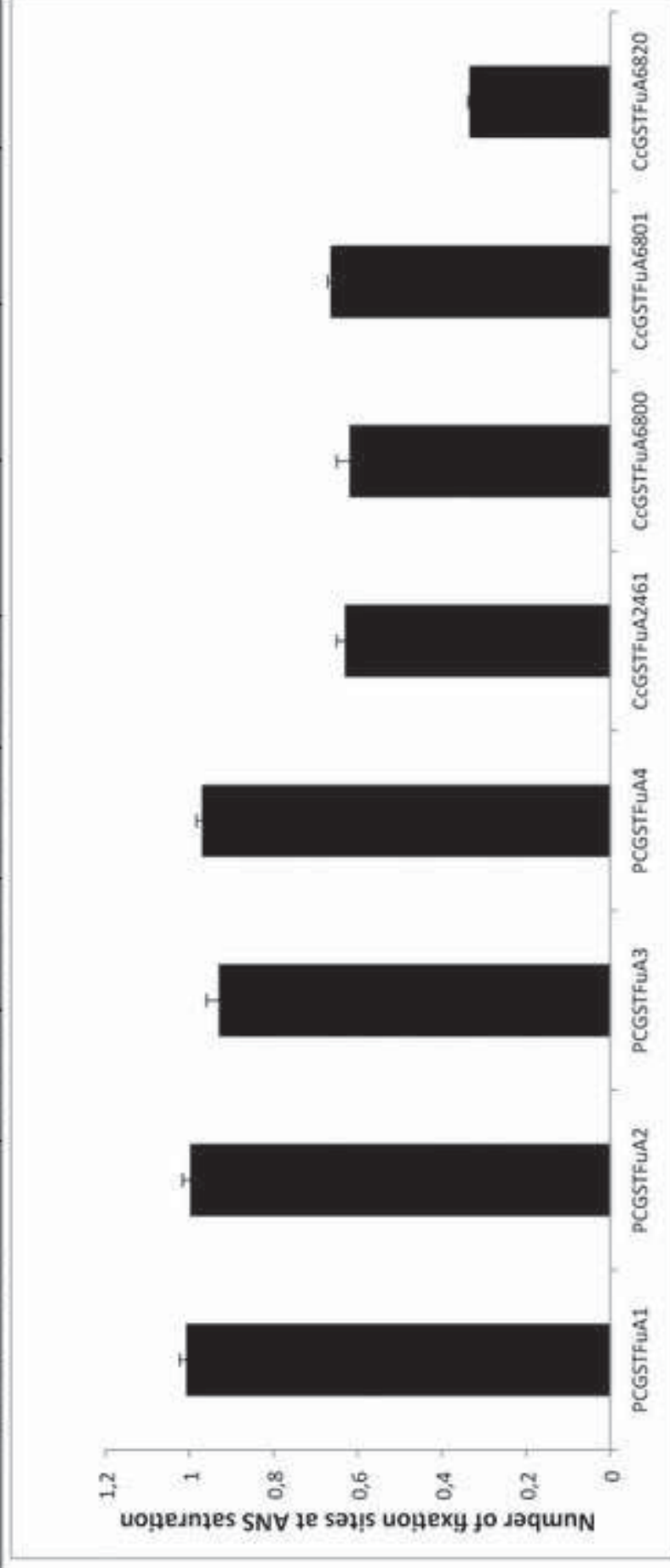
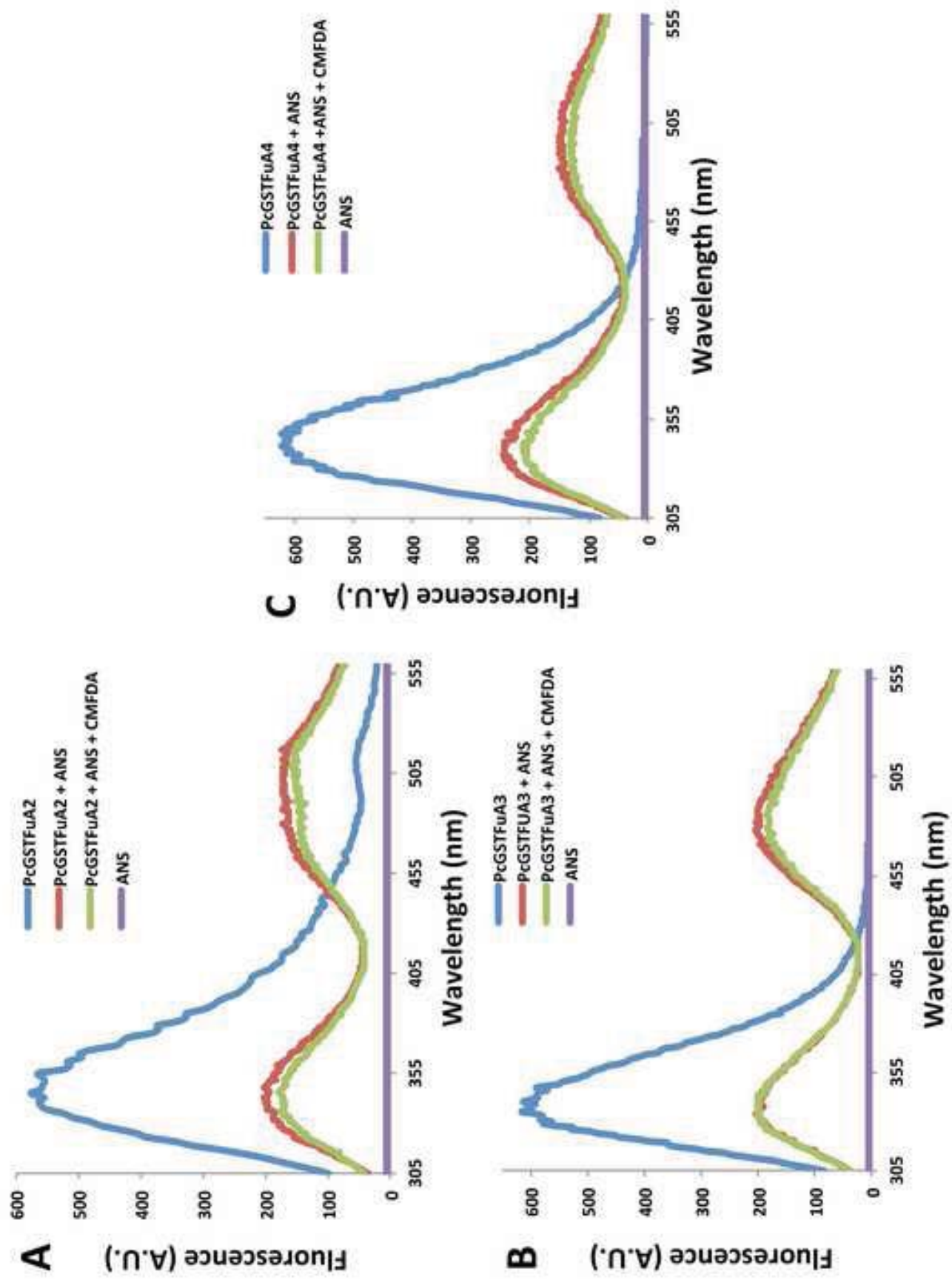
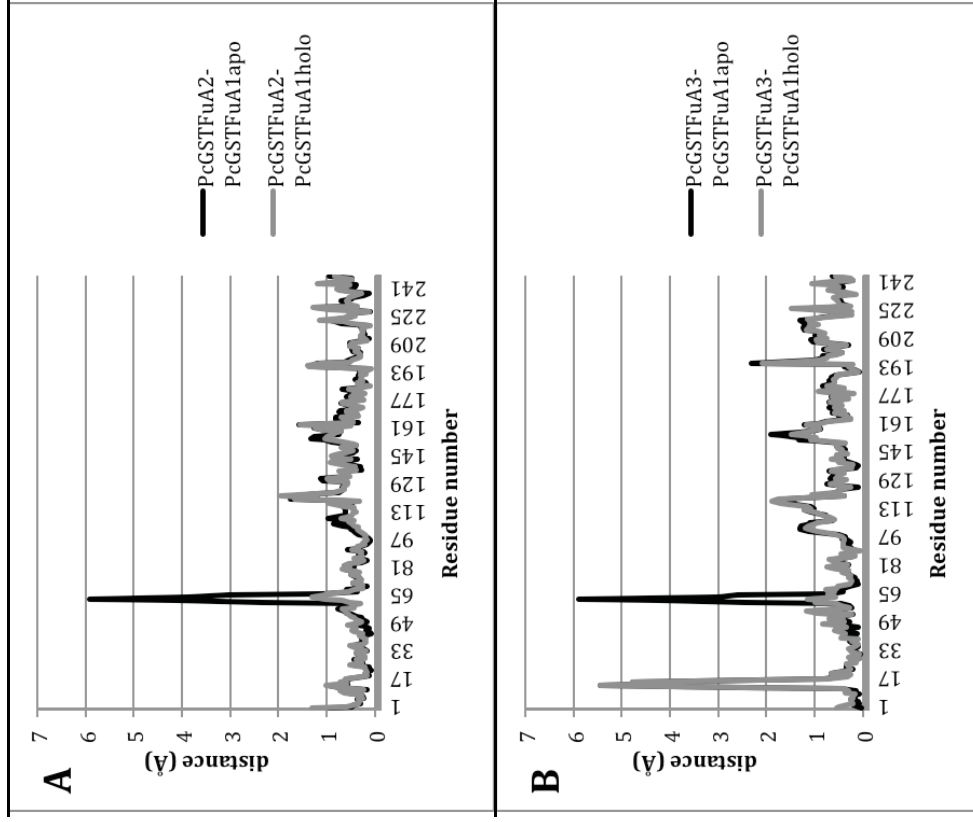


Figure6  
[Click here to download high resolution image](#)



Supplemental Figure S1





1 **Supplemental Table S1. Distribution of GSTFuAs.**

2

3 Sequences were obtained from the Joint Genome Institute with Blastp using all *P.*  
 4 *chrysosporium* PcGSTFuA sequences as template in all available genomes of  
 5 chytridiomycotina, mucoromycotina, saccharomycotina, pezizomycotina, puccinomycotina,  
 6 ustilaginomycotina and agaricomycotina.

7

		gene models	number of GSTFuA sequences
<b>Agaricomycotina</b>	<i>Agaricus bisporus</i>	10438	15
	<i>Ceriporiopsis subvermispora</i>	12125	4
	<i>Coniophora puteana</i>	13761	14
	<i>Coprinopsis cinerea</i>	13342	14
	<i>Cryptococcus neoformans</i>	6967	0
	<i>Dacryopinax</i> sp	10242	2
	<i>Dichomitus squalens</i>	12290	6
	<i>Fomitopsis pinicola</i>	14724	10
	<i>Gloeophyllum trabeum</i>	11846	7
	<i>Heterobasidion annosum</i>	13405	7
	<i>Laccaria bicolor</i>	23132	3
	<i>Paxillus involutus</i>	17968	6
	<i>Phanerochaete chrysosporium</i>	10048	5
	<i>Pleurotus ostreatus</i> Pc 15	12330	6
	<i>Postia placenta</i>	17173	11
	<i>Punctularia strigosozonata</i>	11538	5
	<i>Schizophyllum commune</i>	13210	7
	<i>Serpula lacrymans</i> var <i>lacrymans</i> S7,9	12789	20
	<i>Serpula lacrymans</i> var <i>shastensis</i>	13805	7
	<i>Stereum hirsutum</i>	14072	12
<i>Trametes versicolor</i>	14296	6	
<i>Wolfiporia cocos</i>	12746	5	
<b>Ustilaginomycotina</b>	<i>Ustilago maydis</i>	6902	1
<b>Puccinomycotina</b>	<i>Puccinia graminis</i>	20567	0
	<i>Melampsora laricis</i>	16694	0
	<i>Sporobolomyces roseus</i>	5536	1
<b>Pezizomycotina</b>	<i>Aspergillus carbonarius</i>	11624	3
	<i>Magnaporthe grisea</i>	11054	2
	<i>Trichoderma atroviride</i>	11863	3
	<i>Trichoderma virens</i>	12427	2
	<i>Neurospora discreta</i>	10620	1
	<i>Neurospora tetrasperma</i>	10380	0
	<i>Neurospora crassa</i>	9733	1
	<i>Chaetomium globosum</i>	11124	1
	<i>Cochliobus heterostrophus</i>	13336	2
	<i>Aspergillus clavatus</i>	9121	2
	<i>Aspergillus fumigatus</i>	9887	2
	<i>Aspergillus nidulans</i>	10701	3
	<i>Fusarium verticillioides</i>	14179	3
	<i>Fusarium graminearum</i>	13332	3
	<i>Mycosphaerella fijensis</i>	10327	3
	<i>Sclerotinia sclerotinum</i>	14522	1
	<i>Aspergillus terreus</i>	10406	4
<i>Botrytis cinera</i>	16448	3	
<i>Mycosphaerella graminicola</i>	10952	3	
<i>Cryphonectria parasitica</i>	11609	3	
<i>Trichoderma reesei</i>	9129	4	
<b>Saccharomycotina</b>	<i>Spathaspora panalidarum</i>	5983	0
	<i>Pichia stipitis</i>	5841	0
	<i>Candida tenuis</i>	5533	0
	<i>Saccharomyces cerevisiae</i>	5695	0
<b>Mucoromycotina</b>	<i>Mucor circinelloides</i>	11719	1
	<i>Phycomyces blakesleeanae</i>	14792	9
	<i>Rhizopus oryzae</i>	17459	2
<b>Chytridiomycota</b>	<i>Batrachochytrium dendrobatidis</i>	8732	0

8  
9

## **B.IV Etudes structurales des Lig**

L'étude phylogénétique menée par Morel et collaborateurs (2009) à partir des génomes de différents basidiomycètes, en se focalisant plus précisément sur *P. chrysosporium*, a révélé que les champignons saprophytes possèdent plusieurs GST formant une classe indépendante et présentant des homologies avec l'éthérase LigE de *Sphingobium* sp. SYK-6. Cette enzyme LigE est l'un des maillons d'une cascade de réactions permettant à la bactérie la dégradation de fragments de lignine (Masai *et al.*, 2007, et Illustration 18), dont les produits sont utilisés dans le cycle de l'acide tricarboxylique. L'observation chez *P. chrysosporium* de GST homologues à des protéines Lig dégradant la lignine semble en accord avec le caractère saprophyte du champignon et nous a conduits à étudier tant cette classe de GST chez *P. chrysosporium* (initialement baptisée PcGTE, puis renommée PcGSTFuA, voir paragraphe B.III) que les protéines de type GST du système Lig chez la bactérie *Sphingobium* sp. SYK-6.

Les GST du système Lig de *Sphingobium* sp. SYK-6 sont respectivement les Lig E, F et G appartenant au cluster de gènes *ligDFEG* (Masai *et al.*, 2003). Dans la cascade de réactions étudiée par Masai et collaborateurs sur un modèle de lignine dimérique, les protéines LigE et LigF catalysent la première étape de la coupure du lien éther de l' $\alpha$ -(2-méthoxyphénoxy)- $\beta$ -hydroxypropiovanillone (MPHPV) pour produire la  $\beta$ -hydroxypropiovanillone glutathionylée (GS-HPV) et le gaïacol (voir Illustration 19). Masai et collaborateurs remarquent cependant que l'attaque nucléophile du glutathion est sélectivement dirigée vers l'une ou l'autre des deux configurations possibles de l'atome de carbone en position  $\beta$  du MPHPV, selon que LigE ou LigF est utilisée. Ainsi, LigF semble stéréospécifique de l'énantiomère portant le carbone  $\beta$  de configuration *S*, et LigE agit sur la forme *R* (Sato *et al.*, 2009). La GS-HPV issue de cette première étape de réaction subit ensuite une déglutathionylation. A nouveau, la GST impliquée est énantiosélective. La protéine LigG a ainsi été démontrée stéréospécifique du produit de la réaction catalysée par LigF. Par contre à ce jour, aucune protéine n'a encore été identifiée chez *Sphingobium* sp. SYK-6 qui prenne en charge le produit de la première étape catalysée par LigE.

De nombreuses questions restent en suspens quant aux propriétés biochimiques de ces Lig et aucune étude structurale n'en a été réalisée avant les présents travaux. Les études menées sur les PcGSTFuA supposées avoir des propriétés éthérases appellent donc logiquement l'étude conjointe de ces Lig de type GST, avec pour ambition notable l'analyse structurale de complexes enzyme/substrat : identifier les spécificités de site actif d'une éthérase avérée intervenant dans la dégradation de composés ligniniques revêt une importance tant sur le plan théorique qu'appliqué.

## **B.IV.1 LigE**

### **B.IV.1.a Introduction**

LigE (281 résidus) est la première protéine du système Lig isolée par Masai et collaborateurs (1991). Bien que catalysant une réaction similaire à LigF, elle présente moins de 20 % d'identité de séquence avec cette enzyme. Une analyse des GST bactériennes réalisée en 2009 par Allocati et collaborateurs montre que LigE s'apparente aux GST de la classe Zêta. Selon cette observation, LigE présenterait un résidu catalytique sérine. LigE ne montre qu'une très faible activité glutathion transférase vis-à-vis du substrat classique CDNB. Son activité éthérase vis-à-vis du (-)-( $\beta$ R)-MPPHV conduit à l'obtention d'uniquement l'un des deux diastéréoisomères du GS-HPV. Le glutathion ne peut en être éliminé par LigG. Aucune voie enzymatique permettant cette élimination n'a à ce jour été identifiée chez *Sphingobium* sp. SYK-6. D'autre part, il existe probablement une autre voie de coupure du lien  $\beta$ -aryl éther de la forme R de MPPHV (elle aussi nécessitant le glutathion), puisque l'inactivation du gène *ligE* n'affecte pratiquement pas la capacité de la bactérie à dégrader le MPPHV.

La protéine surexprimée et purifiée par l'équipe d'E. Gelhaye a été cristallisée et une étude cristallographique préliminaire en a été réalisée, malheureusement elle n'a pu aboutir compte tenu de la faible qualité des cristaux de la forme sélénée de la protéine, et du fait de l'échec des méthodes du remplacement moléculaire sur les données de diffraction de la forme native.

### **B.IV.1.b Matériel et Méthodes**

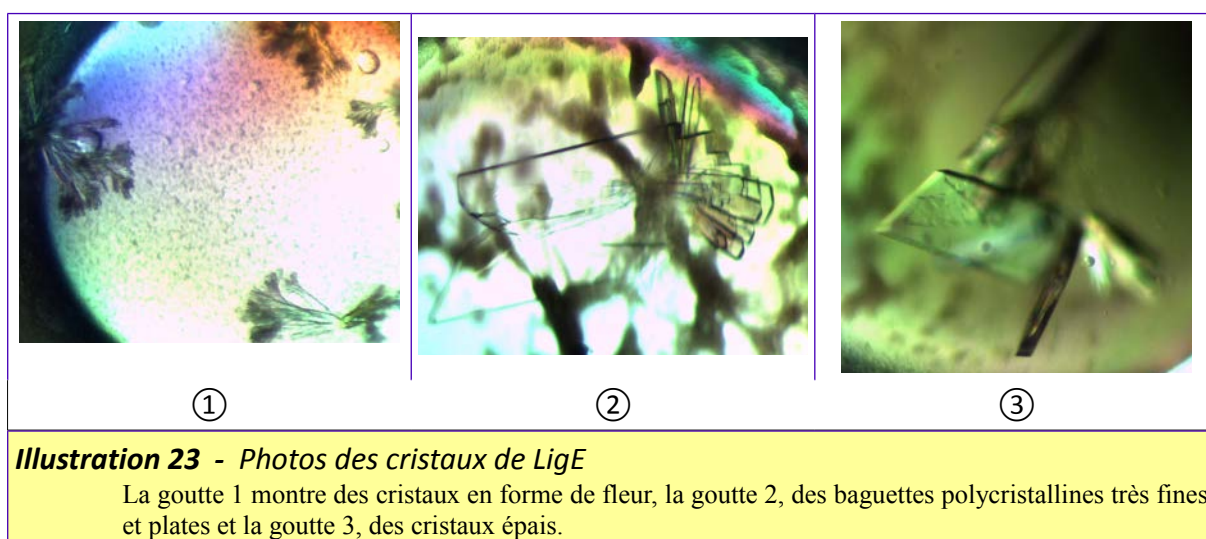
#### ***i* Cristallisation**

La protéine SpLigE concentrée à 32 mg/ml est stable dans le tampon TE et peut être conservée plusieurs semaines à -20°C dans ce tampon. Les conditions de cristallisation ont été déterminées en utilisant les méthodes de microbatch dans l'huile et de la goutte suspendue à 4°C et à 20°C. Les solutions d'agents précipitants des kits commerciaux JBS (1 à 10) de Jena Bioscience et Hampton I et II d'Hampton Research ont été testées. A ces températures, seulement des cristaux de sels ont été obtenus accompagnés d'une précipitation de la protéine mais pas de germination. Cependant en utilisant un kit d'optimisation initialement destiné à cristalliser SpLigF, nous avons observé une germination de SpLigE à l'aide de la méthode batch à 20°C. La condition contient 0,2 M de chlorure de magnésium, 0,1 M d'HEPES à pH 7,0 et 20% w/v de polyéthylène glycol 8000. Un mélange équivalent des solutions protéique et cristallisante permet l'obtention en 2 à 7 jours de cristaux creux en forme de fleur à température ambiante (Illustration 23). De plus, l'ajout dans les gouttes de 0,5  $\mu$ L de phénol (0,2M) comme additif permet l'apparition de fines baguettes

monocristallines en une semaine. Cette condition de cristallisation a été utilisée pour l'obtention des cristaux de la protéine SpLigE sélénométhionylée.

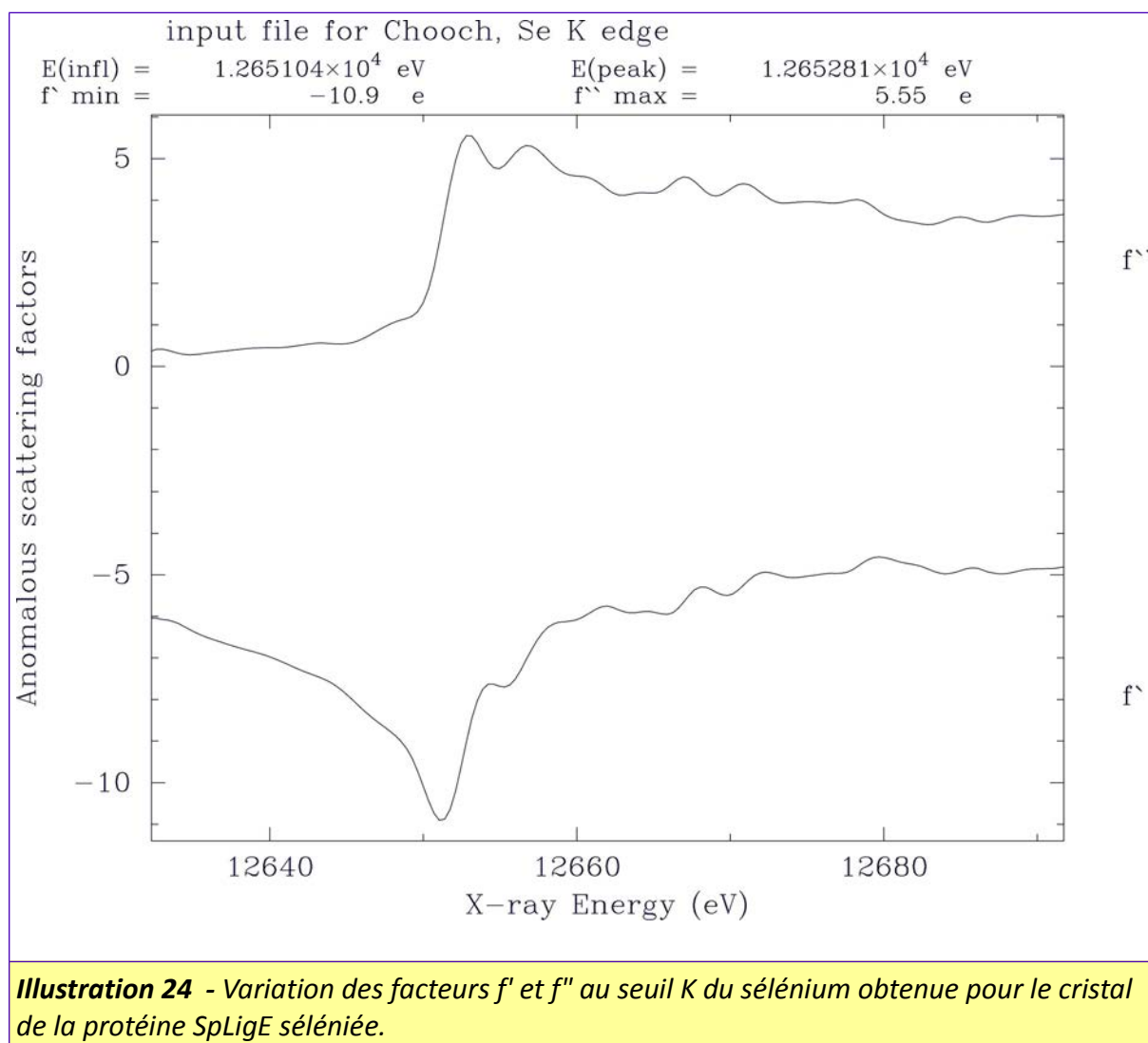
Nous avons réussi à améliorer la taille et la forme des cristaux de la protéine native à l'aide de la méthode de la goutte assise qui permet de ralentir la croissance des cristaux. Nous avons utilisé dans les gouttes la même solution cristallisante en ajoutant deux additifs (0,5  $\mu$ L de phénol (0,2 M) et 0,5  $\mu$ L de n-propanol (0,1 M)) et tout en faisant varier la concentration des agents précipitants de la solution du puits (1 mL). Des cristaux épais de taille convenable ont été obtenus pour la collecte des données de diffraction.

Les cristaux ont été congelés dans une solution de cryoprotection contenant la solution cristallisante supplémentée de 20% de glycérol.



## ***ii Collecte des données de diffraction, phasage et construction du modèle***

Les expériences de diffraction avec les cristaux de la protéine sélénée ont été faites à basse température (100 K) au synchrotron ESRF à Grenoble. Le test de fluorescence réalisé sur le cristal choisi a permis de déterminer la longueur d'onde du seuil du sélénium de l'échantillon (0,97918 Å). Le programme CHOOCH (Evans & Pettifer, 2001) a permis d'estimer les valeurs de  $f'_{\min} = -10,9$  e et  $f'_{\max} = 5,5$  e (Illustration 24). La diffraction du cristal a été enregistrée pour une détermination des phases par la méthode MAD (Multiple Anomalous Dispersion) à trois longueurs d'onde correspondant au pic, à l'inflexion du seuil du sélénium et à une position éloignée du pic (remote). La protéine sélénée cristallise dans le groupe d'espace  $P2_1$  et la limite de diffraction des cristaux est de 4 Å. Les données ont été traitées avec les logiciels XDS (Kabsch, 2010) et SCALA (Evans, 2006) de la suite de programmes CCP4 (Winn *et al.*, 2011). Nous n'avons pas réussi à phaser la structure par la méthode MAD probablement car le signal anomal n'est complet qu'à 6 Å de résolution et la qualité des données de diffraction est médiocre (Tableau 5).



La collecte de données de diffraction à partir des cristaux optimisés de la protéine SpLigE native a été effectuée à basse température de 100 K sur la ligne PROXIMA1 à la longueur d'onde de 0,97911 Å. Les cristaux de la protéine native (groupe d'espace  $P1$ ) ne sont pas isomorphes de ceux de la forme sélénée. La limite de diffraction de la protéine native est de 2,35 Å. Les statistiques des données traitées sont résumées dans le Tableau 6. Le programme MOLREP (Vagin & Teplyakov, 2010) de CCP4 (Winn *et al.*, 2011) a été utilisé pour tenter d'obtenir des phases par la méthode du remplacement moléculaire. La structure en cours d'affinement de la LigF de *Sphingobium* sp SYK-6 et celles des GSTFuA ont été utilisées comme modèle. Aucune solution de remplacement moléculaire claire n'a été déterminée à ce jour.

Caractéristiques du cristal			
Groupe d'espace	$P2_1$		
Paramètres de maille	a=108.58 Å b=58.99 Å c=181.94 Å; $\beta=91.52^\circ$		
Collecte de données	pic	inflexion	lointain
Longueur d'onde (Å)	0.97989	0.98001	0.97777
Resolution (Å)	49.49-4.00 (4.21-4.00)	49.65-3.99 (4.21-3.99)	49.74-3.99 (4.21-3.99)
Rmerge (%)	25.6 (72.2)	27.8 (91.5)	35.5 (142.4)
Réflexions uniques	19317 (2539)	19792 (2692)	19912 (2726)
Complétude (%)	96.5 (87.1)	97.7 (92.0)	98.0 (93.4)
Redondance moyenne	6.3 (4.9)	6.6 (5.4)	6.4 (4.7)
I/ $\sigma$ moyen	6.2 (2.3)	6.2 (1.9)	5.1 (1.3)

**Tableau 5** - Résumé des principales statistiques de collecte des données de diffraction de la protéine SpLigE-Se.  
Les valeurs entre parenthèses concernent les statistiques de la plus haute couronne de résolution.

Caractéristiques du cristal	
Groupe d'espace	$P1$
Paramètres de maille	a=80.46 Å b=94.32 Å c=187.87 Å ; $\alpha=99.76^\circ$ $\beta=92.03^\circ$ $\gamma=89.97^\circ$
Collecte de données	
Longueur d'onde (Å)	0.97911
Resolution (Å)	49.05-2.35 (2.47-2.35)
Rmerge (%)	4.4 (39.9)
Réflexions uniques	222940 (32073)
Complétude (%)	98.1 (96.5)
Redondance moyenne	2.0 (2.0)
I/ $\sigma$ moyen	11.2 (2.1)

**Tableau 6** - Résumé des principales statistiques de collecte des données de diffraction de la SpLigE.  
Les valeurs entre parenthèses concernent les statistiques de la plus haute couronne de résolution.

## B.IV.2 LigF

### B.IV.2.a Introduction

LigF comporte 257 acides aminés. A sa découverte (Masai *et al.*, 1993), une comparaison de séquences l'avait détectée proche de la GST-III de *Zea mays*, et l'alignement alors réalisé laissait supposer un résidu catalytique de type tyrosine. A la lumière des connaissances actuelles, cette hypothèse était évidemment en désaccord avec le fait que la GST de maïs appartient à la classe Phi

pour lequel un résidu catalytique sérine est observé. Le nombre de GST caractérisées augmentant, une étude plus récente des GST bactériennes apparente en fait LigF et LigG aux GST Oméga, ce qui leur conférerait un résidu catalytique cystéine (Allocati *et al.*, 2009). Tout comme LigE, LigF ne présente pas d'activité glutathion transférase vis-à-vis du CDNB. L'inactivation du gène ligF résulte en une bactérie ne transformant que 50 % d'un mélange racémique de MPHPV, le composé restant étant bien dégradé par ajout de LigF dans le milieu. Ce résultat indique que la bactérie ne dispose d'aucune voie alternative dans la dégradation de (+)-(βS)-MPHPV (Masai *et al.*, 2003). Rappelons enfin que LigG est stéréospécifique du GS-HPV obtenu via LigF.

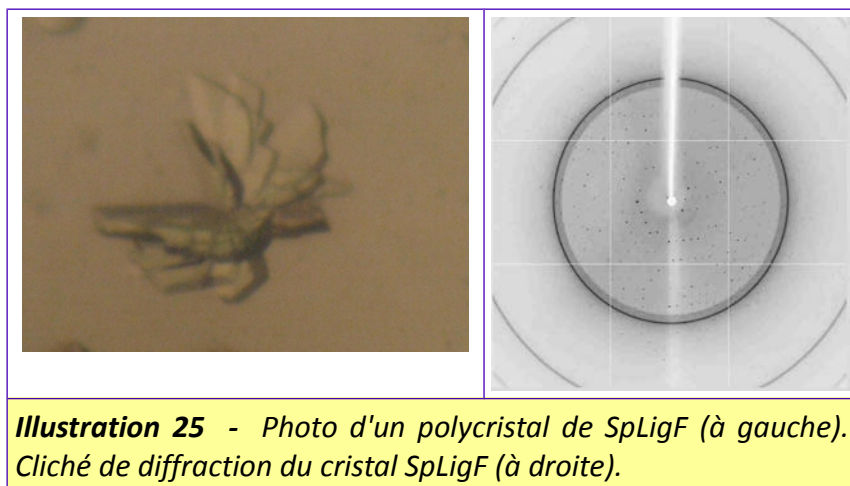
La protéine LigF de *Sphingobium* sp. SYK-6 (notée SpLigF) a été surexprimée et purifiée dans l'équipe du Pr. E. Gelhaye. Les lignes qui suivent présentent la cristallisation, les collectes de données, et la résolution préliminaire de la structure, pour lequel un premier modèle a été obtenu. Son affinement complet n'est actuellement pas achevé, mais une première analyse du modèle provisoire de SpLigF peut néanmoins être réalisée.

#### **B.IV.2.b Matériel et Méthodes**

##### ***i* Cristallisation de la protéine**

Les conditions de cristallisation pour la protéine SpLigF concentrée à 23mg/ml dans du tampon TE (TRIS, EDTA, pH 8) ont été déterminées en utilisant la méthode microbatch dans l'huile et celle de la goutte suspendue à 4°C, à 20°C et à température ambiante (température non régulée). Les solutions d'agents précipitants des kits commerciaux JBS (1 à 10) de Jena Bioscience et Hampton I et II d'Hampton research ont été testées. A température ambiante uniquement et à l'aide de la méthode microbatch, il a été possible d'obtenir une condition de cristallisation donnant un précipité cristallin. Cette condition (D2 du screen commercial JBS2 (Jena BioScience)) contient 0,2M de chlorure de calcium, 0,1M d'HEPES à pH 7.0 et 30% w/v de polyéthylène glycol 4000. Des polycristaux apparaissent au bout d'une semaine en ajoutant dans les gouttes 0.5 µL de phénol (0.1 M) (screen commercial Additif, Hampton Research), ce groupement étant présent dans les substrats naturels de type lignine (Illustration 25). Finalement, des monocristaux de SpLigF ont été obtenus en déplaçant les boîtes de cristallisation de la température ambiante à 20°C.

Les cristaux ont été congelés sans être immergés dans une solution de cryoprotection ce qui explique les anneaux d'eau sur les clichés de diffraction (Illustration 25).



**Illustration 25** - Photo d'un polycristal de SpLigF (à gauche). Cliché de diffraction du cristal SpLigF (à droite).

## ***ii Enregistrement des données de diffraction, phasage et construction du modèle***

Les données de diffraction d'un cristal de SpLigF a été collecté sur la ligne FIP/BM30A à la longueur d'onde de 0,97854 Å. Les statistiques issues du traitement des données sont listées dans le Tableau 7. Nous avons utilisé la méthode du remplacement moléculaire pour phaser les facteurs de structure de SpLigF à l'aide du programme PHASER MR (McCoy *et al.*, 2007) de CCP4 (Winn *et al.*, 2011). La recherche des plus proches homologues structuraux a révélé deux candidats, à savoir la GST Tau de blé (code pdb 1GWC) et la GST Bêta d' *E. coli* (code pdb 2GDR). Les structures primaires de ces deux GST partagent environ 20 % d'identité et 40% d'homologie avec celle de la SpLigF. La structure de la GST Tau s'est avérée le meilleur modèle car 4 monomères peuvent être placés correctement par PHASER MR alors que seulement deux le sont avec le modèle de la GST Bêta. Un modèle incomplet de SpLigF a ensuite été construit automatiquement avec BUCCANEER (Cowtan, 2006). Un travail de reconstruction manuelle a été entrepris à l'aide de COOT (Emsley *et al.*, 2010) suivi de quelques étapes d'affinement avec PHENIX (Adams *et al.*, 2010) en utilisant les symétries non-cristallographiques comme restreinte. Les cartes de densité électronique ne révèlent pas de pics dans le site de fixation du glutation de SpLigF qui semble donc dans une forme apo. L'affinement n'a pas été achevé, faute de temps.

## **B.IV.2.c Résultats**

### ***i Analyse globale du modèle préliminaire***

La SpLigF cristallise dans le groupe d'espace *C2* avec dans l'unité asymétrique quatre monomères qui s'agencent sous la forme de deux dimères classiques de GST. Dans chaque chaîne du modèle préliminaire, 205 résidus ont été construits sur les 257 de SpLigF. Le monomère présente le repliement canonique des GST avec au niveau N-terminal les deux motifs ( $\beta_1\alpha_1\beta_2$ ) et ( $\beta_3\beta_4\alpha_3$ ) du



Caractéristiques du cristal		
Groupe d'espace	C2	
Paramètres de maille	a=159.151 Å b=65.098 Å c=95.568 Å ; $\beta=96.288^\circ$	
Collecte de données		
Longueur d'onde (Å)	0.97854	
Resolution (Å)	47.50-2.50 (2.63-2.50)	
Rmerge (%)	10.0 (52.7)	
Réflexions uniques	31863 (12517)	
Complétude (%)	94.0 (92.3)	
Redondance moyenne	2.9 (2.7)	
I/ $\sigma$ moyen	6.5 (1.6)	
Méthode de phasage		
Remplacement Moléculaire		
Phaser MR	1GWC	
Affinement		
Rwork	25.91 (31.44)	
Rfree	30.71 (38.35)	
RMSD	Longueur de liaison (Å)	0.011
	Angle de liaison ( $^\circ$ )	1.731
	Facteur B moyen (Å <sup>2</sup> )	50.29
Contenu de l'unité asymétrique		
Monomères	4	
Résidus/monomère	A/B/C/D: 205	
Distribution de Ramachandran		
Favorable (%)	86.57	
Accepté (%)	9.33	
Outliers (%)	4.10	
<b>Tableau 7 - Résumé des principales statistiques d'enregistrement, résolution et affinement de la structure de SpLigF.</b>		
Les valeurs entre parenthèses concernent les statistiques de la plus haute couronne de résolution.		

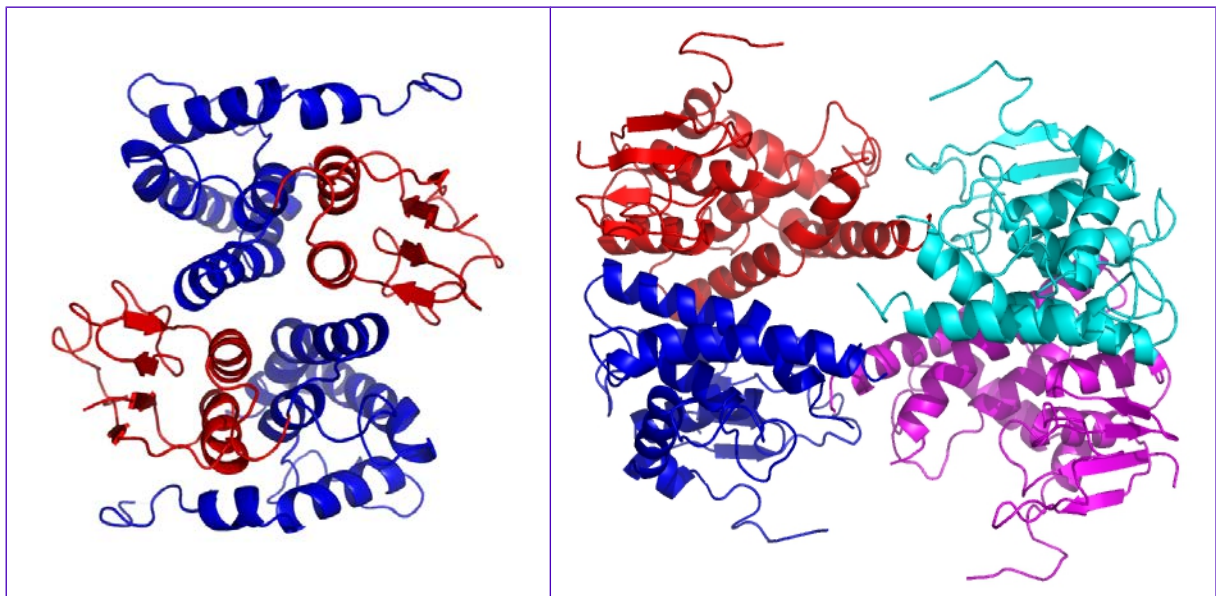
domaine thiorédoxine bien définis dans la densité électronique. L'hélice  $\alpha_2$  les reliant semble absente ou déstructurée. Le domaine C-terminal est composé des cinq hélices habituelles ( $\alpha_4\alpha_5\alpha_6\alpha_7\alpha_8$ ). Toutefois, il n'a pas été possible de construire une quarantaine de résidus situés entre les hélices  $\alpha_4$  et  $\alpha_5$  (119 à 156). Cette région n'est pas présente dans la plupart des GST et correspond donc à une insertion par rapport au repliement canonique. La prédiction des structures secondaires sur la base de la structure primaire à l'aide de Jpred3 (Cole *et al.*, 2008) révèle que cette région manquante devrait se replier sous la forme de deux hélices. Pour finir, 12 résidus de la partie C-

terminale (246 à 257) n'ont pas été construits. Dans le modèle préliminaire, il reste la reconstruction de boucles pour régulariser les angles de torsion phi et psi des résidus qui sont très éloignés des zones favorables de la carte de Ramachandran (voir Tableau 7). Actuellement 86,6% des résidus construits se trouvent dans les régions autorisées. Le modèle préliminaire a été utilisé afin de déterminer les plus proches homologues structuraux de SpLigF à l'aide de DALI (Holm & Rosenström, 2010). L'enzyme la plus proche est la protéine YfcG d'*E. coli* appartenant à la classe des Ure2p et ayant la propriété de fixer des molécules de glutathion par monomère (Wadington *et al.*, 2009). Les deux suivantes (codes pdb 4L8E et 4ECJ) semblent être homologues à YfcG car elles présentent une bonne identité de séquence entre elles (> 45%), néanmoins leurs caractérisations biochimiques et interprétations structurales n'ont pas été publiées à ce jour. La GST Bêta de *Burkholderia xenovorans LB400*, dont le modèle a été utilisé dans le remplacement moléculaire, se trouve en quatrième position. Le modèle de la GST Tau de blé (code pdb 1GWC) qui a permis la résolution de la structure ne se trouve qu'à la 46<sup>ème</sup> place même si cette GST présente le meilleur pourcentage d'identité.

## ***ii Analyse de la structure quaternaire***

L'unité asymétrique du cristal de SpLigF est composée de deux dimères classiques de GST. Ces dimères reliés par un axe deux non cristallographique semblent former un tétramère. Les régions C-terminales des hélices  $\alpha_4$  et N-terminales des hélice  $\alpha_5$  se trouvent au niveau de l'interface entre les deux dimères canoniques. C'est dans cette région (119 à 156) qu'une quarantaine de résidus n'ont pas été modélisés par manque de densité électronique. La stabilisation de cette interface semble se faire par des contacts entre hélices. Le tétramère observé pourrait être la forme biologique active de l'enzyme SpLigF, comme cela a été proposé en solution par Masai *et al.* (2003), par électrophorèse en gel non dénaturant.

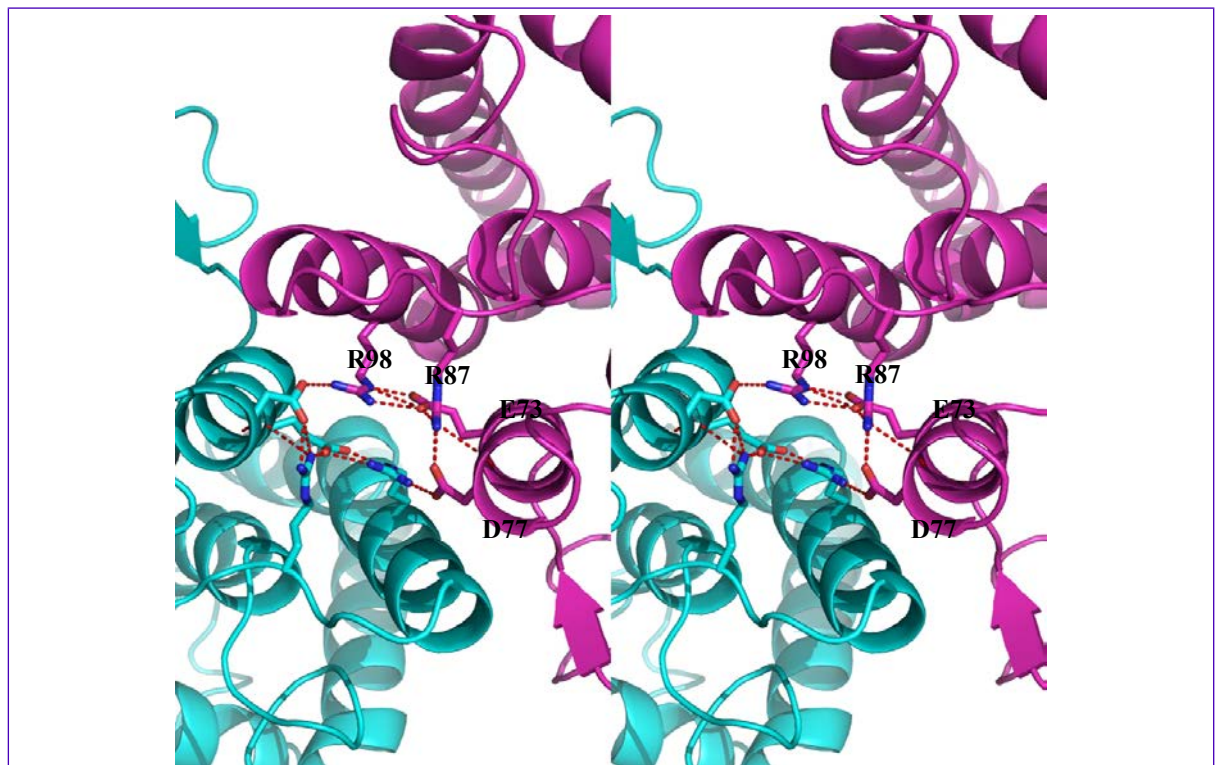
Dans le dimère classique, l'assemblage des sous-unités se fait principalement par l'hélice  $\alpha_3$  du domaine N-terminal d'un monomère avec l'hélice  $\alpha_4$  du domaine C-terminal de l'autre monomère. Dans le cas de SpLigF, l'interface est composée d'un réseau hydrophile impliquant les résidus Glu73 et Asp77 de l'hélice  $\alpha_3$ , l'Arg87 de la boucle  $\alpha_3\alpha_4$  et l'Arg98 de l'hélice  $\alpha_4$ . Il est intéressant de noter d'un réseau hydrophile impliquant des résidus similaires est présent dans la structure de la GST Tau de blé (Thom *et al.*, 2002). Même si l'interface reste hydrophile dans la GST bêta de *Burkholderia xenovorans LB400*, le réseau de liaisons hydrogène implique uniquement des fonctions carboxylate ou hydroxyl (Tocheva *et al.*, 2006). Ainsi l'interface du dimère canonique de SpLigF est hydrophile comme c'est le cas pour la plupart des GST à sérine.



**Illustration 26** - Représentation schématique de la structure de la SpLigF.

Gauche : représentation en flèches et rubans du dimère, avec le domaine N-terminal en rouge et le domaine C-terminal en bleu.

Droite : représentation en flèches et rubans du tétramère, dont chaque monomère est représenté par une couleur différente.



**Illustration 27** - Vue stéréographique de l'interface du dimère canonique de la spLigF

Les deux monomères ont été coloriés respectivement en cyan et magenta

### ***iii Analyse préliminaire de la structure***

Au départ de notre étude biochimique et structurale de SpLigF, la nature de son résidu catalytique n'avait pas pu être déterminée sur la base de sa séquence. Comme nous l'avons montré ci-dessus, les membres de la classe des Ure2p sont les plus proches homologues de SpLigF d'un point de vue structural. Dans la classe des Ure2p, plusieurs propositions ont été suggérées concernant la nature du résidu catalytique. Le groupe d'Armstrong a proposé une thréonine se trouvant à la position du résidu catalytique des GST à sérine (Wadington *et al.*, 2009, Stourman *et al.*, 2011). Cette position est occupée par une glycine (G12) dans la SpLigF. Thuillier et collaborateurs (2013) ont proposé une asparagine deux résidus plus loin (N14) pour l'Ure2p1 de *P. chrysosporium*. Même si cette asparagine est présente dans la SpLigF, seulement des études plus approfondies permettraient de confirmer cette hypothèse.

Plusieurs résidus caractéristiques des GST sont présents dans la SpLigF. Dans le domaine N-terminal, il y a la *cis*-proline 56 dont le résidu précédent stabilise la chaîne principale du glutathion. Il y a également l'acide glutamique 67 et la sérine qui sont les points d'ancrage de la partie N-terminale du glutathion. Au centre de la boucle liant les domaines N-terminal et C-terminal le résidu Leu86 est un résidu important pour le repliement GST car il est situé au cœur hydrophobe impliquant des résidus des deux domaines de l'enzyme. Et enfin le résidu Asp194 du motif (S/TXXD) qui permet la stabilisation du début de l'hélice  $\alpha_6$ .

## **B.IV.3 LigG**

### **B.IV.3.a Introduction**

Dans les voies cataboliques de dégradation de la lignine qu'utilise la bactérie *Sphingobium* sp. SYK-6 pour obtenir les produits nécessaires au cycle de l'acide tricarboxylique, LigG intervient à la suite de LigE et LigF qui permettent de rompre le lien  $\beta$ -aryl éther unissant les monomères constitutifs de la lignine entre eux (voir paragraphe A.III.3). Ces deux enzymes produisent des intermédiaires  $\beta$ -thioéthers conjugués au glutathion, et LigG est l'enzyme qui permettrait leur déglutathionylation, via une probable cystéine catalytique. Les composés modèles utilisés dans l'équipe du Pr. Masai pour étudier cette voie catabolique montrent cependant que LigG ne prendrait en charge que le produit de la réaction catalysée par LigF, et non celui obtenu avec LigE, ces deux enzymes étant énantiosélectifs (Masai *et al.*, 2003). Tout comme LigF, LigG est prédite appartenir à la famille des GST de classe Oméga. Avec un petit groupe de séquences issues de protéobactéries, l'enzyme de *Sphingobium* sp. SYK-6 (noté SpLigG) semble toutefois définir un groupe indépendant au sein des GST présentant une cystéine catalytique.

Les études réalisées lors de ce stage doctoral ayant pour but la caractérisation structurale des GST de *P. chrysosporium*, un champignon saprophyte capable de croître sur différents bois en décomposition, il semblait naturel d'établir la structure de SpLigG et d'étudier ses propriétés biochimiques, puis d'analyser les résultats par comparaison aux propriétés des GST Oméga d'une part, et aux GHR appartenant à la nouvelle classe structurale Xi mise en évidence chez *P. chrysosporium* d'autre part.

### **B.IV.3.b Résultats et discussion**

Une forme de SpLigG portant des résidus sélénométhionine a été utilisée pour permettre la détermination de la structure par la méthode MAD. Des cristaux tétraonaux comportant un unique monomère dans l'unité asymétrique ont été exploités jusqu'à une résolution de 1,2 Å. Compte tenu de cette excellente limite de résolution pour une GST, 95 % du modèle a pu être construit automatiquement dans la densité électronique résultante et le reste facilement complété manuellement lors des cycles de reconstruction/affinement.

SpLigG présente un repliement GST formant des dimères classiques (via un axe 2 cristallographique). La structure possède les caractéristiques du repliement canonique, auquel s'ajoute toutefois une hélice supplémentaire  $\alpha 8$  comparable à celle des GST Oméga humaines HsGSTO1-1 et HsGSTO2-2, suivie d'une élongation C-terminale de 44 résidus. Il n'est donc pas étonnant que les plus proches structures connues de GST s'avèrent appartenir à la famille Oméga. En revanche, la nouvelle famille des GST Xi identifiée chez *P. chrysosporium* (correspondant aux enzymes initialement appelées PcGSTO1 et PcGSTO2) ne représente qu'un homologue structural au même titre que toute autre GST. Malgré la similitude relative avec les GST Oméga, SpLigG présente des caractéristiques propres. D'une part l'extension C-terminale comporte 44 résidus dont les premiers définissent une petite boucle  ${}_{220}\text{GGGNG}_{224}$  qui amène la chaîne latérale de Asn223 vers le site G. D'autre part, le résidu Asn234 de l'extension C-terminale insère sa chaîne latérale dans le réseau de liaisons hydrogène de l'hélice  $\alpha 8$  de sorte à le rompre et à courber cette hélice de 60°, dont certains résidus participent à la formation du site actif.

Parallèlement à ces études structurales, les études biochimiques réalisées par le groupe du Pr. Gelhaye démontrent une activité de type glutaredoxine de SpLigG, abolie si le résidu catalytique potentiel, Cys15, est muté en sérine. Une activité de déglutathionylation est également observée, en accord avec l'homologie de séquence des LigG avec les GST Oméga. Son étude fine sur des substrats variés démontre que, comme PcGSTO3 homologue aux GSTO humaines et contrairement à PcGSTO1 présentant une activité GHR et appartenant à la classe Xi (et donc rebaptisée PcGSTX1, voir paragraphe B.I.2), SpLigG déglutathionyle la phénylacétophénone-SG mais pas la ménadione-SG. Des différences apparaissent cependant avec PcGSTO3, puisque

SpLigG est inactive pour déglutathionyler la tétralone-SG ; de plus, elle n'acquière pas d'activité glutathion transférase vis-à-vis du CDNB si Cys15 est mutée en sérine.

Le site actif est particulièrement large et ouvert. Le GSH semble se poser contre sa paroi, et est stabilisé par des interactions très classiques auxquelles s'ajoute une liaison hydrogène avec la chaîne latérale d'Asn223 via une molécule d'eau. La chaîne latérale de son résidu cystéinyl approche celle de Cys15 démontrée être le résidu catalytique. La région du site actif avoisinant ces deux cystéines présente une poche formée par des donneurs de liaison hydrogène, notamment Tyr214 appartenant à  $\alpha 8$  et la molécule d'eau stabilisée par Asn223. Ce site hydrophile (qui dans la structure accueille un ion acétate issu du milieu de cristallisation) est voisin d'une région plus hydrophobe pouvant probablement stabiliser un substrat tel que la  $\beta$ -hydroxypropiovanillone glutathionylée (GS-HPV) issue de la réaction catalysée par LigF dans l'étude modèle réalisée par Masai *et al.* (2003) (voir paragraphe A.III.3). Une modélisation du complexe LigG/GS-HPV a été réalisée, qui valide cette hypothèse. Seul l'un des deux diastéréoisomères de GS-HPV adopte une position plausible dans le site actif, en accord avec la stéréosélectivité observée pour LigG (Masai *et al.*, 2003).

En conclusion, ces résultats combinés démontrent que SpLigG ne s'apparente pas du tout aux GHR de classe Xi, et ressemble plutôt aux GSTO, avec toutefois des différences de profil d'activité probablement liées à des différences structurales au niveau du site actif. Le groupe des LigG est donc à rattacher à la classe des GSTO qui forment ainsi un groupe plus étendu, mais dans lequel les LigG sont marginales.

#### **B.IV.3.c Publication**

Voir texte page suivante.



ELSEVIER

# FEBS Letters

journal homepage: [www.FEBSLetters.org](http://www.FEBSLetters.org)

## *Sphingobium* sp. SYK-6 LigG involved in lignin degradation is structurally and biochemically related to the glutathione transferase omega class

Edgar Meux<sup>a,b,1</sup>, Pascalita Prosper<sup>c,d,1</sup>, Eiji Masai<sup>e</sup>, Guillermo Mulliert<sup>c,d</sup>, Stéphane Dumarçay<sup>f</sup>,  
Mélanie Morel<sup>a,b</sup>, Claude Didierjean<sup>c,d</sup>, Eric Gelhaye<sup>a,b,\*</sup>, Frédérique Favier<sup>c,d,\*</sup>

<sup>a</sup> Université de Lorraine, IAM, UMR 1136, IFR 110 EFABA, Vandoeuvre-les-Nancy, F-54506, France

<sup>b</sup> INRA, IAM, UMR 1136, Vandoeuvre-les-Nancy, F-54506, France

<sup>c</sup> Université de Lorraine, CRM2, UMR 7036, Vandoeuvre-les-Nancy, F-54506, France

<sup>d</sup> CNRS, CRM2, UMR 7036, Vandoeuvre-les-Nancy, F-54506, France

<sup>e</sup> Department of Bioengineering, Nagaoka University of Technology, Kamitomioka, Nagaoka, Niigata 940-2188, Japan

<sup>f</sup> Université de Lorraine, LERMAB, EA 1093, Vandoeuvre-les-Nancy, F-54506, France

### ARTICLE INFO

#### Article history:

Received 10 July 2012

Revised 13 September 2012

Accepted 21 September 2012

Available online 8 October 2012

Edited by Stuart Ferguson

#### Keywords:

Glutathione transferase

LigG

Lignin

X-ray structure

Omega class

Thiol transferase

### ABSTRACT

**SpligG is one of the three glutathione transferases (GSTs) involved in the process of lignin breakdown in the soil bacterium *Sphingobium* sp. SYK-6. Sequence comparisons showed that SpligG and several proteobacteria homologues form an independent cluster within cysteine-containing GSTs. The relationship between SpligG and other GSTs was investigated. The X-ray structure and biochemical properties of SpligG indicate that this enzyme belongs to the omega class of glutathione transferases. However, the hydrophilic substrate binding site of SpligG, together with its known ability to stereoselectively deglutathionylate the physiological substrate  $\alpha$ -glutathionyl- $\beta$ -hydroxypropiovanillone, argues for broadening the definition of the omega class.**

*Structured summary of protein interactions:*

**SpligG** and **SpligG** bind by X-ray crystallography (View interaction).

© 2012 Federation of European Biochemical Societies. Published by Elsevier B.V. All rights reserved.

### 1. Introduction

Glutathione transferases (GSTs) constitute a multigenic family with fundamental roles in cellular detoxification of a wide range of exogenous and endogenous compounds. Most GSTs characterized until now catalyze the conjugation of the tripeptide glutathione (GSH) to compounds containing an electrophilic center, generally increasing the solubility and diminishing the toxicity of the resulting compounds. The highly reactive thiolate anion of GSH is usually stabilized by a serinyl or tyrosinyl residue conserved in the GSH binding site (G-site) of GSTs [1]. In contrast, several classes of GSTs contain at this place a catalytic cysteine conferring

them thiol transferase activity [1]. Whereas a number of cysteine-containing GST classes span several kingdoms and phyla (for example omega (O) in mammals, fungi and insects [2]; S-glutathionyl hydroquinone reductase (GHR) in fungi and prokaryotes), others are limited to a specific kingdom or phylum (lambda (L) and DHAR in plants; beta (B) and LigG in bacteria). Although the principle in vivo activities of cysteine-containing GSTs remain unclear, recent advances indicate that they are able to bind a large range of small molecules and could thus be involved in various metabolic pathways. Supporting this view, human omega class GSTs could participate in arsenic biotransformation [3], while bacterial and fungal GHRs could be involved in the catabolism of chlorinated quinones such as pentachlorophenol, a fungicide used for wood preservation [4]. In addition *Arabidopsis thaliana* lambda GSTs were recently shown to bind oxidized anthocyanidins and tocopherols, suggesting putative roles in response to oxidative stress [5].

In *Sphingobium* sp. strain SYK-6 (an aerobic Gram-negative soil bacillus), *ligE*, *ligF* and *ligG* are the three GST encoding genes found in an operon involved in the process of lignin breakdown. SpligE and SpligF catalyze the transfer of GSH to an aryl ether substrate resulting in a  $\beta$ -thioether intermediate and SpligG catalyzes elimination of GSH from the conjugate product [6,7]. Sequence

\* Corresponding authors. Addresses: Université de Lorraine, Interactions Arbres/Micro-Organismes (IAM), UMR 1136, IFR 110 EFABA, Faculté des Sciences et Technologies, Bld des Aiguillettes, BP 70239, Vandoeuvre-les-Nancy, F-54506, France. Fax: +33 3 68 42 28 (E. Gelhaye), Université de Lorraine, Laboratoire de Cristallographie, Résonance Magnétique et Modélisations (CRM2), Equipe Biocristallographie, UMR 7036, Faculté des Sciences et Technologies, Bld des Aiguillettes, BP239, Vandoeuvre-les-Nancy, F-54506, France. Fax: +33 3 68 43 00 (F. Favier).

E-mail addresses: [eric.gelhaye@lcb.uhp-nancy.fr](mailto:eric.gelhaye@lcb.uhp-nancy.fr) (E. Gelhaye), [frederique.favier@crm2.uhp-nancy.fr](mailto:frederique.favier@crm2.uhp-nancy.fr) (F. Favier).

<sup>1</sup> Both authors contributed equally to this work.

comparison shows that SpLigG possesses several proteobacteria homologues with which it forms an independent cluster among cysteine-containing GSTs (Fig. 1).

In this study we present the three-dimensional structure of SpLigG, and the comparison of its biochemical properties with omega and xi (GHR) GSTs from *Phanerochaete chrysosporium* characterized recently by our group [8]. The results show a clear relationship between SpLigG and omega GSTs.

## 2. Materials and methods

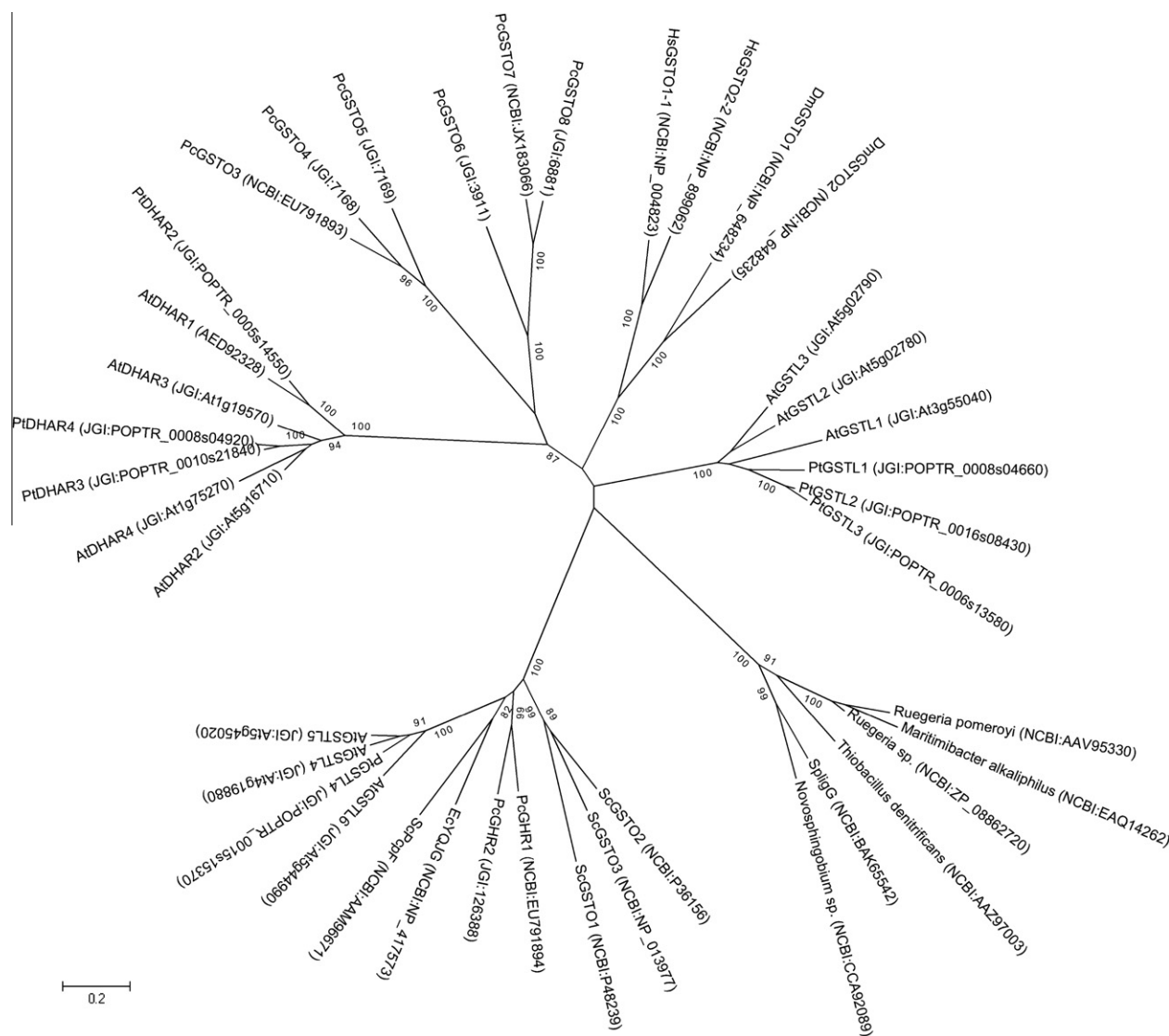
### 2.1. Cloning, site-directed mutagenesis, production and purification of recombinant proteins

The cloning of the *Spingobium* sp. SYK-6 LigG and LigF into pET-21 (+) has been described previously [9]. The LigG\_C15S mutant of the active site cysteine has been obtained by PCR-based site-directed mutagenesis using two complementary mutagenic primers (5'TCACATTCCCGCAGCCCTTCTCCGAGCG3' and 5'CGCT

CGGAGAAGGGGCTGCCGGGAATGTGA3'). *Escherichia coli* BL21(DE3) culture conditions and protein purification procedures were similar to the ones previously described [8]. For the production of the selenomethionine tagged SpLigG, recombinant plasmid carrying the gene of interest was electroporated into methionine auxotrophic strain of *E. coli* BL21 (DE3) and further production and purification steps were performed as described by Koh et al. [10]. After purification to homogeneity, electrospray mass analysis of SpLigG (30830 Da) showed that the N-terminal methionine residue had been cleaved, that a GSH adduct was present and that all eight methioninyl residues had been substituted.

### 2.2. Structure determination of SpLigG

The selenomethionylated form of SpLigG (19 mg/ml in 30 mM Tris-HCl pH 8.0, 1 mM EDTA buffer) was mixed (ratio 1:1) with DMSO containing 10 mM of glutathionylated-menadione. Droplets for the batch method under oil were prepared by mixing 1.5 µl of this solution with 1.5 µl of a solution containing 0.2 M ammonium



**Fig. 1.** Unrooted phylogenetic tree of cysteine-containing glutathione transferases from different phyla (At = *Arabidopsis thaliana*; Pt = *Populus trichocarpa*; Pc = *Phanerochaete chrysosporium*; Hs = *Homo sapiens*; Dm = *Drosophila melanogaster*; Sc = *Sphingobium chlorophenolicum*; Ec = *Escherichia coli*; Sp = *Spingobium* sp. strain SYK-6). SpLigG homologues (bottom right) share less than 20% of sequence identity with the other represented GSTs. Sequence alignments were done by ClustalW. The tree was constructed with the neighbour-joining method in MEGA 5.0 software. The robustness of the branches was assessed by the bootstrap method with 1000 replications. Bootstrap values are placed at the nodes where the value is >70.



sulfate, 0.1 M sodium acetate pH 4.6, 25% polyethylene glycol 4,000 (corresponding to reagent #20 of the Crystal Screen 1 by Hampton Research) The samples stored at 277 K gave crystals suitable for X-ray measurement within 4–7 days. They were flash-frozen at 100 K using glycerol (10%) as the cryoprotectant.

The diffraction data (Table 1) were collected at the FIP-BM30A synchrotron beamline (ESRF, Grenoble) at two wavelengths corresponding to the peak and inflection of the Se K-edge, and at a remote one. The three data sets were processed with XDS [11] and SCALA [12]. Se atom search (eight atoms) and phasing by multiple anomalous diffraction (MAD) were performed at 2.2 Å resolution with SHELXC/D/E [13], leading to a correlation coefficient of 57.7% at the end of SHELXD. Further processing with SHELXE revealed a pseudo-free CC of 83.9% allowing for identification of the correct hand (48.0% for the reverse one). A preliminary model was built with BUCANEER (251 residues) [14]. It was completed/corrected manually with COOT [15] and refined with PHENIX [16]. Complete statistics are reported in Table 1.

### 2.3. Enzyme assays

Guaiacylglycerol- $\beta$ -O-4-methylumbelliferone (GOU) has been prepared as described previously [17] and used as a  $\beta$ -aryl ether substrate. SpLigF (5  $\mu$ M) was mixed with 2 mM GSH, 50  $\mu$ M GOU in 30 mM Tris-HCl pH 8.0 and stirred for 20 min at room temperature. Thus, all GOU molecules were expected to be transformed in 4-methylumbelliferone and  $\alpha$ -glutathionyl- $\beta$ -hydroxypropiovanillone (GS-HPV) [18]. Then cysteine-containing GSTs (1  $\mu$ M) have been added to the mix with NADPH (200  $\mu$ M) and 0.5 IU glutathione reductase. The de-glutathionylation activity was followed by monitoring the decrease in absorbance arising from NADPH oxidation in this coupled enzyme assay system which reflects the formation of glutathione disulfide (GSSG).

All experiments were performed with enzymes as purified, or after reduction with a 10-fold excess of dithiothreitol (DTT) at 25 °C during 30 min. In the latter case, the mixture was passed through a Sephadex G-25 gel-filtration column (GE Healthcare)

to eliminate the DTT excess. Thiol transferase and glutathione transferase activities against various substrates have been tested according to Meux et al. [8].

Cysteine  $pK_a$  values were determined by measuring the reactivity of the enzymes toward 5-nitrofluoresceine at various pH ranging from 3 to 10.

### 2.4. Substrate docking

In order to dock GS-HPV onto SpLigG, (*R*- or (*S*)-GS-HPV molecules, together with the refined coordinates of SpLigG (with the non-bonded molecules deleted) were submitted to the SwissDock server [19] using the default parameters. The possible GS-HPV positions returned by the software were filtered using two different criteria. First, the positions of the glutamyl residues of GS-HPV in the modeled complex were compared to that of GSH in the SpLigG crystal structure. Solutions in which the rmsd was larger than 1.6 Å were eliminated. Second, a maximal distance between the sulfur atoms of GS-HPV and Cys15 was set to 2.5 Å, in order to retain positions proximal enough for attack by Cys15. These filters resulted in only one solution for (*R*)-GS-HPV, and no solution for (*S*)-GS-HPV.

## 3. Results and discussion

### 3.1. X-ray structure of SpLigG

The three-dimensional structure of SpLigG was solved at a resolution of 1.2 Å, using the MAD method on tetragonal crystals of the selenomethionylated enzyme (see Section 2 and Table 1).

SpLigG possesses the characteristics of the canonical GST fold [20] (Fig. 2): it is a homo-dimer (generated by a crystallographic twofold axis) where the N-terminal thioredoxin-like domain of the monomer contained in the asymmetric unit interacts with the  $\alpha$ -helical C-terminal domain of the related monomer in a neighboring asymmetric unit. PDBeFold [21], DALI [22] and VAST [23] were used to search the PDB for similar structures. The

**Table 1**  
Data collection and refinement statistics.

SpLigG crystal parameters			
Space group		P4 <sub>1</sub> 2 <sub>1</sub> 2	
<i>a</i> , <i>b</i> , <i>c</i> (Å)		56.801, 56.801, 183.730	
Protein molecule in a.u.		1	
Matthews coefficient (Å <sup>3</sup> /Da)		2.55	
Data collection statistics	Peak	Inflection	Remote
Wavelength (Å)	0.98007	0.98020	0.97856
Resolution (Å)	48.26–1.75 (1.85–1.75) <sup>a</sup>	48.27–1.73 (1.83–1.73)	48.31–1.20 (1.27–1.20)
Unique reflections	23598 (1108)	23661 (863)	93486 (13146)
Completeness <sup>b</sup> (%)	75.0/73.7 (25.2/21.3)	72.9/71.5 (19.1/15.0)	98.7/98.6 (96.7/96.0)
<i>R</i> -Merge <sup>c</sup> (%)	3.8 (7.8)	3.6 (8.0)	5.8 (4.8)
<i>I</i> / $\sigma$ <i>I</i>	44.6 (11.0)	47.6 (10.4)	16.8 (4.0)
Redundancy <sup>b</sup>	11.3/6.1 (3.5/2.0)	11.3/6.1 (3.2/1.8)	6.8/3.6 (6.7/3.4)
Refinement statistics			
# Residues/water molecules/ligands	263/376/1 GSH, 1 ACT, 1 SO <sub>4</sub> <sup>d</sup>		
<i>R</i> -All/ <i>R</i> -work/ <i>R</i> -free <sup>e</sup> (%)	14.04/13.96/15.47		
RMSD bond length (Å)/angles (°)	0.011/1.443		
Mean B factor (Å <sup>2</sup> )	14.9 (protein)/25.1 (H <sub>2</sub> O)		
Ramachandran plot regions (%)	97.81 (most favored)/2.19 (additional allowed)/0.0 (disallowed)		
PDB ID	4C10		

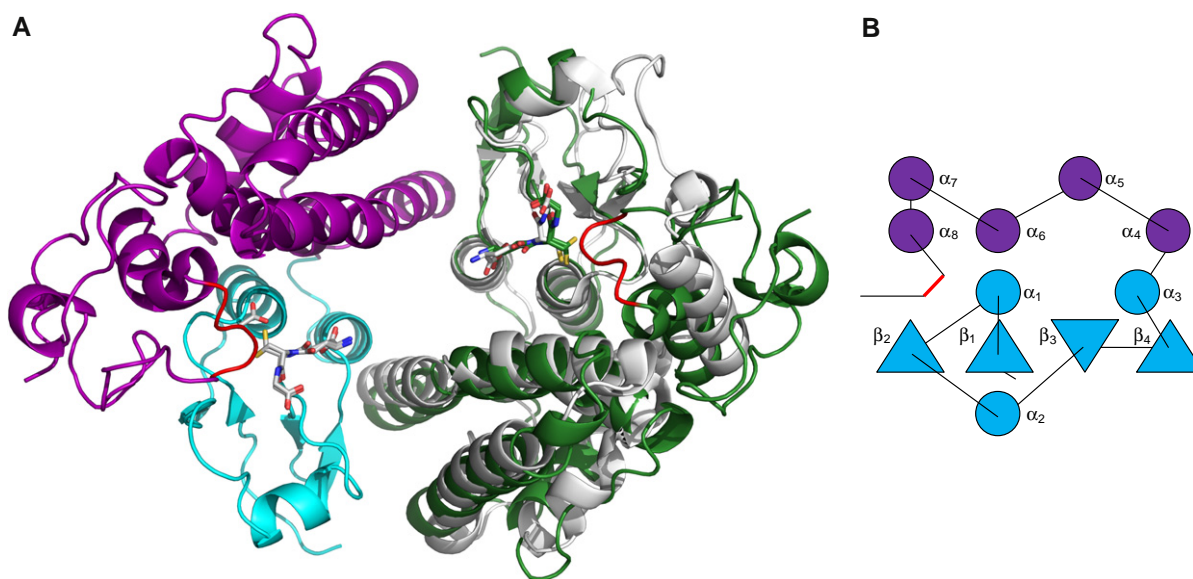
<sup>a</sup> Values in parentheses are for the highest resolution shell.

<sup>b</sup> In the first value of the completeness (resp. redundancy), the anomalous signal is ignored (*I*mean), in the second value it has been taken into account (unmerged *I*+ and *I*-).

<sup>c</sup>  $R$ -Merge =  $\sum_{hkl} \sum_j |<I_{hkl}> - I_{hkl,j}| / \sum_{hkl} <I_{hkl}>$ , where  $I_{hkl,j}$  is the intensity of the  $j^{\text{th}}$  measurement of a reflection  $hkl$  and  $<I_{hkl}>$  is the average intensity of that reflection on its total number of measurements.

<sup>d</sup> GSH: glutathione, ACT: acetate ion, SO<sub>4</sub>: sulfate ion.

<sup>e</sup>  $R_{\text{all}}$  is determined for all reflections (working set + test set),  $R_{\text{work}}$  corresponds to a subset of 95% of reflections (working set),  $R_{\text{free}}$  is for the complementary 5% (test set).



**Fig. 2.** Schematic drawings of the SpLigG structure. (A) The dimer in ribbon mode. The corresponding amino acid sequences are shown in Fig. S1. In the left monomer, the thioredoxin-like domain is colored in cyan, the  $\alpha$ -helix domain is in purple, the gray sticks correspond to GSH (with two alternative conformations) and to the acetate ion. The symmetry-related SpLigG monomer (green) is shown superimposed to the human GSTO2 (in gray, PDB ID: 3Q19), and their corresponding GSH are colored accordingly. The loop specific to SpLigG homologues is in red; (B) topology diagram with the secondary structures numbered.

ranking changed significantly depending on the program used, but regardless the program, hGSTO1-1 (PDB IDs: 3LFL, 1EEM) and hGSTO2-2 (3Q18, 3Q19, 3QAG) followed by GSTO3 from *Bombyx mori* (3RBT) clearly appear as the closest structural neighbors while xi GSTs (3PPU, 3M1G), as well as beta GSTs (1A0F, 1F2E, 1N2A, 1PMT, 2DSA, 2GDR, 2NTO, 2PMT, 2PVQ, 2X64) were found far from the top of the list. The distant ranking of xi GSTs is consistent with the distinctive features characterizing this class of structures, notably where significant loop elongations lead to an atypical dimer [8]. In the same way, beta class GSTs lack both the  $\alpha$ -helix  $\alpha 8$  (Fig. 2) conserved in SpLigG and omega GSTs, and the long C-terminal loop specific to SpLigG.

Interestingly, the invariant aspartyl residue, that belongs to a N-capping box motif present at the beginning of  $\alpha 6$  in the second domain of GSTs [24] is replaced by a glutamate residue (in this case Glu161) which is strictly conserved in close SpLigG homologues. Among the roughly one hundred independent structures available in the Protein Data Bank, only the uncharacterized GST-like proteins from *Pseudomonas fluorescens* (PDB IDs: 3LXT, 3MOF) and *E. coli* (3R2Q) present a natural mutation of Asp in the motif (Gly and Thr, respectively). Despite this difference in its N-capping box motif, SpLigG did not reveal any problems in overexpression or stability, although a critical role in rapid enzyme folding and stability was attributed to the usual aspartyl residue in studies concerning the human pi and alpha class GSTs [25–27].

Despite the high structural similarity between SpLigG and hGSTO1-1 [28,29] or hGSTO2-2 [30], major differences can be observed. In SpLigG,  $\alpha 8$  shows a strong bend ( $\sim 60^\circ$ ), due to the insertion of the Ser234 side chain which breaks the hydrogen bond network between the second and third turns of the helix. Furthermore,  $\alpha 8$  is followed by a long extension of 44 amino acids the beginning of which forms a very specific cap on one side of the GSH binding site. Its residues  $^{220}\text{GGGNG}^{224}$  display a loop which directs the Asn223 side chain towards GSH: a hydrogen bond is formed between its atom OD1 and a water molecule, itself hydrogen bonded to the nitrogen atom of the glycyl moiety of GSH. Sequence comparison shows the high conservation of the C-terminal end in SpLigG homologues (Fig. S1).

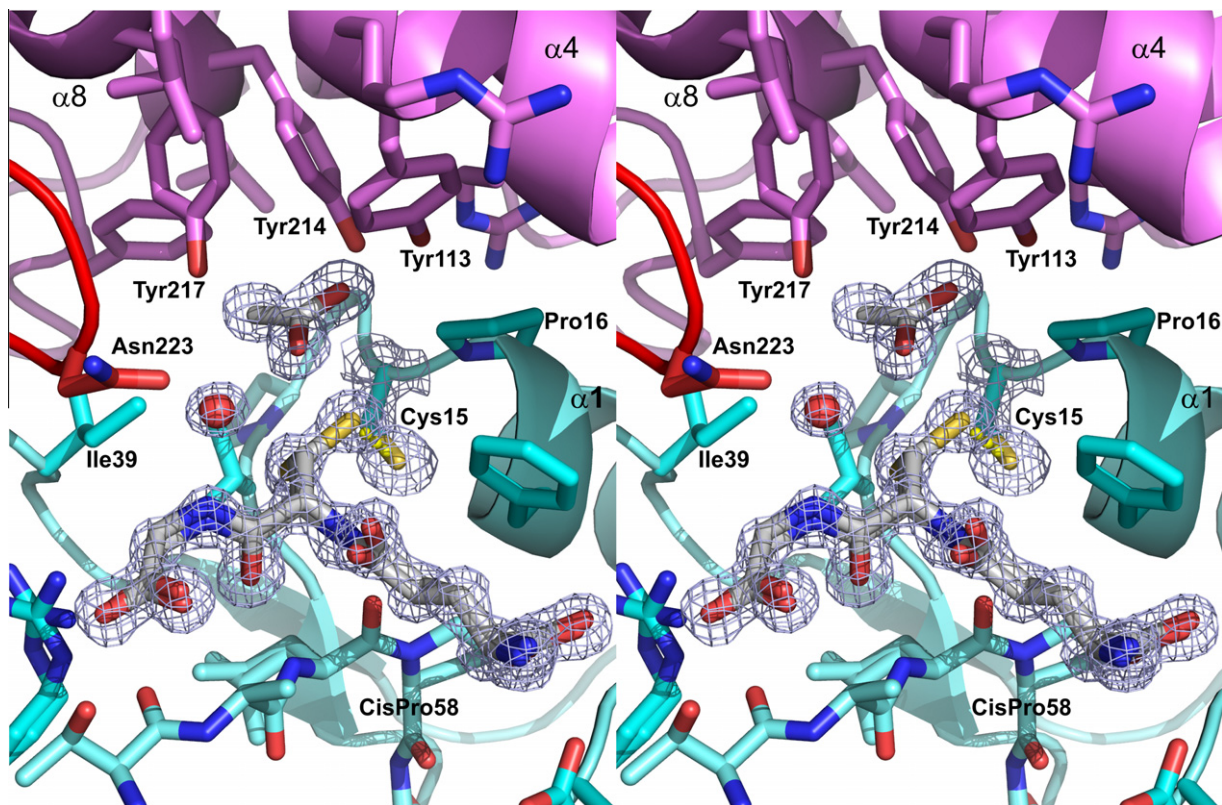
### 3.2. Active site of SpLigG

The anchoring motif of GSH in the G-site of SpLigG is conserved in many GSTs, with the exception of the supplemental hydrogen bond provided by the water molecule stabilized by the Asn223 side chain (see Section 3.1). The side chain of the GSH cysteinyl residue is observed in two distinct positions (Fig. 3), one with its sulfur atom at 3 Å of the catalytic Cys15, the second at 2.4 Å with the highest occupancy (0.63). The latter corresponds possibly to a disulfide bridge disrupted upon X-ray exposure. The presence of a 2/3:1/3 mix of two forms with and without the disulfide bridge is corroborated by the mass spectrometry measurements on the native enzyme (data not shown).

GSH positions itself at the surface of the huge basin observed at the interface between the two SpLigG monomers. This positioning orients the GSH sulfur atom towards a cavity defined by the N-terminal residues of helix  $\alpha 1$  (with the sequence  $^{14}\text{GCPF}_{17}$ , among which is the catalytic Cys15), Ile39 at the C-terminal end of strand  $\beta 2$ , Tyr113 from helix  $\alpha 4$ , Tyr214 and Tyr217 from  $\alpha 8$ , and finally Asn223 and Phe235 from the C-terminal loop. This constitutes an area less hydrophobic than in many GSTs (in particular in GSTOs), and forms an environment that favors hydrogen bonding and ion binding, as attested by the presence of an acetate anion observed in the vicinity. One oxygen atom forms two hydrogen bonds with the hydroxyl groups of Tyr113 and Tyr214, and the second with the water molecule that bridges the nitrogen atoms of the GSH glycyl residue and of the Asn223 side chain (see first Section 3). Conversely, the methyl group of the acetate molecule faces the side chains of Ile39 and Phe235. All of the described residues are strictly conserved in the sequences homologous to SpLigG (Fig. S1). They could constitute one part of the H-site, so called because it corresponds to a usually hydrophobic region of GSTs where the electrophilic substrate or the non-GSH moiety of a glutathionylated substrate can bind.

### 3.3. Biochemical assays using GST substrates

SpLigG exhibits thiol transferase activity towards hydroxyethyl-disulfide (HED) and dehydroascorbate (DHA), with kinetic param-



**Fig. 3.** Stereo view of the active site of SpLigG. The GSH and acetate molecules are shown as gray sticks; the red sphere corresponds to a water molecule; the main residues that form the active site together with the ribbon which figures the SpLigG polypeptide main chain are displayed with the same color code as in Fig. 2. For clarity, only some selected residues are labeled. The 2Fo-Fc electron density map is shown in light blue (contour level 1.2  $\sigma$ ) at the ligand and Cys15 positions. The disulfide bond between Cys15 and GSH probably disrupted upon X-ray exposure is figured as a yellow dashed line.

eters comparable to those obtained with omega 3 GST from *P. chrysosporium* (PcGSTO3) and ten times higher than those obtained with fungal GHR 1 (PcGHR1) towards HED (Table 2). However, as expected, the mutation of Cys15 to serine abolishes thioltransferase activity. It indicates that the active site cysteine is necessary for both DHA and HED reduction as observed for other cysteine-containing GSTs.

Additionally, we used *S*-(phenylacetophenone)-glutathione (PAP-SG), *S*-(tetralone)-glutathione and *S*-(menadione)-glutathione to investigate the ability of SpLigG to catalyze deglutathionylation of small aromatic molecules. SpLigG as well as PcGSTO3

**Table 2**  
Activity of the recombinant fungal PcGHR1, PcGSTO3 and the bacterial SpLigG with various "classical" substrates.

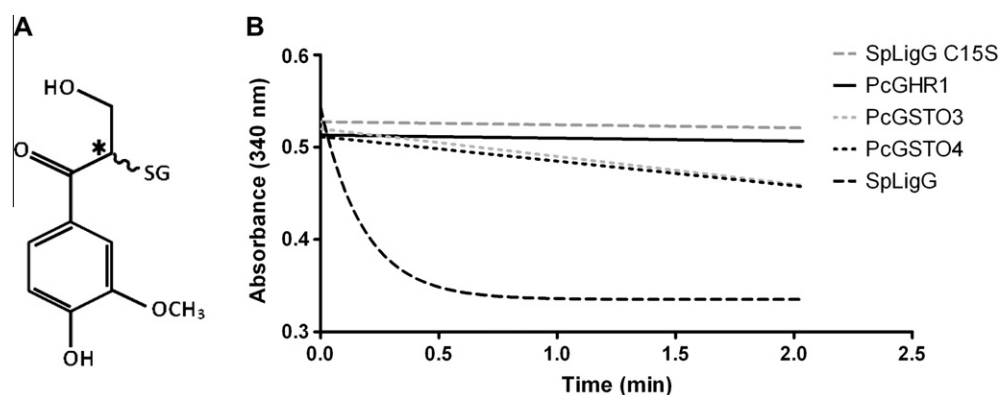
	PcGHR1	SpLigG	PcGSTO3
<i>S</i> -(Menadione)-glutathione	1 185	–	–
DHA	108.6	356.8	58.2
HED	160.2	1 350	943
PAP-SG	–	1 025	1 141
Tetralone-SG	–	–	449.2
Cumene hydroperoxide	–	–	–
Hydrogene peroxide	–	–	–
<i>t</i> -Butyl hydroperoxide	–	–	–
1-Chloro-2,4-dinitrobenzene	–	–	–
Isothiocyanate	–	–	–
	PcGHR1 C117S	SpLiG C15S	PcGT03 C37S
1-Chloro-2,4-dinitrobenzene	–	–	1 105
Isothiocyanate	–	–	10.4

Specific activity is shown in  $\mu\text{mol}/\text{min}/\mu\text{mol}$ . Data were compiled from Meux et al. [8]. Symbol (–) means not detectable.

[8] are active against PAP-SG from which they remove GSH. The reaction between DTT-treated SpLigG (allowing removal of GSH found in the untreated protein) and PAP-SG in the absence of GSH led to the formation of phenylacetophenone detected by HPLC and of glutathionylated SpLigG as shown by Q-TOF analysis of the reaction products (data not shown). In these conditions the reaction occurs only once and does not proceed any further. The deglutathionylated product is neither detected using the native purified protein (i.e. with a GSH adduct) nor with the C15S mutant. These results are in accordance with the involvement of Cys15 in the deglutathionylation reaction leading to the formation of a disulfide bridge with the GSH molecule removed from PAP-SG. When GSH is added to the mix, glutathione disulfide is detected by mass spectrometry (data not shown), in accordance with a role of GSH in the enzyme regeneration. SpLigG as well as PcGSTO3 are inactive towards *S*-(menadione)-glutathione which is a specific substrate of GHRs [4,8] (Table 2). On the other hand, SpLigG is also inactive against tetralone conjugates in contrast to PcGSTO3. Furthermore, subsequent studies showed that SpLigG is largely inactive as a glutathione transferase against 1-chloro-2,4-dinitrobenzene or hydroperoxides. The mutation of the catalytic cysteine to a serine does not confer a transferase activity as previously described for GSTOs [8]. Thus from a biochemical point of view, it appears clearly that SpLigG does not possess GHR activity and that it is functionally related to the omega class although it is not able to use all described GSTO substrates.

#### 3.4. $\beta$ -Thioether lignin intermediate as a natural substrate of SpLigG

SpLigG is able to catalyze the second step of the  $\beta$ -aryl ether bond cleavage initiated by SpLigF and SpLigE [6]. The standard sub-



**Fig. 4.** Degradation of the  $\beta$ -thioether intermediate (GS-HPV) by recombinant PcGHR1, PcGSTO3, PcGSTO4, SpLigG and the C15S mutant in the presence of GSH. In (A), SG represent the glutathione linked to the asymmetric carbon (\*) by its sulfur atom. In (B), the activity was followed by monitoring the decrease in absorbance arising from NADPH oxidation (B).

strate GOU was first incubated with the purified  $\beta$ -etherase SpLigF and reduced GSH in order to generate a  $\beta$ -thioether intermediate (GS-HPV) and then different cysteine-containing GSTs were added to the mix. The addition of SpLigG (0.25  $\mu$ M) led to a decrease in the OD at 340 nm which reflects the production and the subsequent reduction of GSSG by the coupled system NADPH/glutathione reductase (Fig. 4). As expected, the mutant SpLigG C15S was inactive. PcGHR1 (2  $\mu$ M) was also inactive, whereas a weak activity was detected using PcGSTO3 or PcGSTO4. In complementary experiments, comparable cysteine  $pK_a$  values around 6.3 were obtained for the different enzymes. As a consequence, the better turnover of SpLigG towards its natural substrate likely results from specific structural features of the active site.

In order to confirm this hypothesis, docking of GS-HPV to the SpLigG model was attempted. Two distinct ligands were tested in SwissDock [19], corresponding to the *R* and *S* configurations of the HPV carbon atom linked to GSH (Fig. 4). In both cases, multiple positions for the ligand in the region of the active site were obtained. Among them, a unique one fulfills the requirements for a roughly conserved position of the GSH part of the ligand and for a distance of less than 2.5 Å between the sulfur atoms of GS-HPV and Cys15. Only the *R* form of GS-HPV provides a position compatible with the reaction catalyzed at the SpLigG active site. These results are in agreement with the stereoselectivity of SpLigG previously observed by Masai et al. [6,7]. This suggests that the formation of GS-HPV catalyzed by SpLigF was accompanied by an inversion of the configuration of the asymmetric carbon of the SpLigF substrate during the nucleophilic substitution [6,7]. In the modeled complex between SpLigG and GS-HPV (Fig. S2), the HPV part of the substrate would take advantage of the hydrophilic region formed by Tyr 113, 214 and 217 and Asn223 to stabilize its carbonyl and hydroxymethyl groups linked to the asymmetric center (Fig. 4). The HPV aromatic ring, located at the dimer interface, could interact with Tyr113 through a  $\pi$ - $\pi$  stacking.

#### 4. Conclusion

SpLigG belongs to a subfamily of GSTs that possess a catalytic cysteine residue, as confirmed by its loss of activity after Cys15 mutation. Accordingly, it displays activity with the classical glutaredoxin substrates DHA and HED. Sequence analysis links SpLigG to a few proteins that form a group separated from other known GST classes. Our structural and biochemical studies demonstrate that SpLigG clearly does not belong to the xi or beta class GSTs. On the other hand, the overall three-dimensional structure of SpLigG together with its activity with phenyl-acetophenone links

it to the omega class GSTs. However, structural and sequence particularities, as well as biochemical properties specific to SpLigG, redefine the boundaries of the omega class of GSTs. Indeed, SpLigG possesses a small cavity that can provide hydrogen bonds near Cys15 and an original loop that contains the conserved motif Gly-Ala/Gly-Gly-Asn-Gly in which Asn223 belongs to the active site. Furthermore, it stereoselectively deglutathionylates  $\alpha$ -glutathionyl- $\beta$ -hydroxypropiovanillone proposed to be its main physiological substrate, and does not gain GSH transferase activity upon mutation of its catalytic cysteine to a serine residue. These properties make SpLigG a marginal omega class GST and highlight the unexpected extent of this class. On a more general note, this study demonstrates the challenge provided by GSTs regarding the prediction of their classification solely from sequence, and emphasizes the advantage of coupling structural and biochemical analyses.

#### 4.1. PDB accession number

Atomic coordinates and structure factors of SpLigG have been deposited in the Protein Data Bank under accession number 4G10.

#### Acknowledgements

We thank the staff and particularly David Cobessi at ESRF beamline FIP-BM30A for assistance with synchrotron data collection. We are grateful to Lynn Pauron, Jean-Pierre Jacquot and the reviewers for their kind help in improving the manuscript. Financial support for this project was provided by the Agence Nationale pour la Recherche (ANR-09-BLAN-0012).

#### Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.febslet.2012.09.036>.

#### References

- [1] Oakley, A. (2011) Glutathione transferases: a structural perspective. *Drug Metab. Rev.* 43, 138–151.
- [2] Board, P.G. (2011) The omega-class glutathione transferases: structure, function, and genetics. *Drug Metab. Rev.* 43, 226–235.
- [3] Zakharyan, R.A., Sampayo-Reyes, A., Healy, S.M., Tsaprilis, G., Board, P.G., Liebler, D.C. and Aposhian, H.V. (2001) Human monomethylarsonic acid (MMA(V)) reductase is a member of the glutathione-S-transferase superfamily. *Chem. Res. Toxicol.* 14, 1051–1057.
- [4] Xun, L.Y., Belchik, S.M., Xun, R., Huang, Y., Zhou, H.N., Sanchez, E., Kang, C. and Board, P.G. (2010) S-Glutathionyl-(chloro)hydroquinone reductases: a novel class of glutathione transferases. *Biochem. J.* 428, 419–427.

- [5] Dixon, D.P. and Edwards, R. (2010) Roles for stress-inducible lambda glutathione transferases in flavonoid metabolism in plants as identified by ligand fishing. *J. Biol. Chem.* 285, 36322–36329.
- [6] Masai, E., Ichimura, A., Sato, Y., Miyauchi, K., Katayama, Y. and Fukuda, M. (2003) Roles of the enantioselective glutathione S-transferases in cleavage of beta-aryl ether. *J. Bacteriol.* 185, 1768–1775.
- [7] Sato, Y. et al. (2009) Identification of three alcohol dehydrogenase genes involved in the stereospecific catabolism of arylglycerol-beta-aryl ether by *Sphingobium* sp. strain SYK-6. *Appl. Environ. Microbiol.* 75, 5195–5201.
- [8] Meux, E. et al. (2011) Glutathione transferases of *Phanerochaete chrysosporium*: S-glutathionyl-p-hydroquinone reductase belongs to a new structural class. *J. Biol. Chem.* 286, 9162–9173.
- [9] Peng, X., Masai, E., Kitayama, H., Harada, K., Katayama, Y. and Fukuda, M. (2002) Characterization of the 5-carboxyvanillate decarboxylase gene and its role in lignin-related biphenyl catabolism in *Sphingomonas paucimobilis* SYK-6. *Appl. Environ. Microbiol.* 68, 4407–4415.
- [10] Koh, C.S. et al. (2007) Crystal structures of a poplar thioredoxin peroxidase that exhibits the structure of glutathione peroxidases: insights into redox-driven conformational changes. *J. Mol. Biol.* 370, 512–529.
- [11] Kabsch, W. (2010) XDS. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* 66, 125–132.
- [12] Evans, P. (2006) Scaling and assessment of data quality. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* 62, 72–82.
- [13] Sheldrick, G.M. (2008) A short history of SHELX. *Acta Crystallogr., Sect. A* 64, 112–122.
- [14] Cowtan, K. (2006) The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* 62, 1002–1011.
- [15] Emsley, P., Lohkamp, B., Scott, W.G. and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* 66, 486–501.
- [16] Adams, P.D. et al. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* 66, 213–221.
- [17] Masai, E., Katayama, Y., Nishikawa, S., Yamasaki, M., Morohoshi, N. and Haraguchi, T. (1989) Detection and localization of a new enzyme catalyzing the beta-aryl ether cleavage in the soil bacterium (*Pseudomonas paucimobilis* SYK-6). *FEBS Lett.* 249, 348–352.
- [18] Otsuka, Y., Sonoki, T., Ikeda, S., Kajita, S., Nakamura, M. and Katayama, Y. (2003) Detection and characterization of a novel extracellular fungal enzyme that catalyzes the specific and hydrolytic cleavage of lignin guaiacylglycerol beta-aryl ether linkages. *Eur. J. Biochem.* 270, 2353–2362.
- [19] Grosdidier, A., Zoete, V. and Michielin, O. (2011) SwissDock, a protein-small molecule docking web service based on EADock DSS. *Nucleic Acids Res.* 39, W270–W277.
- [20] Armstrong, R.N. (1997) Structure, catalytic mechanism, and evolution of the glutathione transferases. *Chem Res Toxicol* 10, 2–18.
- [21] Krissinel, E. and Henrick, K. (2004) Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* 60, 2256–2268.
- [22] Holm, L. and Sander, C. (1993) Protein-structure comparison by alignment of distance matrices. *J. Mol. Biol.* 233, 123–138.
- [23] Gibrat, J.F., Madej, T. and Bryant, S.H. (1996) Surprising similarities in structure comparison. *Curr. Opin. Struct. Biol.* 6, 377–385.
- [24] Aceto, A., Dragani, B., Melino, S., Allocati, N., Masulli, M., Dillio, C. and Petruzzelli, R. (1997) Identification of an N-capping box that affects the alpha 6-helix propensity in glutathione S-transferase superfamily proteins: a role for an invariant aspartic residue. *Biochem. J.* 322, 229–234.
- [25] Cocco, R., Stenberg, G., Dragani, B., Principe, D.R., Paludi, D., Mannervik, B. and Aceto, A. (2001) 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.* 276, 32177–32183.
- [26] Dragani, B., Stenberg, G., Melino, S., Petruzzelli, R., Mannervik, B. and Aceto, A. (1997) The conserved N-capping box in the hydrophobic core of glutathione S-transferase P1-1 is essential for refolding – identification of a buried and conserved hydrogen bond important for protein stability. *J. Biol. Chem.* 272, 25518–25523.
- [27] Rossjohn, J. et al. (2000) Structures of thermolabile mutants of human glutathione transferase P1-1. *J. Mol. Biol.* 302, 295–302.
- [28] Board, P.G. et al. (2000) Identification, characterization, and crystal structure of the omega class glutathione transferases. *J. Biol. Chem.* 275, 24798–24806.
- [29] Zhou, H.N., Brock, J., Casarotto, M.G., Oakley, A.J. and Board, P.G. (2011) Novel folding and stability defects cause a deficiency of human glutathione transferase omega 1. *J. Biol. Chem.* 286, 4271–4279.
- [30] Zhou, H., Brock, J., Liu, D., Board, P.G. and Oakley, A.J. (2012) Structural insights into the dehydroascorbate reductase activity of human omega-class glutathione transferases. *J. Mol. Biol.* 420, 190–203.

$\beta_1$                        $\alpha_1$                        $\beta_2$                        $\alpha_2$                        $\beta_3$                        $\beta_4$

	1	10	20	30	40	50	60	70		
08862720	-----	MTGKHIVYHIPV	CPFSQRLEILLALRGLTDAVEFRVVDITKPRDPELLAKTRGTTALPVLETSDGRILKE						70	
AAV95330	-----	MIRFRQHDPGAKTLSGKFI	IHHIPVCPFSQRLEILLALRGLTDAVEFRVVDITRPRDPELLAKTRGTTALPVLETSDGRILKE						83	
AAZ97003	-----	MTRPTVYHIPV	CPFCQVEILLSLKGRREDVDFRMDITAPRPDWLLQKTRGTTALPVLETADGRVIKE						69	
CCA92089	-----	MAWHEEPPPGALRMYHIPG	CPFSERVEILLDLKGLSGIMDDHEVDISKPRPDWLLSKTRGTTSLPALELENGETLKE						77	
SpLigG	-----	maepqe	-----	LTIIYHIPG	CPFSERVEIMLELKGLR	-----	MKDVEIDISKPRPDWLLAKTGgTTALPLLDVEnGESLKE		70	
hgSTO1-1	msgesarslgkgsappgpV-pegs	IRIYSMRFC	CPFAERTRLVLKAKGIR	-----	HEVININLKN	-----	KPEWFFKKNP	-----	FGLVPLVLENSqGQLIYE	85
hgSTO2-2	-sgdatrtlkgksqppgpV-pegl	IRIYSMRFC	CPYSHRTRLVLKAKDIR	-----	HEVNVINLRN	-----	KPEWYYTKHP	-----	FGHIPVLETsqSQLIYE	85

$\alpha_3$                        $\alpha_4$                        $\alpha_5$

	80	90	100	110	120	130	140	150					
08862720	SLVILRYLDEALPGAPL	RRSDPTEHAIESMLIA	REGPFTMAGYLFVMN	-----	QDQ	-----	ARRDEYLNKLLALYRDLDDFLSHYNTGPFL	FE	154				
AAV95330	SLVILRYLDEVLPGAPL	RRIDPAEHAIESMLIA	KEGAFVTAGYLFVMN	-----	QDR	-----	SQREAHLDRLGLYRGLDDFLVEHNPKGTWL	FD	167				
AAZ97003	SLVILRYFEDIYREPI	AQTDYRRAVENMLTTMDRDF	VAAAGYGLWMN	-----	QDP	-----	KQDALRENMLKQYAQLDDFLLEHGAPGPF	FE	153				
CCA92089	SMVILRYFEERFPERPV	AQRDPFEHAVEAMLCATDGQFTGAGYRMLN	-----	RDR	-----	K	RDDCKAEVDAQYARLDDFLRHYS	PDGYL	FA	161			
SpLigG	SMVILRYLEQRYPEPAV	AHPDPFCHAVEGMLAELAGPFSGAGYRMLN	-----	ReI	-----	g	K	REEMRAAVDAEFKVD	AF	LkrYATgsDFL	Fd	154	
hgSTO1-1	SAITCEYLDEAYPGKKL	LPDDPYEKACQKMLELFSKVP	SLVGSFIRS	-----	Q	-----	Nke	-----	DYAGLKEEFRKEFTKLEEV	tnKKT	-----	TFFgGn	168
hgSTO2-2	SVIACEYLLDAYPGRKL	FPYDPYERARQKMLELFSKVP	PHLTKELVALrS	-----	G	-----	Re	-----	STNLKAALRQEFNSLEE	ILeyQNT	-----	TFFgGt	169

$\alpha_6$                        $\alpha_7$                        $\alpha_8$

	160	170	180	190	200	210	220	230						
08862720	D-FGLAEAVFTPMFKR	FWFLDYEGF	KL	-----	PVGKEYDRVRAWRD	ACMAHPETS	QV	-----	TEEEIVKLYDYALGAGNGALDGRSVSFAF	238				
AAV95330	G-FGLAEAVYTPFKR	FWFLDYEDF	AL	-----	PDDPAYDRVRWRAA	CIAHPATD	QV	-----	SEEEIVKLYDYALGAGNGALVDGRVTVSSFAF	251				
AAZ97003	T-FGWAETVTFPFQR	FWFLDYEDF	EL	-----	PTASRYARVREWID	ACVAHPAAQ	QT	-----	TREEVIKLYDYACGAGNGALPGRSRSSFAF	237				
CCA92089	S-FGWAEEVVFAPMFKR	LWFLEYYEGY	EV	-----	P	-----	QHLARVLRWREACTEHP	-----	AVANRHSR	ELMTLYDY	-----	SQGGGNGRLPQGRHISSTL	244	
SpLigG	dRFGWAEVAVFTPMFKR	LWFLDyedy	EV	-----	P	-----	a	-----	NFDRVLRWRAACTAHPAAQ	YR	-----	SKEElklyDYdtgggggripeg	rsissfsp	237
hgSTO1-1	-SISIMIDYLLWPFER	LLEAMK	---	L	-----	NEcvd	---	HTPKLKLWMAAMKED	PTVsaLLtSEKD	---	W	-----	Q	223
hgSTO2-2	-SISIMIDYLLWPFER	LLDVYG	---	I	-----	LdcvS	---	HTPALRLWISAMKWD	PTVsaLLmDKSI	---	F	-----	Q	224

	240	250	260		
08862720	EPPWQSRPMPPREKYAV	PATDKELGLVA	-----	266	
AAV95330	EPDWRSRPWPPRDKYD	GTASDAALGLLSV	-----	280	
AAZ97003	EPDWRARPWPPRSKYR	QPAGDTELGL	-----	263	
CCA92089	DPANEKRPMPPRDKWG	TPATDRELGLVTATVGA		277	
SpLigG	dvdwrtrpmPprdkwgha	atdaelgltr	-----	265	
hgSTO1-1	-----	Gf	lelylqnspeacygl	-----	241
hgSTO2-2	-----	Gf	lnlyfqmnpafd	-----	239

### **Supplementary Figure 1 : Sequence and structure alignment**

The top four sequences correspond to hypothetical proteins found to be SpLigG homologues, and aligned to it by ClustalW. NCBI:ZP\_08862720 is from *Ruegeria* sp. TW15; NCBI:AAV95330 from *Ruegeria pomeroyi* DSS3; NCBI:AAZ97003 from *Thiobacillus denitrificans* ATCC 25259; NCBI:CCA92089 from *Novosphingobium* sp. PP1Y; SpLigG is the product of the *Sphingobium* sp. strain SYK-6 gene NCBI:CCA92089. The bottom two sequences are derived from the PDB entries 1eem (hGSTO1-1) and 3qag (hGSTO2-2), and were aligned to SpLigG according to the three dimensional structures, using PDBeFold [2]. Aminoacid numbering indicated above the alignment corresponds to SpLigG. The secondary structures (arrows:  $\beta$ -strands; coils: helices) were determined by PDBeFold (black: SpLigG; gray: hGSTO1-1 and hGSTO2-2). The catalytic cysteine is coloured in red, residues that are conserved in all the sequences are in yellow, residues that are conserved in SpLigG homologues only are in gray, and residues that delineate the putative H-site of SpLigG are in green. The lower-case characters correspond to residues that do not superimpose in the 3-D structures. The blue box highlights the residues involved in the N-capping box [3]. The red box corresponds to the residues highlighted in red in Fig. 2.

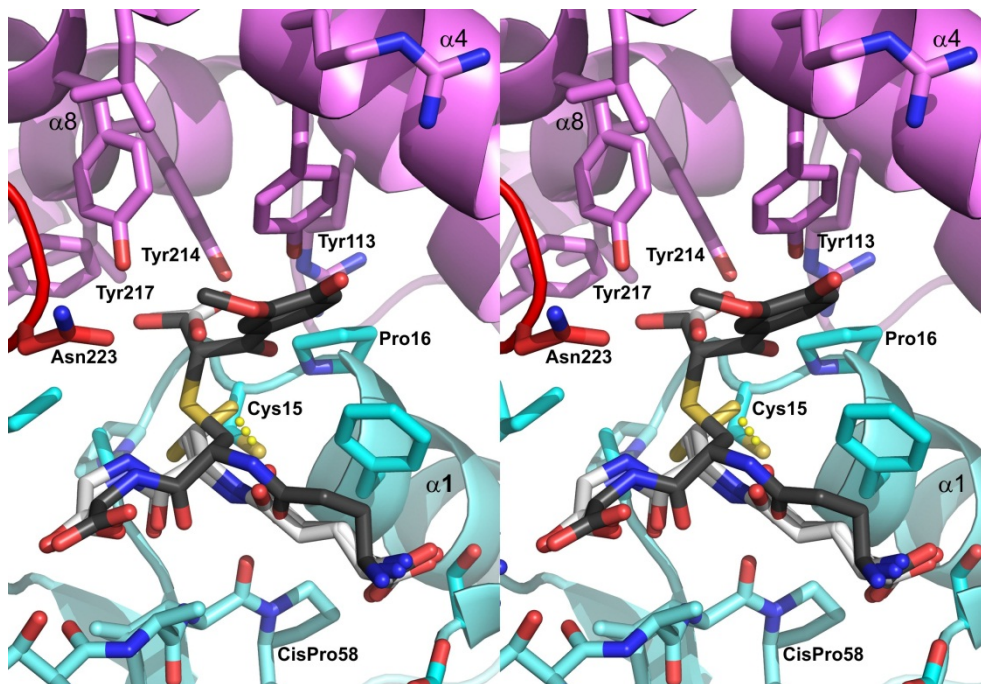
[1]: Larkin, M.A. et al. (2007). Clustal W and clustal X version 2.0. *Bioinformatics* 23, 2947-2948.

[2]: PDBeFold: Protein structure comparison service Fold at European Bioinformatics Institute (<http://www.ebi.ac.uk/msd-srv/ssm>), authored by E. Krissinel and K. Henrick [Krissinel, E. and Henrick, K. (2004). Secondary-structure matching (PDBeFold), a new tool for fast protein structure alignment in three dimensions. *Acta Cryst. D60*, 2256-2268].

[3]: Dragani, B., Stenberg, G., Melino, S., Petruzzelli, R., Mannervik, B. and Aceto, A. (1997). The conserved N-capping box in the hydrophobic core of glutathione S-transferase P1-1 is essential for refolding. *J. Biol. Chem.* 272, 25518-25523.

**Supplementary Figure 2 : GS-HPV docked in the active site of SpLigG**

The active site of SpLigG is displayed as it is shown in Fig. 4, with the same color code and details but the electron density. The substrate GS-HPV (Fig. 3, here represented in gray sticks) is docked as described in sections 2.4 and 3.4.





## C Discussion

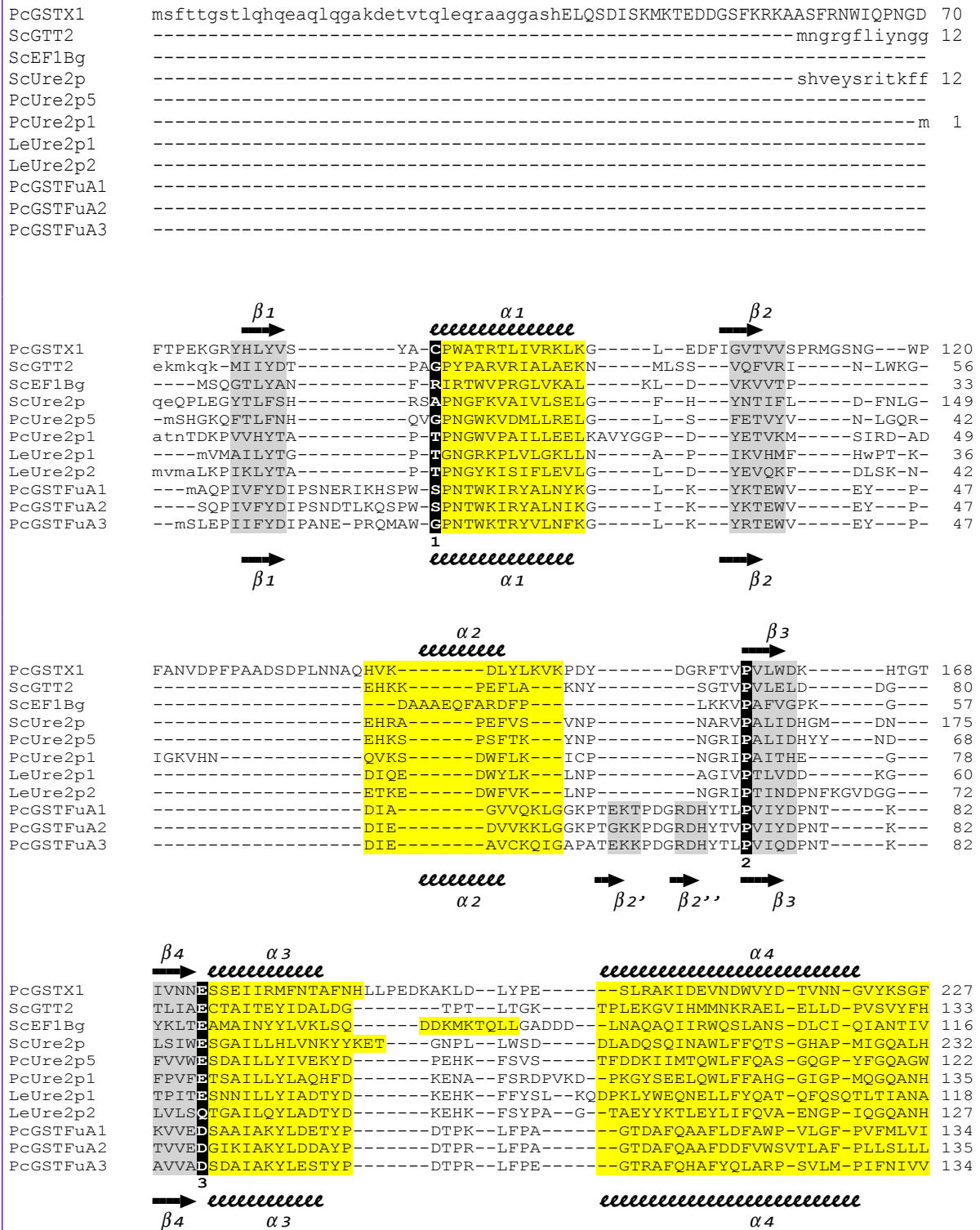
Les glutathion transférases (GST) sont principalement connues pour leur rôle dans la phase II de la détoxification cellulaire où elles catalysent l'attaque nucléophile du glutathion sur le centre électrophile de substrats xénobiotiques essentiellement hydrophobes afin de faciliter leur élimination. Depuis le début des années 1990, cette famille de protéine a fait l'objet de nombreuses études biochimiques et cristallographiques à des fins fondamentales et appliquées. Ainsi de nombreux articles de revues sont parus, (Dirr *et al.*, 1994 ; Wilce & Parker, 1994 ; Armstrong, 1997 ; Sheehan *et al.*, 2001 ; Hayes *et al.*, 2005 ; Oakley, 2005 ; Frova, 2006 ; Öztetik, 2008 ; Allocati *et al.*, 2009 ; Oakley, 2011 ; Allocati *et al.*, 2012, Board *et al.*, 2013) dont le volume 401 de « Methods in Enzymology » intitulé « Gluthione Transferases and Gamma-Glutamyl Transpeptidases » (2005). Ces travaux ont notamment mis en évidence la diversité fonctionnelle de cette famille de protéine (Oakley, 2005). Trois grandes familles sont différenciées : les GST microsomales, les GST mitochondriales et les GST cytosoliques. Nous nous sommes intéressés à cette dernière famille qui est la plus abondante. Les membres de cette famille ont été divisés en classes qui partagent une structure globale connue sous le nom de repliement GST. Une classe peut être spécifique à des organismes tout en possédant des particularités de séquences et/ou structurales (Sheehan *et al.*, 2001). Ces GST ont été également séparées suivant la position du résidu catalytique : les GST à tyrosine, les GST à sérine ou cystéine et les GST atypiques (Ma *et al.*, 2009). L'analyse des génomes nouvellement séquencés ne cessent de révéler de nouvelles classes potentielles de GST cytosoliques. Actuellement, la superfamille des GST est l'une des cibles du programme de recherche « Enzyme Function Initiative » où l'un des objectifs est d'accumuler des données bioinformatiques, structurales et biochimiques pour être capable de faire des prédictions fonctionnelles précises (site web : <http://enzymefunction.org/>, Gerlt *et al.*, 2011). De notre côté, nous nous sommes intéressés aux GST fongiques encore peu étudiées. Les GST du règne des champignons ont été divisées en sept classes : les MAK16, les EF1B $\gamma$ , les Ure2p, les Oméga, les Glutathionyl-Hydroquinone Réductases (GHR ou GST Xi), les GTT et les GSTFuA (initialement appelées GTE), (Morel *et al.*, 2009). Avant de débiter ces travaux, des études structurales et biochimiques avaient été reportées sur un membre des classes Ure2p, EF1B $\gamma$  et GTT, tous provenant de la levure *Saccharomyces cerevisiae* (Bousset *et al.*, 2001 ; Jeppesen *et al.*, 2003 ; Ma *et al.*, 2009). Dans le cadre de ce travail, nous avons réalisé les études cristallographiques d'un

membre de la classe Xi et de trois membres de la classe GSTFuA provenant du basidiomycète *P. chrysosporium*. Par ailleurs, quatre nouvelles structures de la classe Ure2p ont été déterminées récemment dans notre groupe et le programme EFI. Toutes ces structures sont listées dans le Tableau 8 et la figure 28 illustre l'alignement des séquences basé sur la superposition des structures. L'objectif de cette dernière partie est de discuter des particularités structurales des GST fongiques dans le cadre de la famille des glutathion transférases cytosoliques et en relation avec les données biochimiques obtenues notamment par nos collaborateurs de l'équipe du professeur E. Gelhaye (IAM, UMR1136).

Classe	Organisme	Nom	Code PDB / Référence
<b>Ure2p</b>			
	<i>S. cerevisiae</i>	ScUre2p	1K0A, 1K0B, 1K0C, 1K0D, 1JZR, 1G6W, 1G6Y, 1HQO (Bousset <i>et al.</i> , 2001 ;Umland <i>et al.</i> , 2001)
	<i>Lodderomyces elongisporus</i>	LeUre2p1	4EXJ
	<i>L. elongisporus</i>	LeUre2p2	4IVF
	<i>P. chrysosporium</i>	PcUre2p1	4FOB (Thuillier <i>et al.</i> , 2013)
	<i>P. chrysosporium</i>	PcUre2p5	4FOC
<b>EFBY</b>			
	<i>S. cerevisiae</i>	Sc EFB $\gamma$	1NHY (Jeppesen <i>et al.</i> , 2003)
<b>GTT</b>			
	<i>S. cerevisiae</i>	ScGTT2	3IBH (Ma <i>et al.</i> , 2009)
<b>GSTFuA</b>			
	<i>P. chrysosporium</i>	PcGSTFuA1	4F03, 4G19 (Mathieu <i>et al.</i> , 2012)
	<i>P. chrysosporium</i>	PcGSTFuA2	4LMV
	<i>P. chrysosporium</i>	PcGSTFuA3	4LMW
<b>GST Xi</b>			
	<i>P. chrysosporium</i>	PcGSTX1	3PPU (Meux <i>et al.</i> , 2011)
<b>Tableau 8 - Les GST fongiques dont la structure tridimensionnelle est déposée dans la Protein Data Bank.</b>			

**Illustration 28 - Alignement de séquences basé sur l'alignement structural des structures de GST fongiques disponibles dans la PDB.**

Les noms des différentes protéines sont explicités dans le tableau 8. Les structures secondaires de la PcGSTX1 et des PcGSTFuA sont représentés respectivement en haut et en bas de chaque bloc d'alignement en utilisant des flèches pour les brins  $\beta$  et des boucles pour les hélices. Dans chaque séquence les régions repliées sont surlignées en jaune (hélices) ou en gris (brins  $\beta$ ). Les quatre positions (1, 2, 3 et 4) mises en valeur sont surlignées en noir et coloriées en blanc.

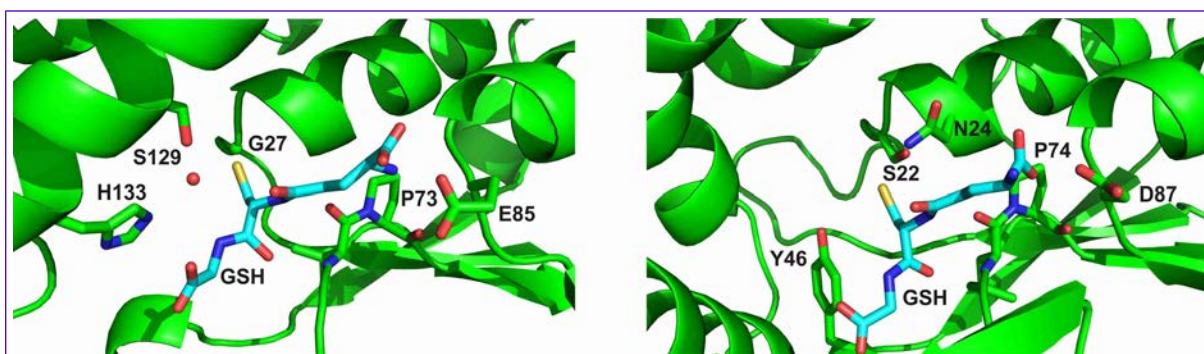


			$\alpha 5$	
			<i>ccccccccc</i>	
PcGSTX1	-----ASTQKA-----		--YEAAVIPLFES	244
ScGTT2	-----HATPGLGPEVELYQNK-----		EWGLRQRDKALHG	162
ScEF1Bg	-----PL-----		KGGAPYNKKSVDSAMDAVDKI	139
ScUre2p	-----FRYF-----		H---SQ---KIASAVERYTDEVRV	255
PcUre2p5	-----FLAV-----		APPEER---NPT--IAERYQKEILRV	147
PcUre2p1	-----FN-----		LYAPEKIPYAINRYLINESKRL	158
LeUre2p1	-----NY-----		QNG-HIDENIAQYVLSSEFKV	140
LeUre2p2	-----FV-----		FAAKEKVPYGINRYITDTKRI	150
PcGSTFuA1	<b>LDTANSLLPRSHDYFRSTREQKFGKKLEELA</b>		-----TEEWAKVEAG	176
PcGSTFuA2	<b>LDVNSLPRSSAYFRATREQQFGKRLEEQG</b>		-----GEERWQQLEAG	177
PcGSTFuA3	<b>ARVWKLRLPRSQEYFRATREQMLGKKLEELIG</b>		-----SEDDWNALESG	176
	<i>ccc</i>	<i>ccccccccc</i>	<i>ccccccccc</i>	
		$\alpha 4'$	$\alpha 5$	
			$\alpha 6$	
			<i>ccccccccc</i>	
PcGSTX1	<b>LDRLEKMLE</b> -----GQ-----		DYLIGG--QL <b>TEADIRLFVTIV</b>	275
ScGTT2	MHYFDTVLR-----ER-----		PYVAG--DSFSMA <b>DTVIAGLI</b>	193
ScEF1Bg	VDIFENRLK-----NY-----		TYLATE--NISL <b>ADLVAASIFT</b>	170
ScUre2p	<b>YGVVEMALA</b> -----ERREALVMeldtenaaysagtpmsqsrffDYVPVWLVG--		DKL <b>TIADLAFVPWNN</b>	318
PcUre2p5	<b>YGVLESVLS</b> -----QR-----		QWLVA--DKL <b>TIADISFVIWNA</b>	178
PcUre2p1	<b>YRVLDDRLLK</b> -----GR-----		EYILG--TY <b>GIADIKIFGWAR</b>	188
LeUre2p1	<b>YAFMETKLS</b> -----GR-----		DWFVG--DKF <b>TIVDIAFLVGEH</b>	171
LeUre2p2	<b>YGVFEDILSRNKANDS</b> -----		KYLVG--DRY <b>TVADFALLGWAY</b>	186
PcGSTFuA1	<b>LAKLKGYLE</b> DANGKND-----		LLMGAQGGI <b>TYSDIQIASFFV</b>	214
PcGSTFuA2	<b>LGKFKGYLE</b> RNGAGND-----		LLMGTQGGI <b>TYSDVQIASLFV</b>	215
PcGSTFuA3	<b>LARIKSSLE</b> ANGAGKD-----		LLMGD--RV <b>TFADLQLASLFI</b>	212
	<i>ccccccccc</i>		<i>ccccccccc</i>	
			$\alpha 6$	
			$\alpha 7$	
			<i>ccccccccc</i>	
			$\alpha 8$	
			<i>ccccccccc</i>	
PcGSTX1	<b>RE</b> DPVY---VTHFKC--NL----R---TIRD-GY <b>PNLHRWMRKLYWGNPAFKDTCNFEHIKTHYFWSH</b>			330
ScGTT2	<b>FA</b> -----AIV---KL----QVPEEC- <b>EALRAWYKRMQQRPS-VKLL</b> Leirskss.....			233
ScEF1Bg	<b>RYF</b> -----E---SL----FGTEWRAQ-H- <b>PAIVRWENTVRASPF-LKDEYKDFKPADKPLSPP</b>			218
ScUre2p	<b>VV</b> -----DRI---GI----NIKIEF- <b>PEVYKWKHMMRRPA-VI</b> KALrge.....			354
PcUre2p5	<b>TAV</b> -----NLL----VKGYKGF--DFEKDF- <b>PSVHRWHTALITRPA-IAESLKTKEAIAQMNR.</b>			229
PcUre2p1	<b>IAP</b> -----R---TG----L---DLDE-F- <b>PNVKAWVERIEKRPA-VQAGINSCN.</b>			224
LeUre2p1	<b>RRR</b> -----ERLHNSPIW----I---DLKENF- <b>PNVEKWFQRAIAFEN-VEEILKEHAAENLYFQsh</b>			223
LeUre2p2	<b>RLS</b> -----R---LE----I---DINQ-W- <b>PLLGKWDLSLLKLPV-VQKGFVPPKNAenlyfq</b>			231
PcGSTFuA1	<b>WAKIIWGESEK</b> -----WK----R--- <b>LISL</b> -HDGKWAQFYAQFTK----FEQVDv.....			253
PcGSTFuA2	<b>WAKVVWGESEK</b> -----WK----R--- <b>LMGF</b> -HGKWAQFCAQFAE----YERAD.....			253
PcGSTFuA3	<b>WLRVSAGESEK</b> -----WK----R--- <b>FLSL</b> -HEGKWAQFCAQFAA----YEFVDV.....			251
	<i>cc</i>	<i>ccccccccc</i>	<i>ccccccccc</i>	
		$\alpha 6'$	$\alpha 7$	
PcGSTX1	TFINPHRIVPIGPIPDILPLDa			352
ScGTT2	.....			
ScEF1Bg	Q.....			219
ScUre2p	.....			
PcUre2p5	.....			
PcUre2p1	.....			
LeUre2p1	hhhhhwhspqfek.....			236
LeUre2p2	.....			
PcGSTFuA1	.....			
PcGSTFuA2	.....			
PcGSTFuA3	.....			

## C.I Quelles sont les caractéristiques structurales communes aux GST ?

La plupart des glutathion transférases cytosoliques connues sont des protéines dimériques où chaque monomère adopte un repliement canonique appelé « repliement GST ». Ce protomère

caractéristique possède deux domaines reliés par une boucle d'une dizaine de résidus. Le domaine N-terminal adopte le repliement typique du motif thiorédoxine ( $\beta_1\alpha_1\beta_2\alpha_2\beta_3\beta_4\alpha_3$ ) composé d'un feuillet  $\beta$  avec le brin  $\beta_3$  antiparallèle aux trois autres et entouré de trois hélices. Le domaine C-terminal est un paquet d'au moins quatre hélices ( $\alpha_4\alpha_5\alpha_6\alpha_7$ ) avec la dernière hélice perpendiculaire aux trois autres. Cette description rapide de la structure tertiaire d'une GST cytosolique est le motif minimal retrouvé dans toutes les structures de GST révélées à ce jour, y compris dans celles des GST fongiques. Les GST cytosoliques sont toutefois séparées en classes principalement sur la base de leurs séquences primaires, ainsi deux GST appartiennent à la même classe si l'identité de séquence est supérieure à 40% et n'appartiennent pas à la même classe si l'identité est inférieure à 25 % (Oakley, 2011). Une classe de GST se différencie souvent des autres par une ou plusieurs originalité(s) structurale(s) commune(s) à ses membres (extensions N- ou/et C-terminales, une ou des insertions de résidus structurées ou non).

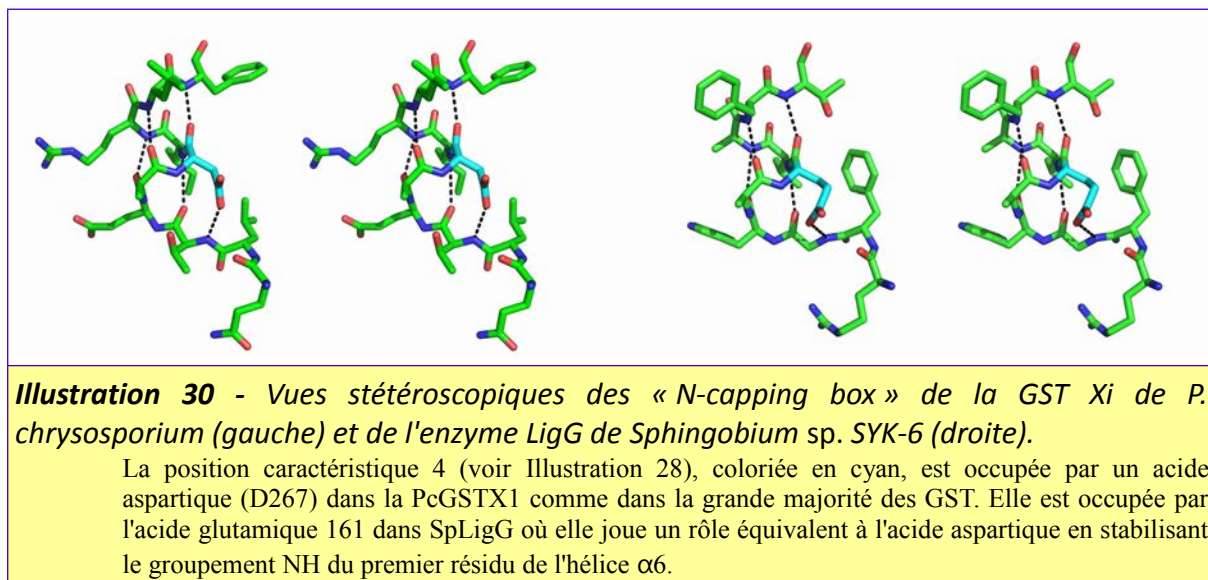


**Illustration 29** - Sites actif de la GTT2 de *S. cerevisiae* (gauche) et de la GSTFuA1 de *P. chrysosporium* (droite)

Les positions caractéristiques 1, 2 et 3 (voir Illustration 28) sont respectivement G27, P73 et E85 chez ScGTT2 et S22, P74 et D87 dans la PcGSTFuA1. Les résidus catalytiques potentiels au voisinage de l'atome de soufre du glutathion ont été représentés (H133 et S129 pour la ScGTT2 et S22, N24 et Y46 pour la PcGSTFuA1).

De nombreuses études ont tenté de déterminer les facteurs moléculaires communs à toutes les GST. La position et la nature du résidu catalytique, si la GST est une enzyme, sont relativement variables et seront discutées dans la dernière partie de cette discussion. Seul un résidu est invariant. Il s'agit d'une proline en conformation *cis* localisée à la fin de la boucle  $\alpha_2\beta_3$  (position 1, Illustration 28 et Illustration 29). Le résidu précédent cette proline reçoit la chaîne principale de la cystéine du glutathion. Même si la mutation de cette proline affecte la stabilité et la réactivité de l'enzyme (Allocati *et al.*, 1999), son rôle structural précis n'a pas vraiment été établi (Atkinson & Babbitt, 2009). Une deuxième position dans le domaine N-terminal mérite attention ; elle est majoritairement occupée par un glutamate ou une glutamine, et constitue le point d'ancrage du groupement amine N-terminal du glutathion (position 2, Illustration 28 et Illustration 29). De plus il

a été supposé que ce résidu joue un rôle dans l'activation du groupement thiol du glutathion via un réseau de partage d'électrons ( Winayanuwattikun & Ketterman, 2007). Dans toutes les structures des GST fongiques connues, cette deuxième position est occupée par un glutamate sauf dans les GSTFuA où un aspartate joue le même rôle (Illustration 29) (Mathieu *et al.*, 2012).



Dans le domaine C-terminal, un résidu aspartate situé dans le premier tour de l'hélice  $\alpha_6$  semblait, jusqu'à nos études, être invariant chez les GST. Il est impliqué dans la stabilisation du cœur hydrophobe de l'enzyme (Dragani *et al.*, 1997). Il fait partie d'un motif structural, de signature S/TXXD et appelé motif « N-capping box », retrouvé dans de nombreuses structures de protéines (Richardson & Richardson, 1988) car important pour la stabilisation des groupements amines n'ayant pas d'accepteur de liaison hydrogène puisque situés au niveau N-terminal d'une hélice alpha. La première position du motif est le premier résidu de l'hélice. Les chaînes latérales de la sérine ou théronine et de l'aspartate forment des liaisons hydrogène, respectivement avec les groupements NH du quatrième et premier résidus du motif. Comme dans les autres GST, la première position du motif n'est pas toujours occupée par une sérine ou une thréonine dans les GST fongiques (Illustration 28). Dragani *et al.* (1997) ont montré que la mutation de l'acide aspartique en alanine affectait la stabilité de la protéine. L'analyse des structures des GST fongiques nous a montré que cet acide aspartique n'est finalement que « quasi-invariant » car nous avons remarqué que dans quatre structures de GST il est remplacé par un acide glutamique, une thréonine ou une glycine. Dans la protéine LigG de *Sphingobium* sp. SYK 6 dont nous avons étudié la structure, il est remplacé par un acide glutamique qui joue le même rôle que l'aspartate quasi invariant (Illustration 30) (Meux *et al.*, 2012). Il en est de même lorsqu'il est remplacé par une thréonine (structure de la GST yibF d'*E. coli*, code pdb 3R2Q). Dans les deux derniers cas l'acide aspartique est remplacé par

une glycine (structures de GST de *Pseudomonas fluorescens*, codes PDB 3LXT et 3MOF). Dans ces deux protéines, la deuxième position du motif N-capping box est occupée par un acide glutamique et la position après la glycine par une sérine qui interagissent ensemble assurant ainsi la stabilisation de la région N-terminale de l'hélice  $\alpha_6$ .

## **C.II Quelles sont les caractéristiques structurales distinguant les GST ?**

### **C.II.1 La structure du monomère**

Un alignement structural a été réalisé à partir de toutes les structures de GST non redondantes et déposées dans la PDB (voir détails en Annexe, partie E). Cet alignement avait pour but de créer un arbre de similarité afin de regrouper les GST à la fois d'après leur structure tridimensionnelle (base de l'alignement des séquences) et d'après leur similitude (base de la construction de l'arbre) (Illustration 31). Pareillement à ce qui a été montré récemment (Atkinson & Babitt, 2009 ; Stourman *et al.*, 2011) uniquement sur la base des structures primaires, notre arbre regroupe toutes les GST à tyrosine dans une branche où l'on retrouve la majorité des GST de mammifères regroupées classe par classe (Alpha, Mu et Pi). Cette branche de GST à tyrosine possède tout de même plusieurs exceptions non classées et dont la nature du résidu catalytique n'a pas été clairement déterminée à ce jour (codes PDB : 3lsz, 4hz4, 2fno). Parmi ces trois structures, seule celle de la GST Atu5508 d'*Agrobacterium tumefaciens* a été décrite dans la littérature (Kosloff *et al.*, 2006). Les auteurs mettent effectivement en évidence une proximité structurale de cette GST avec celles de la famille des GST à tyrosine. La nature du résidu catalytique proposée est une arginine se trouvant dans le premier tour de l'hélice  $\alpha_1$ . Il est intéressant de noter qu'une autre GST d'*Agrobacterium tumefaciens* (classe Eta, code PDB 2ycd) a été récemment caractérisée d'un point de vue structural et biochimique, démontrant que le résidu catalytique est une arginine (Skopelitou *et al.*, 2012) toujours localisée dans le premier tour de l'hélice  $\alpha_1$ . Cette dernière se trouve dans une branche peu éloignée du groupe des GST à tyrosine (Illustration 31).

#### **Illustration 31 - Arbre de similarité des structures de GST.**

Le protocole de la construction de l'arbre est détaillé en Annexe E.

Lorsque plusieurs GST appartenant à une même classe se regroupent pour présenter une racine commune, les différents représentants ne sont plus listés, à la place un triangle coloré en gris figure l'ensemble du groupe.

Les structures de GST de *P. chrysosporium* déterminées au CRM2 sont représentées en rouge, de même que la nouvelle classe Xi identifiée grâce à la résolution de la structure de PcGSTX1 et comportant actuellement 4 membres.

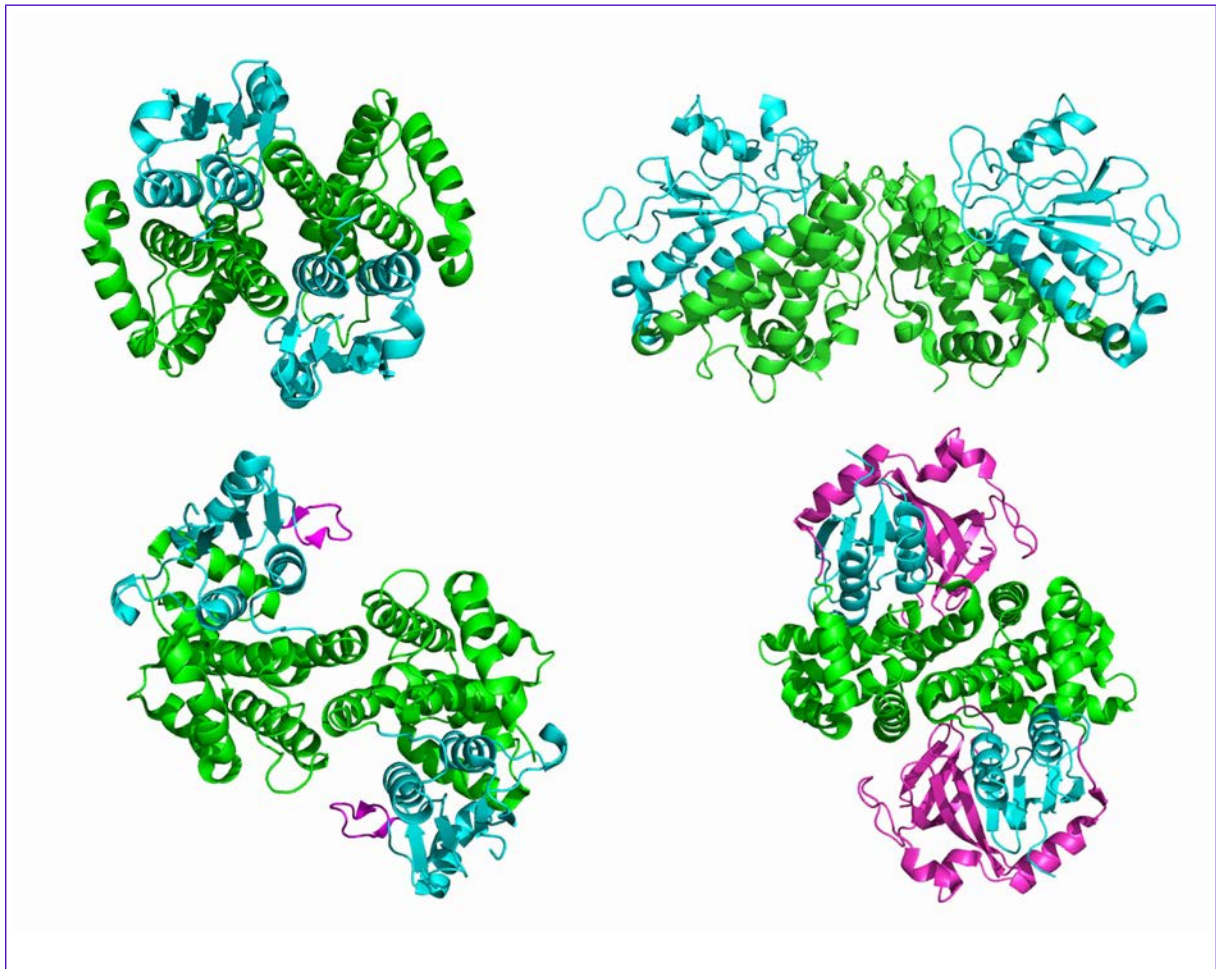




Les autres branches de l'arbre de similarité regroupent la plupart des protéines des classes connues de GST à sérine ou à cystéine et de GST sans activité transférase ou de déglutathionylation telles que les CLIC. C'est dans cette partie de l'arbre que se trouvent toutes les structures des GST fongiques déposées dans la PDB et c'est également dans cette partie qu'il y a le plus grand nombre de GST esseulées (Illustration 31). Bon nombre sont non-classées car les structures proviennent de programmes de la génomique structurale qui publient rarement les analyses structurales de leurs cibles. Cet arbre met en exergue à quel point la diversité structurale des GST est importante. Même si la plupart des GST non classées n'ont pas été caractérisées d'un point de vue biologique, il faut s'attendre à la découverte de nouvelles fonctions pour cette famille de protéines. De façon intéressante, l'arbre de similarité construit semble en accord avec l'ensemble des observations émanant des comparaisons structurales que nous avons réalisées. La protéine LigG (code PDB 4g10) et les GST Oméga sont proches dans l'arbre comme elles le sont d'un point de vue structural et biochimique (Meux *et al.*, 2012). Même si la GSTFuA (code PDB 4g19) est très particulière, celle-ci a comme plus proche voisin la gutarédoxine 2 (codes PDB 1g7o et 3ir4) avec qui elle partage la particularité structurale d'avoir une hélice supplémentaire entre les hélices  $\alpha_4$  et  $\alpha_5$  (Mathieu *et al.*, 2012). La PcGSTX1 se regroupe bien avec les autres glutathionyl-hydroquinone réductase (GHR) en formant une nouvelle classe structurale de GST à cystéine (GST Xi) qui possède notamment comme spécificités structurales des extensions N- et C-terminales et une élongation de la boucle  $\beta_2\alpha_2$  (Meux *et al.*, 2011 ; Green *et al.*, 2012 ; Lam *et al.*, 2012). Pour finir, les Ure2p fongiques sont relativement bien rassemblées, celles-ci sont notamment proches des GST YfcG et YghU d'*E. coli* caractérisées récemment (Wadington *et al.*, 2009 ; Stourmann *et al.*, 2011 ; codes pdb 3gx0 et 3c8e). Un travail publié cette année par notre groupe (ne faisant pas partie de cette thèse, Thuillier *et al.*, 2013) expose l'analyse cristallographique et biochimique d'une Ure2p de *P. chrysosporium* qui est homologue des protéines YfcG et YghU. Ces enzymes qui présentent une activité de déglutathionylation sont originales car elles fixent deux molécules de glutathion par monomère (Wadington *et al.*, 2009). L'Ure2p de *S. cerevisiae* reste particulière dans cette classe puisque c'est la seule à posséder une extension N-terminale impliquée dans la formation de fibrilles amyloïdes et à avoir une longue extension de la boucle  $\alpha_5\alpha_6$  (Bousset *et al.*, 2001, Morel *et al.*, 2009). De nombreux membres de la classe Ure2p sont présents dans le tableau de similarité car cette classe est l'une des cibles du programme de recherche de génomique structurale EFI (Gerlt *et al.*, 2011). Ils ont en particulier déposés récemment deux nouveaux membres provenant de la levure *Lodderomyces elongisporus* (Tableau 8 et Illustration 31). La plupart des GST de la classe Ure2p ont en commun une arginine stabilisant la deuxième molécule de glutathion (sachant que la première se fixe dans le site classique G).

## C.II.2 La structure quaternaire

Les GST cytosoliques sont pour la plupart des protéines dimériques, seules quelques classes sont monomériques comme les GST Lambda et les DéHydroAscorbate Réductases (DHAR) qui sont des GST à cystéine spécifiques aux plantes (Dixon *et al.*, 2002). Quand nous avons commencé les études structurales des GST Xi et GSTFuA de *P. chrysosporium*, un seul type de dimère avait été reporté dans la littérature pour les GST cytosoliques ; nous l'appellerons dimère canonique. Ce dimère possède une symétrie axiale du type C2 où le domaine N-terminal du premier monomère est en face du domaine tout hélice de l'autre monomère (Illustration 32). Nous avons été relativement surpris de découvrir deux nouveaux types de dimères pour la PcGSTX1 et la PcGSTFuA1 (Meux *et al.*, 2011 ; Mathieu *et al.*, 2013) (Illustration 32). Sur la base des particularités structurales de ces deux enzymes, nous avons pu expliquer pourquoi elles ne peuvent pas former le dimère canonique des GST. Alors que pour la PcGSTX1, le nouveau dimère est totalement original, dans le cas de la PcGSTFuA1, l'arrangement est voisin de celui observé dans le dimère classique, les deux protomères sont translétés dans le plan de l'interface de façon à rapprocher les domaines tout hélice. Le motif épingle à cheveux  $\beta_2'\beta_2''$  observé pour la première fois dans une GST empêche la formation du dimère classique des GST. Une comparaison structurale réalisée par ailleurs nous a permis de découvrir qu'une GST de *Ralstonia solanacearum* (bactérie du sol pathogène pour les plantes) non caractérisée d'un point de vue biochimique s'homodimérise comme les GSTFuA (code pdb 4KF9, provenant du programme EFI, Gerlt *et al.*, 2011). De façon intéressante cette GST possède notamment l'hélice supplémentaire entre les hélices  $\alpha_4$  et  $\alpha_5$  comme les GSTFuA. Par contre, elle ne possède pas le motif  $\beta_2'\beta_2''$  mais à la place une longue extension C-terminale qui allonge le feuillet  $\beta$  de la structure (Illustration 32). Pour conclure, il semble évident que la diversité des structures quaternaires des GST ne pourra que s'accroître et quelle n'est absolument pas spécifique au règne des champignons. Pour preuves, une GST de *Leishmania* a récemment été découverte trimérique (Fyfe *et al.*, 2012) ; comme montré au paragraphe B.IV.2, la protéine SpLigF forme un tétramère qui peut être décrit comme l'association de deux dimères classiques. Toutefois, il semble également évident que la grande majorité des GST adopte la structure quaternaire canonique.



**Illustration 32 - Vues de 4 modes de dimérisation des GST.**

Dans chaque cas, deux monomères (un à gauche et l'autre à droite) sont représentés dans le mode rubans et flèches. Les domaines N- et C-terminaux sont coloriés respectivement en cyan et vert. Des régions particulières sont coloriées en magenta.

En haut à gauche : mode de dimérisation le plus courant des GST dans lequel le domaine N-terminal interagit avec le domaine C-terminal de l'autre monomère. Le dimère représenté est celui de la GTT2 de *S. cerevisiae*.

En haut à droite : dimère des GST Xi où les deux monomères interagissent via le domaine C-terminal de chaque monomère. Le dimère représenté est celui de la GSTX1 de *P. chrysosporium*.

En bas à gauche : dimère des GSTFuA où les deux monomères interagissent via le domaine C-terminal de chaque monomère. Le motif épingle à cheveux  $\beta_2'\beta_2''$  (en magenta) empêche la formation du dimère classique. Le dimère représenté est celui de la GSTFuA1 de *P. chrysosporium*.

En bas à droite : dimère d'une GST de la bactérie du sol *Ralstonia solanacearum*. Ce dimère est similaire de celui des GSTFuA où cette fois c'est l'extension C-terminale (en magenta) qui empêche la formation du dimère classique.

Les cristallographes travaillant sur la superfamille des GST ont tenté de comprendre si la forme dimérique était nécessaire à l'efficacité catalytique. Une crevasse potentiellement importante avait été remarquée au centre du dimère, accessible au solvant et reliant les deux sites actifs (Reinemer *et al.*, 1996). Très rapidement, les études structurales ont mis en évidence une grande diversité au niveau de l'interface du dimère qui s'est révélée hydrophobe (GST à tyrosine) ou hydrophile (GST Thêta, Oméga ou Bêta). Les résidus au niveau de l'interface ne s'avèrent pas strictement conservés dans certaines classes comme dans la classe Thêta ou les facteurs

d'élongation EF1B $\gamma$  (Prade *et al.*, 1997 ; Jeppesen *et al.*, 2003). De plus l'étendue de l'interface est très variable passant de 2000 Å<sup>2</sup> pour la GST Oméga 1 humaine à presque 3600 Å<sup>2</sup> pour l'Ure2p1 de *P. chrysosporium* (Board *et al.*, 2000 ; Thuillier *et al.*, 2013). Hegazy *et al.* (2004) ont montré que la 'clé' (Y50) du motif 'clé serrure' situé à l'interface du dimère était essentielle pour l'efficacité de la GST Pi humaine (Illustration 3). Par ailleurs, Huang *et al.* (2008) ont réussi à déplacer l'équilibre monomère/dimère de cette même GST Pi humaine vers la forme monomérique en modifiant des résidus à l'interface. Ils ont caractérisé une forme monomérique conservant une activité substantielle. L'ensemble de ces résultats semble montrer que la nature dimérique de la majorité des GST est certes importante mais pas essentielle pour les activités spécifiques glutathion transférase ou de déglutathionylation. Probablement, la majorité des GST forment des homo-oligomères pour les mêmes raisons que plus de 50% des protéines dans les organismes (MacKinnon *et al.*, 2013). Godsell & Olson (2000) notent en particulier deux forces opposées poussant les protéines à s'oligomériser. D'un côté, les systèmes de grandes tailles semblent plus adaptés pour assurer les fonctions des protéines et d'un autre côté, la machinerie moléculaire qui synthétise les protéines tend à limiter la longueur des chaînes polypeptidiques.

### **C.II.3 Les résidus permettant la catalyse**

Caractériser une enzyme implique la détermination des facteurs moléculaires responsables de l'activité catalytique. Dans le cas des GST qui présentent une activité avec les substrats classiques des glutarédoxines (HED, DHA) et qui possèdent une cystéine à la même position que la cystéine catalytique des thiorédoxines ou glutarédoxines (c'est-à-dire au début de l'hélice  $\alpha_1$ ), la nature du résidu catalytique est évidente. Le mécanisme catalytique comprend un intermédiaire covalent où le glutathion forme un pont disulfure avec la cystéine catalytique. Ainsi, l'activité est perdue lorsque la cystéine est remplacée par une alanine dans tous les cas, comme cela a été montré dans le cas de la GST Oméga Humaine (hGSTO1, Board *et al.*, 2000) ou par nos collaborateurs pour la GST Xi (PcGSTX1 ou GHR) de *P. chrysosporium* (Meux *et al.*, 2011) et LigG de *Sphingobium* sp SYK-6 (Meux *et al.*, 2012). De façon évidente, d'autres résidus sont importants pour la catalyse. Par exemple dans la structure de la PcGSTXi, nous avons remarqué trois tyrosines potentiellement intéressantes car d'une part à proximité du groupement thiol du glutathion et d'autre part conservés dans les GST Xi. A partir d'études de mutagenèse dirigée, Green *et al.* (2012) ont proposé récemment un mécanisme catalytique impliquant ces trois tyrosines qui pourraient fournir le proton nécessaire lors de la déglutathionylation. D'autres enzymes appartenant à la famille structurale des GST présentent une activité catalytique du type glutarédoxine. Elles ne possèdent pas de cystéine catalytique contrairement aux GST Oméga ou Xi. La première à avoir été caractérisée est l'Ure2p de *S. cerevisiae* qui possède une alanine à la place de la cystéine (Zhang &

Perret, 2009). Deux résidus plus loin, une asparagine (N124) a été supposé potentiellement importante pour la catalyse sur la base d'études biochimiques antérieures (Zhang *et al.*, 2008). Plus tard le groupe d'Armstrong a montré que les enzymes YfcG et YghU d'*E. coli* sont également actives avec les substrats classiques de la glutarédoxine en fixant deux molécules de glutathion dans leur site actif (Wadington *et al.*, 2009 ; Stourman *et al.*, 2011). Ces deux enzymes possèdent, à la place de la cystéine des GST Oméga ou Xi, une thréonine qui a été supposée importante pour la catalyse du fait de sa proximité avec les atomes soufre des deux molécules de glutathion. Même si YghU et YfcG présentent une certaine homologie de séquence avec l'Ure2p de *S. cerevisiae* (respectivement 35% et 37 %), celles-ci ont été définies dans une nouvelle classe appelée Nu. Comme cette nouvelle classe avait été auparavant définie pour des GST de nématodes (Schuller *et al.*, 2005 ; Asojo *et al.*, 2007), YghU et YfcG ont été regroupés dans la classe des Ure2p (Thuillier *et al.*, 2013, étude réalisée sur l'Ure2p1 de *P. chrysosporium* dans le groupe ne faisant pas partie de ce travail). L'Ure2p1 de *P. chrysosporium* présentent les mêmes propriétés biochimiques que YghU et YfcG. Il est intéressant de noter que la mutation de l'asparagine 18 (équivalente à l'asparagine 124 de l'Ure2p de *S. cerevisiae*) conduit à une forme inactive et que cette asparagine est présente dans YghU et YfcG. Cette dernière partie du paragraphe montre un nouveau point de diversité des GST sur la nature des facteurs moléculaires responsables d'une même activité. Cette diversité qui s'avère encore plus importante pour l'activité glutathion transférase est discutée ci-dessous.

Pour les GST possédant l'activité glutathion transférase, les études biochimiques et cristallographiques montrent que la plupart utilisent une tyrosine (GST à tyrosine) ou une sérine (GST à sérine) pour stabiliser la cystéine du glutathion mais aussi pour diminuer son pKa (Armstrong *et al.*, 1997 ; Board *et al.*, 2003). La tyrosine est localisée à la fin du brin  $\beta_1$  alors que la sérine se trouve à la même position que le résidu catalytique des GST à cystéines. La GTT2 de *S. cerevisiae* est la première GST fongique présentant l'activité classique des GST à avoir été caractérisée d'un point de vue structural et biochimique (Ma *et al.*, 2009). Celle-ci est originale car elle ne possède ni tyrosine et ni sérine aux positions attendues. La structure cristallographique ne révèle aucune chaîne latérale dans le voisinage direct de l'atome de soufre du glutathion. Seule une molécule d'eau pontage son interaction avec une histidine (H133) et une sérine (S129), situées au niveau C-terminal de l'hélice  $\alpha_4$  (Illustration 29). Ces deux résidus semblent importants pour la catalyse puisque l'enzyme devient faiblement active ou inactive lorsqu'ils sont remplacés par une alanine (Ma *et al.*, 2009). Le cas de la GTT2 n'est pas isolé puisque récemment Skopelitou *et al.* (2012) ont mis en évidence une nouvelle classe de GST appelée Eta dont l'activité catalytique serait assurée par une arginine et un réseau de partage d'électrons. Ce dernier implique en particulier le glutamate (ou glutamine) stabilisant l'extrémité N-terminale du glutathion (position 3 sur

l'illustration 28). Ce réseau est supposé jouer le rôle d'assistance basique pour activer la cystéine du glutathion (Winayanuwattikun & Ketterman, 2005 ; Winayanuwattikun & Ketterman, 2007). Concernant notre étude sur la PcGSTFuA1, sa structure en présence de glutathion montre clairement que la sérine 22, située à la position attendue des GST à sérine, forme une liaison hydrogène avec le groupement thiol du glutathion (Mathieu *et al.*, 2012) (Illustration 29). Cette enzyme catalyse efficacement l'addition du glutathion sur les isothiocyanates. Par contre elle ne présente pas d'activité avec le CDNB qui est le substrat classique des GST. Comme attendu la mutation de la sérine en alanine rend l'enzyme inactive vis-à-vis de l'isothiocyanate, mais de façon surprenante, cela permet à l'enzyme d'accepter le CDNB comme substrat et même de présenter une activité peroxydase. L'analyse structurale que nous avons réalisée nous a permis de mettre en évidence d'autres résidus potentiels capables de remplacer la sérine, telles que l'asparagine 24 et la tyrosine 46 dont la caractérisation biochimique est en cours (Illustration 29). Il est intéressant de noter que ces deux résidus sont conservés dans la PcGSTFuA3 qui présente une activité peroxydase et où la position de la sérine 'catalytique' est occupée par une glycine (Mathieu *et al.*, 2013). Il est toutefois difficile de prédire l'activité puisque la PcGSTFuA2, qui est homologue à la PcGSTFuA1 d'un point de vue séquence, est inactive avec les substrats testés. Même s'il est possible que les GSTFuA possèdent un substrat commun non identifié à ce jour, ces observations montrent une certaine diversité fonctionnelle de la classe GSTFuA dont les membres partagent la même structure.

## **D Conclusion et perspectives**

Ce manuscrit présente la caractérisation cristallographique de deux nouvelles classes de glutathion transférases (GST) : la classe structurale Xi avec comme modèle d'étude la *S*-glutathionyl-(chloro)hydroquinone réductase (GHR) du basidiomycète saprophyte *Phanerochaete chrysosporium* et la classe GSTFuA (pour Fungal specific class A) avec comme modèles d'étude les isoformes 1, 2 et 3 de *P. chrysosporium* (PcGSTFuA1, PcGSTFuA2 et PcGSTFuA3). Il semblait judicieux d'étudier les GST fongiques encore peu décrites dans la littérature par rapport à celles provenant des règnes animal, végétal et bactérien. Le séquençage massif d'organismes fongiques avait permis de révéler une distribution importante de GST dans les champignons saprophytes (Morel *et al.*, 2009). L'homologie entre les GSTFuA et l'enzyme LigE de *Sphingobium* sp. SYK-6 paraissait potentiellement très intéressante car le système d'enzymes LigE, F et G est connue pour sa capacité à dégrader de petits fragments de lignine (Masai *et al.*, 2003). Les études cristallographiques de ces trois Lig de *Sphingobium* sp. SYK-6 (SpLigE, SpLigF et Sp LigG) ont été entreprises d'abord parce que les GSTFuA ne semblent pas posséder les propriétés catalytiques de LigE mais aussi pour le potentiel applicatif des Lig en biotechnologie. Seule la structure de LigG a pu être caractérisée à haute résolution, un modèle en cours d'affinement est présenté pour LigF et les données de diffraction obtenues pour LigE n'ont pas permis la résolution de sa structure.

L'étude cristallographique de la *S*-glutathionyl-(chloro)hydroquinone réductase (GHR) nous a permis de découvrir une nouvelle classe structurale de GST à cystéine que nous avons appelée GST Xi. Cette nouvelle classe présente des éléments structuraux originaux avec en particulier une longue extrémité N-terminale, un allongement de la boucle reliant  $\beta_2$  à  $\alpha_2$ , une hélice additionnelle  $\alpha_9$  suivie d'une boucle C-terminale, et un nouveau mode de dimérisation. Aujourd'hui quatre structures de cette nouvelle classe sont présentes dans la PDB dont une non caractérisée d'un point vue biochimique (code PDB 3M1G). Les trois autres (PcGSTX1, la protéine YqjG d'*E. coli*, et PcpF de *S. chlorophenolicum*) possèdent des profils d'activité proches, en particulier une activité GHR démontrée. Pour deux de ces enzymes (PcGSTX1 dans la présente étude, et EcYqjG (Green *et al.*, 2012)), la structure d'un complexe avec le substrat GS-ménadione a été obtenue. Des observations très proches réalisées au niveau des sites actifs permettent l'identification des acides aminés définissant le site H, de nature très hydrophobe, ainsi que d'acides aminés intervenants probablement dans l'acte catalytique. Toutefois, dans les deux cas, la qualité des densités électroniques au niveau du site actif ne permet pas une définition précise des facteurs moléculaires

mis en jeu pour fixer le substrat. L'obtention d'une structure avec un intermédiaire covalent permettrait peut-être de travailler avec un échantillon plus homogène au niveau conformationnel et ainsi d'obtenir une définition précise des relations entre la structure et la fonction de cette nouvelle classe de GST.

L'ubiquité de la classe GSTFuA dans les basidiomycètes saprophytes implique que cette famille de protéines possède un rôle physiologique dans ces organismes. Nos études cristallographiques sur plusieurs GSTFuA de *P. chrysosporium* font apparaître une faible diversité structurale dans cette nouvelle classe alors que cette diversité est importante en termes de spécificité de substrat et de spécificité de ligand. Toutefois, les GSTFuA possèdent une singularité structurale dans la superfamille des GST, à savoir une épingle à cheveux ( $\beta_2'\beta_2''$ ) entre l'hélice  $\alpha_2$  et le brin  $\beta_3$  qui empêche la formation du dimère canonique des GST d'une part et qui d'autre part couvre le site de fixation du glutathion. Par ailleurs, la présence d'une hélice supplémentaire ( $\alpha_4'$ ) entre les hélices  $\alpha_4$  et  $\alpha_5$  est remarquable car celle-ci, très rarement observée dans une GST, est présente dans la glutarédoxine 2 qui est considérée comme l'ancêtre commun des GST. La découverte d'un site anionique au cœur du site actif des GSTFuA1 a permis de s'assurer que l'acide 8-Anilino-1-naphtalène sulfonique était une sonde fluorescente adéquate pour effectuer des tests de compétition avec des ligands. Ces analyses font apparaître une autre singularité dans la superfamille des GST, car il semblerait que les GSTFuA utilisent le site de fixation du glutathion plutôt que le site H pour fixer des molécules essentiellement hydrophobes issues de la dégradation du bois. Malgré les nombreux essais réalisés à la fois en cocrystallisation et par diffusion de ligands dans les cristaux avec l'enzyme de type sauvage ou modifiée, nous n'avons malheureusement obtenu qu'une seule forme complexée, à savoir la PcGSTFuA1 avec le glutathion. Les ligands restent difficiles à manipuler car peu solubles dans l'eau. L'obtention d'une structure d'une GSTFuA avec un ligand hydrophobe permettrait de vérifier les hypothèses émises sur la base des structures résolues et des données en solution.

Les GST du système Lig de *Sphingobium* sp. SYK-6 sont respectivement les Lig E, F et G appartenant au cluster de gènes ligDFEG (Masai *et al.*, 2003). Dans la cascade de réactions étudiées par Masai et collaborateurs sur un modèle de lignine dimérique, les protéines LigE et LigF catalysent stéréosélectivement la première étape de la coupure du lien éther de l' $\alpha$ -(2-méthoxyphénoxy)- $\beta$ -hydroxypropiovanillone (MHPV) pour produire la  $\beta$ -hydroxypropiovanillone glutathionylée (GS-HPV) et le gaïacol (voir Illustration 19). La protéine LigG a été démontrée stéréospécifique du produit de la réaction catalysée par LigF. Par contre à ce jour, aucune protéine n'a encore été identifiée chez *Sphingobium* sp. SYK-6 qui prenne en charge le produit de la première étape catalysée par LigE. L'étude des relations entre la structure et la fonction de SpLigG révèle une enzyme appartenant à la famille des GST à cystéines qui peut être considérée comme



marginale dans la classe Oméga car l'homologie de séquence reste très faible nonobstant des structures et des profils d'activité globalement proches. Cela dit, les légères différences structurales permettent d'apporter des hypothèses sur la spécificité de SpLigG pour le substrat GS-HPV. Concernant SpLigE et SpLigF les études cristallographiques sont préliminaires. Les premiers résultats structuraux sur SpLigF laissent présager des propriétés structurales uniques notamment en terme d'oligomérisation. La détermination d'une structure en présence de glutathion permettrait peut-être de stabiliser les régions désordonnées non observées, comme cela a été montré récemment dans le cas de la protéine Ure2p6 de *P. chryso sporium* (C. Didierjean, communication personnelle). L'obtention de structures à haute résolution des LigE et F reste nécessaire pour comprendre la stéréosélectivité de ces enzymes vis-à-vis de leur substrat. En conclusion, les premiers résultats structuraux obtenus sur les Lig montrent à nouveau l'extrême modularité de la superfamille des GST, tant d'un point de vue fonctionnel que structural.



## **E Annexe : l'arbre de similarité des structures de GST**

Les paragraphes suivants décrivent la procédure qui a été utilisée pour réaliser l'arbre de similarité des structures de GST présenté dans l'illustration 31 de la partie « Discussion ».

En bref, l'ensemble des structures de GST a été sélectionné. Pour chacune de ces GST, une recherche d'homologie de séquences a été effectuée sur la PDB de sorte à identifier d'éventuelles structures déjà résolues pour la même protéine. Lorsque plusieurs structures d'une même GST existent, une seule d'entre elles a été retenue. Le texte qui suit précise les critères de sa sélection. Le jeu des GST retenues est donc un jeu « non redondant ». Les structures de ces GST ont été superposées, et l'alignement des séquences qui en découle a ensuite été utilisé pour construire un arbre sur la base des similarités observées dans cet alignement.

### **E.I Recherche des structures de GST dans la Protein Data Bank**

Toutes les entrées de la Protein Data Bank ayant « Glutathione transferase » ou « GST » dans les champs « Structure Title » ou « Molecule Name » ont été retenues (405 entrées). Les GST MAPEG et Kappa n'ayant pas le repliement canonique des GST cytosoliques ont été éliminées (9 entrées), de même que 2 GST de la classe Mu dont seules les coordonnées des atomes de la chaîne principale sont listées dans le fichier PDB, et une GST correspondant à une protéine chimère entre deux GST de la classe Alpha.

La base de données des structures de GST contient donc 392 entrées (voir Tableau 9).

### **E.II Constitution d'une base de données non redondante de structures de GST**

Pour chacune des 392 entrées de GST retenues, une recherche des séquences homologues de la PDB a été réalisée à l'aide de BlastP (Altschul *et al.*, 1990). Cette étape a permis d'une part de regrouper sans ambiguïté les entrées de la PDB correspondant à différentes structures d'une même protéine (un seuil de 97 % d'identité a permis de tenir compte des mutations), d'autre part d'assigner à chaque entrée la classe d'appartenance de la protéine quand celle-ci n'était pas indiquée par les auteurs de la structure (en utilisant un seuil de 50 % d'identité avec des GST de classe connue).

Lorsque plusieurs structures d'une même protéine existent dans la PDB, une seule d'entre elles a été choisie comme représentant. Les critères utilisés, sont, par ordre de priorité :

- privilégier les entrées pour lesquels la chaîne polypeptidique est la plus complète possible. Au sein même d'une entrée, retenir la chaîne pour laquelle les éventuelles parties manquantes sont plutôt aux extrémités N et C terminales et non au milieu de la chaîne. Retenir la chaîne pour laquelle, si un « tag » a été inséré dans la construction pour faciliter la purification, ce « tag » est si possible absent de la structure.
- privilégier les formes non mutées
- privilégier les formes avec GSH prioritairement, et sinon les formes avec GSH-substrat, et en cas de choix les substrats qui n'engendrent pas de changement conformationnel important (ce qui est vérifié par superposition des structures)
- privilégier la plus haute résolution

La base de données des structures de GST non redondantes contient 145 entrées.

### **E.III Alignement structural des entrées de la base de données non redondante des GST**

Le programme Strap (Gile & Frommel, 2001) a été utilisé pour superposer deux à deux les 145 structures tridimensionnelles précédemment sélectionnées. L'alignement des séquences des 145 GST est ensuite généré à partir de cette superposition structurale. Compte tenu des nombreuses différences dans la nature et la position des singularités des classes (voire sous-groupes) de GST, cet alignement est très étiré, les parties bien alignées correspondant pour la très grande majorité aux éléments de structures secondaires du repliement canonique.

### **E.IV Construction de l'arbre de similarité des séquences de GST alignées selon la superposition structurale.**

L'alignement des séquences des GST sur la base de leur superposition structurale est utilisé pour générer la matrice (145x145) des distances entre ces séquences protéiques. Le programme « Protldist » de la suite Phylip 3.7 (Felsenstein, 1989) disponible sur la plateforme Mobylye 1.5 de l'Institut Pasteur (<http://mobylye.pasteur.fr/cgi-bin/portal.py>) a été utilisé, en choisissant pour modèle de substitution des acides aminés la matrice « Dayhoff PAM », et en réalisant une procédure « bootstrap » de 100 répétitions.

L'arbre est ensuite généré à partir de la matrice précédente avec le programme Quicktree (Howe *et al.*, 2002) disponible sur la plateforme Mobylye 1.5 de l'Institut Pasteur. Sa représentation

est réalisée avec le logiciel FigTree v1.4.0 (Auteur : Andrew Rambaut, Institute of Evolutionary Biology, University of Edinburgh).

**Tableau 9 - Bilan des structures de GST déposées dans la Protein Data Bank**

*Les entrées (listées par leur code PDB) ont été sélectionnées selon le protocole présenté au paragraphe E.I.*

*Lorsqu'aucune mention de la classe d'appartenance ne figurait dans l'entrée PDB ou la publication associée, une classe a pu être indiquée dans certains cas (voir paragraphe E.II), un point d'interrogation est indiqué en colonne 2 sinon.*

*Les organismes dont sont issus les GST sont dans la très grande majorité des cas indiqués par un code 5 lettres, correspondant aux 3 premières du nom de genre et aux 2 premières de l'épithète de spécificité de l'espèce. Exceptionnellement d'autres notations sont utilisées, telles que « human » pour Homo sapiens, « rat » pour Ratus norvegicus, « mouse » pour Mus musculus, « coli » pour E. coli, ou encore « ralstonia » pour ralstonia sp..*

*Les colonnes Ligand 1 et Ligand 2 donnent les codes PDB des ligands importants éventuellement présents dans le (ou proche du) site actif. D'autres ligands non listés ici peuvent être présents ailleurs dans la structure.*

*L'année listée est celle de mise à disposition des coordonnées dans le domaine publique.*

*La colonne « modèle retenu » donne le code PDB et le numéro de chaîne du monomère sélectionné comme représentant de chaque entrée.*

*Lorsqu'une entrée n'a pas de publication associée décrivant le résidu catalytique, une attribution a si possible été réalisée par superposition avec les structures de résidus catalytiques connus, un point d'interrogation a été mentionné sinon.*

Code PDB	Classe	Organisme	Ligand1	Ligand2	Année	Résol. (Å)	Modèle retenu	Résidu catalyt.
<a href="#">4HZ4</a>	?	actpl	-	-	2012	1,62	4HZ4_A	?
<a href="#">2FNO</a>	?	agrtu	-	-	2006	2	2FNO_A	Tyr
<a href="#">3H1N</a>	?	borbr	-	-	2009	1,83	3H1N_A	Tyr
<a href="#">4IEL</a>	?	buram	GSH	-	2012	1,6	4IEL_B	Ser
<a href="#">4ISD</a>	?	burgl	GSH	-	2013	2,65	4ISD_B	Ser
<a href="#">4KDU</a>	?	burgr	-	-	2013	1,6	4KDX_A	Ser
<a href="#">4KDX</a>	?	burgr	GSH	-	2013	1,35	4KDX_A	Ser
<a href="#">4KE3</a>	?	burgr	-	-	2013	1,9	4KDX_A	Ser
<a href="#">3BBY</a>	?	coli	-	ACT	2007	1,85	3BBY_A	Ser
<a href="#">3C8E</a>	?	coli	GSH	-	2008	1,5	3C8E_A	?
<a href="#">3GX0</a>	?	coli	GDS	-	2009	2,3	3GX0_A	?
<a href="#">3R2Q</a>	?	coli	GSH	EDO	2011	1,05	3R2Q_A	Ser
<a href="#">4JBB</a>	?	klepn	GSH	FMT	2013	1,48	4JBB_A	Ser
<a href="#">4AGS</a>	?	leiin	GSH	EDO	2012	2,3	4AGS_A	Cys
<a href="#">3UBK</a>	?	lepin	-	GOL	2011	1,95	3UBL_A	Ser

Code PDB	Classe	Organisme	Ligand1	Ligand2	Année	Résol. (Å)	Modèle retenu	Résidu catalyt.
<a href="#">3UBL</a>	?	lepin	GTT	GOL	2011	2	3UBL_A	Ser
<a href="#">4EXJ</a>	?	lodel	-	-	2013	1,64	4EXJ_B	?
<a href="#">4IVF</a>	?	lodel	GSH	-	2013	2,2	4IVF_A	?
<a href="#">4GLT</a>	?	metfl	GSH	-	2012	2,2	4GLT_B	Ser
<a href="#">4JED</a>	?	metra	GTS	GOL	2013	1,1	4JDE_A	?
<a href="#">3W8S</a>	?	necam	GSH	-	2013	2,07	3W8S_A	Tyr
<a href="#">4HOJ</a>	?	neigo	GSH	-	2012	1,4	4HOJ_A	Cys
<a href="#">4ECI</a>	?	pseae	-	ACT	2012	1,8	4ECJ_A	Asn?
<a href="#">4ECJ</a>	?	pseae	GSH	-	2012	1,76	4ECJ_A	Asn?
<a href="#">3LXT</a>	?	psepr	-	CL	2010	1,76	3M0F_A	Ser
<a href="#">3M0F</a>	?	psepr	GSH	-	2010	1,6	3M0F_A	Ser
<a href="#">3M3M</a>	?	psepr	GTT	-	2010	1,75	3M3M_A	Ser
<a href="#">4IBP</a>	?	psepr	GSH	-	2012	1,8	4IBP_A	Ser
<a href="#">4ID0</a>	?	psepr	GSF	ACT	2012	1,1	4IBP_A	Ser
<a href="#">4IJI</a>	?	psepr	99T	AKR	2012	1,5	4IBP_A	Ser
<a href="#">4IKH</a>	?	psepr	GSH	-	2012	2,1	4IKH_A	Asn
<a href="#">3LXZ</a>	?	psepu	-	-	2010	1,76	3PR8_A	Ser
<a href="#">3PR8</a>	?	psepu	GSH	EPE	2010	1,8	3PR8_A	Ser
<a href="#">3IC8</a>	?	psey	-	-	2009	2,4	3IC8_A	Ser
<a href="#">3CBU</a>	?	raleu	-	EDO	2008	2,05	3CBU_B	Ser
<a href="#">3TOT</a>	?	ralso	-	ACT	2011	1,76	3TOU_A	Ser
<a href="#">3TOU</a>	?	ralso	GSH	-	2011	1,75	3TOU_A	Ser
<a href="#">4KF9</a>	?	ralso	GSH	-	2013	2,3	4KF9_A	?
<a href="#">3M8N</a>	?	rhopa	-	-	2010	2,04	3M8N_D	Ser
<a href="#">3LSZ</a>	?	rhosp	GSH	-	2010	1,7	3LSZ_C	Arg / Ser ?
<a href="#">4KH7</a>	?	salen	GSH	-	2013	1,5	4KH7_A	Ser
<a href="#">4G10</a>	?	sphpa	GSH	-	2012	1,2	4G10_A	Cys
<a href="#">4GF0</a>	?	sulfi	GSH	-	2012	1,75	4GF0_A	?
<a href="#">4HZ2</a>	?	xanau	GSH	BEZ	2012	1,5	4HZ2_B	Ser
<a href="#">1VF1</a>	Alpha	galga	GSH	-	2004	1,77	1VF1_A	Tyr
<a href="#">1VF2</a>	Alpha	galga	-	GTX	2004	2,15	1VF1_A	Tyr
<a href="#">1VF3</a>	Alpha	galga	-	GDN	2004	2,15	1VF1_A	Tyr
<a href="#">1VF4</a>	Alpha	galga	-	-	2004	2,45	1VF1_A	Tyr
<a href="#">1GSD</a>	Alpha	human	-	-	1995	2,5	1K3Y_A	Tyr
<a href="#">1GSE</a>	Alpha	human	GTT	EAA	1995	2	1K3Y_A	Tyr
<a href="#">1GSF</a>	Alpha	human	-	EAA	1995	2,7	1K3Y_A	Tyr

Code PDB	Classe	Organisme	Ligand1	Ligand2	Année	Résol. (Å)	Modèle retenu	Résidu catalyt.
<a href="#">1GUH</a>	Alpha	human	-	GSB	1993	2,6	1K3Y_A	Tyr
<a href="#">1GUL</a>	Alpha	human	-	I	1998	2,7	3IK7_A	Tyr
<a href="#">1GUM</a>	Alpha	human	-	-	1998	3,5	3IK7_A	Tyr
<a href="#">1K3L</a>	Alpha	human	-	GTX	2001	1,5	1K3Y_A	Tyr
<a href="#">1K3O</a>	Alpha	human	-	-	2001	1,8	1K3Y_A	Tyr
<a href="#">1K3Y</a>	Alpha	human	-	GTX	2001	1,3	1K3Y_A	Tyr
<a href="#">1PKW</a>	Alpha	human	GTT	-	2003	2	1K3Y_A	Tyr
<a href="#">1PKZ</a>	Alpha	human	-	-	2003	2,1	1K3Y_A	Tyr
<a href="#">1PL1</a>	Alpha	human	-	ABY	2003	1,75	1K3Y_A	Tyr
<a href="#">1PL2</a>	Alpha	human	-	ABY	2003	1,8	1K3Y_A	Tyr
<a href="#">1TDI</a>	Alpha	human	GSH	-	2004	2,4	1TDI_A	Tyr
<a href="#">1USB</a>	Alpha	human	GSH	-	2003	2,07	1K3Y_A	Tyr
<a href="#">1XWG</a>	Alpha	human	-	-	2004	1,85	1K3Y_A	Tyr
<a href="#">1YDK</a>	Alpha	human	-	GTX	2004	1,95	1K3Y_A	Tyr
<a href="#">2R3X</a>	Alpha	human	-	GTX	2007	1,8	1K3Y_A	Tyr
<a href="#">2R6K</a>	Alpha	human	-	GTX	2007	2,51	1K3Y_A	Tyr
<a href="#">2VCT</a>	Alpha	human	-	ASD	2007	2,1	2WJU_A	Tyr
<a href="#">2VCV</a>	Alpha	human	-	GSW	2007	1,8	1TDI_A	Tyr
<a href="#">2WJU</a>	Alpha	human	-	GSW	2009	2,3	2WJU_A	Tyr
<a href="#">2WJU</a>	Alpha	human	-	GSW	2009	2,3	2WJU_A	Tyr
<a href="#">3I69</a>	Alpha	human	GTT	-	2009	2,38	-	Tyr
<a href="#">3I6A</a>	Alpha	human	GTT	-	2009	1,98	-	Tyr
<a href="#">3IK7</a>	Alpha	human	-	BOB	2009	1,97	3IK7_A	Tyr
<a href="#">3IK9</a>	Alpha	human	-	BOB	2009	2,2	-	Tyr
<a href="#">3KTL</a>	Alpha	human	-	GTX	2009	1,75	1K3Y_A	Tyr
<a href="#">3LOH</a>	Alpha	human	-	GTX	2009	2,13	1K3Y_A	Tyr
<a href="#">3Q74</a>	Alpha	human	-	-	2011	1,79	1K3Y_A	Tyr
<a href="#">3U6V</a>	Alpha	human	-	-	2011	2,2	1K3Y_A	Tyr
<a href="#">3ZFB</a>	Alpha	human	-	-	2012	1,86	1K3Y_A	Tyr
<a href="#">3ZFL</a>	Alpha	human	-	-	2012	1,88	1K3Y_A	Tyr
<a href="#">4ACS</a>	Alpha	human	GSH	-	2011	2,1	2WJU_A	Tyr
<a href="#">4HJ2</a>	Alpha	human	-	LZ6	2012	2,1	1K3Y_A	Tyr
<a href="#">1B48</a>	Alpha	mouse	GSH	HAG	1998	2,6	1B48_B	Tyr
<a href="#">1F3A</a>	Alpha	mouse	GSH	-	2000	1,9	1F3A_A	Tyr
<a href="#">1F3B</a>	Alpha	mouse	-	GBX	2000	2	1F3A_A	Tyr
<a href="#">1GUK</a>	Alpha	mouse	-	-	1997	2,9	1B48_B	Tyr

Code PDB	Classe	Organisme	Ligand1	Ligand2	Année	Résol. (Å)	Modèle retenu	Résidu catalyt.
<a href="#">1ML6</a>	Alpha	mouse	-	GBX	2002	1,9	1ML6_A	Tyr
<a href="#">1EV4</a>	Alpha	rat	GTS	-	2000	2,2	1EV4_A	Tyr
<a href="#">1EV9</a>	Alpha	rat	GTS	-	2000	2,2	1EV4_A	Tyr
<a href="#">2DSA</a>	Beta	burxe	GTT	HPX	2006	2,1	2GDR_A	Cys
<a href="#">2GDR</a>	Beta	burxe	GTT	-	2006	2,1	2GDR_A	Cys
<a href="#">1AOF</a>	Beta	coli	GTS	-	1997	2,1	1N2A_A	Cys
<a href="#">1N2A</a>	Beta	coli	GTS	-	2002	1,9	1N2A_A	Cys
<a href="#">4IQ1</a>	Beta	manha	-	-	2013	1,85	4IW9_A	Ser
<a href="#">4IW9</a>	Beta	manha	GSH	-	2013	1,76	4IW9_A	Cys
<a href="#">3UAP</a>	Beta	metca	-	-	2011	2,8	3UAR_A	Cys
<a href="#">3UAR</a>	Beta	metca	GSH	-	2011	2,6	3UAR_A	Cys
<a href="#">2NTO</a>	Beta	ochan	GTT	-	2006	2,1	2NTO_A	Cys
<a href="#">2PVQ</a>	Beta	ochan	GTT	-	2007	1,8	2NTO_A	Cys
<a href="#">1PMT</a>	Beta	promi	GTT	-	1998	2,5	1PMT_A	Cys
<a href="#">2PMT</a>	Beta	promi	GTT	-	1998	2,7	1PMT_A	Cys
<a href="#">4KGI</a>	Beta	shifl	GSH	-	2013	1,6	4KGI_A	Cys
<a href="#">1F2E</a>	Beta	sphpa	GTT	-	2000	2,3	1F2E_A	Cys
<a href="#">2X64</a>	Beta	xylfa	GTT	-	2010	2,3	2X64_A	Cys
<a href="#">4G9H</a>	Beta	yerpe	GSH	-	2012	2,1	4GCI_A	Cys
<a href="#">4GCI</a>	Beta	yerpe	GSH	-	2012	1,5	4GCI_A	Cys
<a href="#">2YV9</a>	CLIC	caeel	-	-	2007	1,6	2YV9_A	Cys
<a href="#">2YV7</a>	CLIC	drome	-	-	2007	1,7	2YV7_A	Cys
<a href="#">1K0M</a>	CLIC	human	-	-	2001	1,4	3O3T_A	Cys
<a href="#">1K0N</a>	CLIC	human	GTT	-	2001	1,8	3O3T_A	Cys
<a href="#">1K0O</a>	CLIC	human	-	-	2001	1,75	3O3T_A	Cys
<a href="#">1RK4</a>	CLIC	human	-	-	2003	1,79	3O3T_A	Cys
<a href="#">2AHE</a>	CLIC	human	-	-	2005	1,8	2D2Z_B	Cys
<a href="#">2D2Z</a>	CLIC	human	-	-	2005	2,2	2D2Z_B	Cys
<a href="#">2PER</a>	CLIC	human	-	MNB	2007	2	2R5G_A	Cys
<a href="#">2R4V</a>	CLIC	human	GSH	-	2007	1,85	2R5G_A	Cys
<a href="#">2R5G</a>	CLIC	human	-	-	2007	1,86	2R5G_A	Cys
<a href="#">3FY7</a>	CLIC	human	-	-	2009	1,95	3KJY_A	Cys
<a href="#">3KJY</a>	CLIC	human	-	-	2009	1,95	3KJY_A	Cys
<a href="#">3O3T</a>	CLIC	human	-	-	2010	1,7	3O3T_A	Cys
<a href="#">3P8W</a>	CLIC	human	-	-	2010	2	3O3T_A	Cys
<a href="#">3P90</a>	CLIC	human	-	-	2010	2,3	3O3T_A	Cys



Code PDB	Classe	Organisme	Ligand1	Ligand2	Année	Résol. (Å)	Modèle retenu	Résidu catalyt.
<a href="#">3QR6</a>	CLIC	human	-	-	2011	1,78	3O3T_A	Cys
<a href="#">3SWL</a>	CLIC	human	-	-	2011	2,35	3O3T_A	Cys
<a href="#">3TGZ</a>	CLIC	human	-	-	2011	2,3	3O3T_A	Cys
<a href="#">3UVH</a>	CLIC	human	-	-	2011	1,84	3O3T_A	Cys
<a href="#">4IQA</a>	CLIC	human	-	-	2013	2,49	3O3T_A	Cys
<a href="#">1JLV</a>	Delta	anocr	GSH	-	2001	1,75	1JLV_A	Ser
<a href="#">1JLW</a>	Delta	anocr	-	-	2001	2,45	3F63_A	Ser
<a href="#">1R5A</a>	Delta	anocr	GTS	-	2003	2,5	1R5A_A	Ser
<a href="#">1V2A</a>	Delta	anocr	GTS	-	2003	2,15	1V2A_B	Ser
<a href="#">3F63</a>	Delta	anocr	-	GTX	2008	1,8	3F63_A	Ser
<a href="#">3F6D</a>	Delta	anocr	-	GTX	2008	1,7	3F63_A	Ser
<a href="#">3G7I</a>	Delta	anocr	-	GSW	2009	2,05	3F63_A	Ser
<a href="#">3G7J</a>	Delta	anocr	-	GTX	2009	2,2	3F63_A	Ser
<a href="#">1PN9</a>	Delta	anoga	-	GTX	2003	2	1PN9_A	Ser
<a href="#">3AY8</a>	Delta	bommo	-	-	2011	2,1	3AY8_A	Ser
<a href="#">3VK9</a>	Delta	bommo	-	-	2011	2	3VK9_A	Ser
<a href="#">4E8E</a>	Delta	bommo	-	-	2012	2,51	4E8H_B	Ser
<a href="#">4E8H</a>	Delta	bommo	GSH	-	2012	2,12	4E8H_B	Ser
<a href="#">3EIN</a>	Delta	drome	GTT	-	2008	1,13	3EIN_A	Ser
<a href="#">3F6F</a>	Delta	drome	-	-	2008	1,6	3F6F_A	Ser
<a href="#">3GH6</a>	Delta	drome	GTT	-	2009	1,65	3F6F_A	Ser
<a href="#">3MAK</a>	Delta	drome	GTT	-	2010	1,8	3EIN_A	Ser
<a href="#">1NHY</a>	EF1	sacce	-	-	2002	3	1NHY_A	-
<a href="#">2IL3</a>	Epsilon	anoga	-	-	2006	2,2	2IMI_A	Ser
<a href="#">2IMI</a>	Epsilon	anoga	GSH	-	2006	1,4	2IMI_A	Ser
<a href="#">2IMK</a>	Epsilon	anoga	-	GTX	2006	1,9	2IMI_A	Ser
<a href="#">4HI7</a>	Epsilon	dromo	GSH	-	2012	1,25	4HI7_A	Ser
<a href="#">3VWX</a>	Epsilon	musdo	GSH	-	2012	1,8	3VWX_C	Ser
<a href="#">2YCD</a>	Eta	agrtu	GTB	-	2011	1,4	2YCD_A	Arg
<a href="#">3LQ7</a>	Eta	agrtu	-	-	2010	2,3	2YCD_A	Arg
<a href="#">1G7O</a>	GRX2	coli	-	-	2000	nmr	1G7O_A	Cys
<a href="#">3IR4</a>	GRX2	salty	GTT	-	2009	1,2	3IR4_A	Cys
<a href="#">4F03</a>	GSTFuA	phach	-	-	2012	1,8	4G19_A	Ser
<a href="#">4G19</a>	GSTFuA	phach	GSH	-	2012	2	4G19_A	Ser
<a href="#">3ERF</a>	GTT2	sacce	-	-	2008	2,23	3IBH_A	?
<a href="#">3ERG</a>	GTT2	sacce	GTS	-	2008	2,2	3IBH_A	?

Code PDB	Classe	Organisme	Ligand1	Ligand2	Année	Résol. (Å)	Modèle retenu	Résidu catalyt.
<a href="#">3IBH</a>	GTT2	sacce	GTT	-	2009	2,1	3IBH_A	?
<a href="#">3ISO</a>	Mu	closi	GSH	-	2009	1,9	3ISO_A	Tyr
<a href="#">1FHE</a>	Mu	fashe	-	GSW	1997	3	1FHE_A	Tyr
<a href="#">2FHE</a>	Mu	fashe	GTT	-	1998	2,3	2FHE_A	Tyr
<a href="#">2WRT</a>	Mu	fashe	-	-	2009	2,4	2WRT_A	Tyr
<a href="#">1C72</a>	Mu	galga	-	EPY	2000	2,8	1GSU_A	Tyr
<a href="#">1GSU</a>	Mu	galga	-	GTX	1997	1,94	1GSU_A	Tyr
<a href="#">1GTU</a>	Mu	human	-	-	1998	2,68	1XW6_A	Tyr
<a href="#">1HNA</a>	Mu	human	-	GDN	1993	1,85	1XW5_A	Tyr
<a href="#">1HNB</a>	Mu	human	-	GDN	1993	3,5	1XW5_A	Tyr
<a href="#">1HNC</a>	Mu	human	-	GDN	1993	3,25	1XW5_A	Tyr
<a href="#">1XW5</a>	Mu	human	GSH	-	2004	1,8	1XW5_A	Tyr
<a href="#">1XW6</a>	Mu	human	GSH	-	2004	1,9	1XW6_A	Tyr
<a href="#">1XWK</a>	Mu	human	-	GDN	2004	2,3	1XW6_A	Tyr
<a href="#">1YJ6</a>	Mu	human	GSH	-	2005	2,5	1XW6_A	Tyr
<a href="#">1YKC</a>	Mu	human	-	GDS	2005	2,1	1XW5_A	Tyr
<a href="#">2AB6</a>	Mu	human	-	GSM	2005	2,5	1XW5_A	Tyr
<a href="#">2C4J</a>	Mu	human	-	GSO	2005	1,35	1XW5_A	Tyr
<a href="#">2F3M</a>	Mu	human	-	GTD	2005	2,7	1XW6_A	Tyr
<a href="#">2GTU</a>	Mu	human	-	-	1998	2,55	1XW5_A	Tyr
<a href="#">3GTU</a>	Mu	human	-	-	1998	2,8	3GTU_A	Tyr
<a href="#">3GUR</a>	Mu	human	GTT	BYG	2009	2,5	1XW5_A	Tyr
<a href="#">4GTU</a>	Mu	human	-	-	1999	3,3	4GTU_A	Tyr
<a href="#">2DC5</a>	Mu	mouse	-	-	2005	1,6	2DC5_A	Tyr
<a href="#">1MTC</a>	Mu	rat	-	GPR	2002	2,2	6GST_A	Tyr
<a href="#">2GST</a>	Mu	rat	-	GPS	1993	1,8	6GST_A	Tyr
<a href="#">3FYG</a>	Mu	rat	-	GPR	1997	2,2	6GST_A	Tyr
<a href="#">3GST</a>	Mu	rat	-	GPR	1993	1,9	6GST_A	Tyr
<a href="#">4GST</a>	Mu	rat	-	GTD	1993	1,9	6GST_A	Tyr
<a href="#">5FWG</a>	Mu	rat	-	GPR	1997	2	6GST_A	Tyr
<a href="#">5GST</a>	Mu	rat	-	GDN	1993	2	6GST_A	Tyr
<a href="#">6GST</a>	Mu	rat	GTT	-	1996	2,2	6GST_A	Tyr
<a href="#">6GSU</a>	Mu	rat	-	GPS	1996	1,85	6GST_A	Tyr
<a href="#">6GSV</a>	Mu	rat	-	GPS	1996	1,75	6GST_A	Tyr
<a href="#">6GSW</a>	Mu	rat	-	GPS	1996	1,85	6GST_A	Tyr
<a href="#">6GSX</a>	Mu	rat	-	GPS	1996	1,91	6GST_A	Tyr

Code PDB	Classe	Organisme	Ligand1	Ligand2	Année	Résol. (Å)	Modèle retenu	Résidu catalyt.
<a href="#">6GSY</a>	Mu	rat	GTT	-	1996	2,2	6GST_A	Tyr
<a href="#">1GNE</a>	Mu	schja	-	GSW	1994	2,5	1UA5_A	Tyr
<a href="#">1GTA</a>	Mu	schja	-	-	1994	2,4	1UA5_A	Tyr
<a href="#">1GTB</a>	Mu	schja	-	-	1994	2,6	1UA5_A	Tyr
<a href="#">1M99</a>	Mu	schja	GTS	-	2002	2,3	1UA5_A	Tyr
<a href="#">1M9A</a>	Mu	schja	-	GTX	2002	2,1	1UA5_A	Tyr
<a href="#">1M9B</a>	Mu	schja	-	IBG	2002	2,6	1UA5_A	Tyr
<a href="#">1U87</a>	Mu	schja	-	GSW	2004	3,5	1UA5_A	Tyr
<a href="#">1U88</a>	Mu	schja	-	GTY	2004	3,5	1UA5_A	Tyr
<a href="#">1UA5</a>	Mu	schja	GTT	-	2003	2,5	1UA5_A	Tyr
<a href="#">1Y6E</a>	Mu	schja	-	-	2004	3	1UA5_A	Tyr
<a href="#">3CRT</a>	Mu	schja	GSH	-	2008	1,9	1UA5_A	Tyr
<a href="#">3CRU</a>	Mu	schja	GSH	-	2008	2,3	1UA5_A	Tyr
<a href="#">3D0Z</a>	Mu	schja	GSH	-	2008	2,5	1UA5_A	Tyr
<a href="#">1OKT</a>	Mu'	plafa	-	-	2003	1,9	1Q4J_A	Tyr
<a href="#">1PA3</a>	Mu'	plafa	-	-	2003	2,7	1Q4J_A	Tyr
<a href="#">1Q4J</a>	Mu'	plafa	-	GTX	2003	2,2	1Q4J_A	Tyr
<a href="#">2AAW</a>	Mu'	plafa	-	GTX	2005	2,4	1Q4J_A	Tyr
<a href="#">3FR3</a>	Mu'	plafa	GDS	-	2009	1,9	1Q4J_A	Tyr
<a href="#">3FR6</a>	Mu'	plafa	-	-	2009	2,6	1Q4J_A	Tyr
<a href="#">3FR9</a>	Mu'	plafa	-	GSW	2009	2,4	1Q4J_A	Tyr
<a href="#">3FRC</a>	Mu'	plafa	-	OHG	2009	2	1Q4J_A	Tyr
<a href="#">1YQ1</a>	Nu	caeel	-	-	2005	3	1YQ1_A	Tyr
<a href="#">1ZL9</a>	Nu	caeel	GTT	-	2005	2,01	1ZL9_A	Tyr
<a href="#">2WS2</a>	Nu	haeco	-	-	2009	2,01	2WS2_A	Tyr
<a href="#">1TW9</a>	Nu	helpo	-	-	2004	1,71	1TW9_F	Tyr
<a href="#">2ON5</a>	Nu	necam	GTT	-	2007	1,9	2ON5_A	Tyr
<a href="#">2ON7</a>	Nu	necam	-	-	2007	2,4	2ON7_A	Tyr
<a href="#">3RBT</a>	Omega	bommo	-	-	2011	2,2	3RBT_D	Cys
<a href="#">1EEM</a>	Omega	human	GSH	-	2000	2	1EEM_A	Cys
<a href="#">3LFL</a>	Omega	human	GTT	DTT	2010	2,1	1EEM_A	Cys
<a href="#">3Q18</a>	Omega	human	-	-	2010	1,7	3QAG_A	Cys
<a href="#">3Q19</a>	Omega	human	GSH	-	2010	1,9	3QAG_A	Cys
<a href="#">3QAG</a>	Omega	human	GSH	-	2011	2	3QAG_A	Cys
3VLN	Omega	human	-	-	2011	1,7	1EEM_A	Cys
<a href="#">4IS0</a>	Omega	human	GDS	1R4	2013	1,72	1EEM_A	Cys

Code PDB	Classe	Organisme	Ligand1	Ligand2	Année	Résol. (Å)	Modèle retenu	Résidu catalyt.
<a href="#">1IYH</a>	PGDS	human	GSH	-	2002	1,7	2VCQ_A	Tyr
<a href="#">1IYI</a>	PGDS	human	GSH	-	2002	1,8	2VCQ_A	Tyr
<a href="#">1V40</a>	PGDS	human	GSH	O16	2003	1,9	2VCQ_A	Tyr
<a href="#">2CVD</a>	PGDS	human	GSH	HQL	2005	1,45	2VCQ_A	Tyr
<a href="#">2VCQ</a>	PGDS	human	GSH	D25	2007	1,95	2VCQ_A	Tyr
<a href="#">2VCW</a>	PGDS	human	GSH	ZZA	2007	1,95	2VCQ_A	Tyr
<a href="#">2VCX</a>	PGDS	human	GSH	D26	2007	2,1	2VCQ_A	Tyr
<a href="#">2VCZ</a>	PGDS	human	GSH	VC3	2007	1,95	2VCQ_A	Tyr
<a href="#">2VDO</a>	PGDS	human	GSH	D27	2007	2,2	2VCQ_A	Tyr
<a href="#">2VD1</a>	PGDS	human	GSH	D28	2007	2,25	2VCQ_A	Tyr
<a href="#">3EE2</a>	PGDS	human	GTT	NZO	2008	1,91	2VCQ_A	Tyr
<a href="#">3KXO</a>	PGDS	human	-	GSW	2009	2,1	2VCQ_A	Tyr
<a href="#">3VI5</a>	PGDS	human	GSH	M4M	2011	2	2VCQ_A	Tyr
<a href="#">3VI7</a>	PGDS	human	GSH	-	2011	2	2VCQ_A	Tyr
<a href="#">4ECO</a>	PGDS	human	GSH	7PQ	2012	1,85	2VCQ_A	Tyr
<a href="#">4EDY</a>	PGDS	human	GSH	9PQ	2012	1,72	2VCQ_A	Tyr
<a href="#">4EDZ</a>	PGDS	human	GSH	005	2012	2	2VCQ_A	Tyr
<a href="#">4EE0</a>	PGDS	human	GSF	004	2012	1,75	2VCQ_A	Tyr
<a href="#">2HNL</a>	PGDS	oncvo	GTT	-	2006	2	2HNL_A	Tyr
<a href="#">1PD2</a>	PGDS	rat	GTT	-	1998	2,3	1PD2_1	Tyr
<a href="#">1Z9H</a>	PGES	macfa	-	IMN	2005	2,6	2PBJ_A	Cys
<a href="#">2PBJ</a>	PGES	macfa	GSH	HEM	2007	2,8	2PBJ_A	Cys
<a href="#">1BX9</a>	Phi	arath	-	B	1998	2,6	1BX9_A	Ser
<a href="#">1GNW</a>	Phi	arath	-	GTX	1996	2,2	1BX9_A	Ser
<a href="#">1AW9</a>	Phi	zeama	-	-	1997	2,2	1AW9_A	Ser
<a href="#">1AXD</a>	Phi	zeama	-	C	1997	2,5	1AXD_A	Ser
<a href="#">1BYE</a>	Phi	zeama	-	ATA	1998	2,8	1AXD_A	Ser
<a href="#">10GS</a>	Pi	human	-	G	1997	2,2	3GUS_A	Tyr
<a href="#">11GS</a>	Pi	human	GTT	EAA	1997	2,3	3GUS_A	Tyr
<a href="#">12GS</a>	Pi	human	-	C	1997	2,1	3GUS_A	Tyr
<a href="#">13GS</a>	Pi	human	GSH	SAS	1997	1,9	3GUS_A	Tyr
<a href="#">14GS</a>	Pi	human	-	-	1997	2,8	3GUS_A	Tyr
<a href="#">16GS</a>	Pi	human	-	-	1997	1,9	3GUS_A	Tyr
<a href="#">17GS</a>	Pi	human	-	GTX	1997	1,9	3GUS_A	Tyr
<a href="#">18GS</a>	Pi	human	-	GDN	1997	1,9	3GUS_A	Tyr
<a href="#">19GS</a>	Pi	human	GTT	BSP	1997	1,9	3GUS_A	Tyr

Code PDB	Classe	Organisme	Ligand1	Ligand2	Année	Résol. (Å)	Modèle retenu	Résidu catalyt.
<a href="#">1AQV</a>	Pi	human	-	C	1997	1,94	3GUS_A	Tyr
<a href="#">1AQW</a>	Pi	human	GSH	-	1997	1,8	3GUS_A	Tyr
<a href="#">1AQX</a>	Pi	human	GSH	-	1997	2	3GUS_A	Tyr
<a href="#">1EOG</a>	Pi	human	-	-	2000	2,1	3GUS_A	Tyr
<a href="#">1EOH</a>	Pi	human	-	-	2000	2,5	3GUS_A	Tyr
<a href="#">1GSS</a>	Pi	human	-	I	1992	2,8	3GUS_A	Tyr
<a href="#">1KBN</a>	Pi	human	GTT	-	2001	2	3GUS_A	Tyr
<a href="#">1LBK</a>	Pi	human	GSH	-	2002	1,86	3GUS_A	Tyr
<a href="#">1MD3</a>	Pi	human	GSH	-	2002	2,03	3GUS_A	Tyr
<a href="#">1MD4</a>	Pi	human	GSH	-	2002	2,1	3GUS_A	Tyr
<a href="#">1PGT</a>	Pi	human	-	GTX	1997	1,8	3GUS_A	Tyr
<a href="#">1PX6</a>	Pi	human	GSH	-	2003	2,1	3GUS_A	Tyr
<a href="#">1PX7</a>	Pi	human	GSH	-	2003	2,03	3GUS_A	Tyr
<a href="#">1ZGN</a>	Pi	human	GSH	-	2005	2,1	3GUS_A	Tyr
<a href="#">20GS</a>	Pi	human	-	CBD	1997	2,45	3GUS_A	Tyr
<a href="#">22GS</a>	Pi	human	-	-	1998	1,9	3GUS_A	Tyr
<a href="#">2A2R</a>	Pi	human	-	GSN	2005	1,4	3GUS_A	Tyr
<a href="#">2A2S</a>	Pi	human	-	GSN	2005	1,7	3GUS_A	Tyr
<a href="#">2GSS</a>	Pi	human	-	EAA	1996	1,9	3GUS_A	Tyr
<a href="#">2J9H</a>	Pi	human	-	GTX	2006	2,4	3GUS_A	Tyr
<a href="#">2PGT</a>	Pi	human	-	GPR	1997	1,9	3GUS_A	Tyr
<a href="#">3CSH</a>	Pi	human	GTT	LZ6	2008	1,55	3GUS_A	Tyr
<a href="#">3CSI</a>	Pi	human	GTT	LZ6	2008	1,9	3GUS_A	Tyr
<a href="#">3CSJ</a>	Pi	human	-	CBL	2008	1,9	3GUS_A	Tyr
<a href="#">3DD3</a>	Pi	human	GTT	-	2008	2,25	3GUS_A	Tyr
<a href="#">3DGQ</a>	Pi	human	-	EAA	2008	1,6	3GUS_A	Tyr
<a href="#">3GSS</a>	Pi	human	GTT	EAA	1996	1,9	3GUS_A	Tyr
<a href="#">3GUS</a>	Pi	human	GTT	N11	2009	1,53	3GUS_A	Tyr
<a href="#">3HJM</a>	Pi	human	-	-	2009	2,1	3GUS_A	Tyr
<a href="#">3HJO</a>	Pi	human	GTT	EAA	2009	1,95	3GUS_A	Tyr
<a href="#">3HKR</a>	Pi	human	-	-	2009	1,8	3GUS_A	Tyr
<a href="#">3IE3</a>	Pi	human	GTT	N11	2009	1,8	3GUS_A	Tyr
<a href="#">3KM6</a>	Pi	human	GTT	EAA	2009	2,1	3GUS_A	Tyr
<a href="#">3KMN</a>	Pi	human	-	-	2009	1,8	3GUS_A	Tyr
<a href="#">3KMO</a>	Pi	human	GTT	EAA	2009	2,6	3GUS_A	Tyr
<a href="#">3N9J</a>	Pi	human	CL	EAA	2010	1,85	3GUS_A	Tyr

Code PDB	Classe	Organisme	Ligand1	Ligand2	Année	Résol. (Å)	Modèle retenu	Résidu catalyt.
<a href="#">3PGT</a>	Pi	human	-	GBX	1999	2,14	3GUS_A	Tyr
<a href="#">4GSS</a>	Pi	human	-	GTX	1997	2,5	3GUS_A	Tyr
<a href="#">4PGT</a>	Pi	human	-	GBX	1999	2,1	3GUS_A	Tyr
<a href="#">5GSS</a>	Pi	human	GTT	-	1997	1,95	3GUS_A	Tyr
<a href="#">6GSS</a>	Pi	human	GTT	-	1997	1,9	3GUS_A	Tyr
<a href="#">7GSS</a>	Pi	human	GTT	-	1997	2,2	3GUS_A	Tyr
<a href="#">8GSS</a>	Pi	human	GTT	-	1997	1,9	3GUS_A	Tyr
<a href="#">9GSS</a>	Pi	human	-	GTX	1997	1,97	3GUS_A	Tyr
<a href="#">1BAY</a>	Pi	mouse	-	-	1996	2	2OA7_A	Tyr
<a href="#">1GLP</a>	Pi	mouse	GTS	-	1994	1,9	2OA7_A	Tyr
<a href="#">1GLQ</a>	Pi	mouse	-	GTB	1994	1,8	2OA7_A	Tyr
<a href="#">1GSY</a>	Pi	mouse	GTT	-	1996	2,44	2OA7_A	Tyr
<a href="#">1GTI</a>	Pi	mouse	-	GTB	1998	3	2OA7_A	Tyr
<a href="#">2GLR</a>	Pi	mouse	-	GTX	1994	2,2	2OA7_A	Tyr
<a href="#">2OA7</a>	Pi	mouse	-	GTX	2006	2,2	2OA7_A	Tyr
<a href="#">2OAC</a>	Pi	mouse	-	GTB	2006	2,2	2OA7_A	Tyr
<a href="#">2OAD</a>	Pi	mouse	-	GTB	2006	2,5	2OA7_A	Tyr
<a href="#">3O76</a>	Pi	mouse	-	GTB	2010	1,77	2OA7_A	Tyr
<a href="#">1TU7</a>	Pi	oncvo	GTT	-	2004	1,5	1TU7_A	Tyr
<a href="#">1TU8</a>	Pi	oncvo	-	GTX	2004	1,8	1TU7_A	Tyr
<a href="#">2GSR</a>	Pi	suscr	GTT	-	1996	2,11	2GSR_A	Tyr
<a href="#">3T2U</a>	Pi	wucba	GSH	-	2011	2,3	3T2U_A	Tyr
<a href="#">3QAV</a>	Rho	latel	-	-	2011	2,1	3QAW_A	Ser
<a href="#">3QAW</a>	Rho	latel	GSH	-	2011	2,2	3QAW_A	Ser
<a href="#">3VPQ</a>	Sigma	bombo	GSH	PGO	2012	1,7	3VPQ_A	Tyr
<a href="#">3VPT</a>	Sigma	bombo	-	PGO	2012	1,9	3VPQ_A	Tyr
<a href="#">3VUR</a>	Sigma	bommo	GTS	1PE	2013	1,37	3VPQ_A	Tyr
<a href="#">1M0U</a>	Sigma	drome	-	GSW	2002	1,75	1M0U_A	Tyr
<a href="#">2WB9</a>	Sigma	fashe	GSH	CGL	2009	1,6	2WB9_A	Tyr
<a href="#">2WDU</a>	Sigma	fashe	GSH	GDS	2009	1,62	2WB9_A	Tyr
<a href="#">1GSQ</a>	Sigma	squid	-	GDN	1995	2,4	1GSQ_A	Tyr
<a href="#">2GSQ</a>	Sigma	squid	-	GBI	1995	2,2	1GSQ_A	Tyr
<a href="#">1OE7</a>	Sigma'	schha	GSH	-	2003	1,8	1OE8_A	Tyr
<a href="#">1OE8</a>	Sigma'	schha	GSH	-	2003	1,65	1OE8_B	Tyr
<a href="#">2C80</a>	Sigma'	schha	-	GTX	2005	2,3	1OE8_A	Tyr
<a href="#">2C8U</a>	Sigma'	schha	-	-	2005	2	1OE8_A	Tyr

Code PDB	Classe	Organisme	Ligand1	Ligand2	Année	Résol. (Å)	Modèle retenu	Résidu catalyt.
<a href="#">2CA8</a>	Sigma'	schha	-	GSW	2005	2,49	1OE8_A	Tyr
<a href="#">2CAI</a>	Sigma'	schha	-	-	2005	2,26	1OE8_A	Tyr
<a href="#">2CAQ</a>	Sigma'	schha	-	GSW	2005	2	1OE8_A	Tyr
<a href="#">2F8F</a>	Sigma'	schha	GTT	-	2005	2,1	1OE8_A	Tyr
<a href="#">1U3I</a>	Sigma'	schma	-	GSW	2004	1,89	1U3I_A	Tyr
<a href="#">3LYK</a>	SSPA	haein	-	-	2010	2,1	3LYK_A	?
<a href="#">4DEJ</a>	SSPA	idilo	-	-	2012	2,9	4DEJ_A	?
<a href="#">3LYP</a>	SSPA	psepr	-	-	2010	1,6	3LYP_B	?
<a href="#">3MDK</a>	SSPA	psepu	-	-	2010	1,85	3MDK_B	?
<a href="#">1YY7</a>	SSPA	yerpe	-	-	2005	2,02	1YY7_A	?
<a href="#">1GWC</a>	Tau	aegta	-	GTX	2002	2,25	1GWC_C	Ser
<a href="#">1OYJ</a>	Tau	orysa	GTT	GSW	2003	1,95	1OYJ_C	Ser
<a href="#">4J2F</a>	Tau	ricco	-	-	2013	1,9	4J2F_A	Ser
<a href="#">2VO4</a>	Tau	glyma	-	GTB	2008	1,75	3FHS_A	Ser
<a href="#">3FHS</a>	Tau	glyma	GSH	-	2008	2,71	3FHS_A	Ser
<a href="#">1LJR</a>	Thêta	human	GTT	-	1998	3,2	1LJR_A	Ser
<a href="#">2C3N</a>	Thêta	human	-	-	2005	1,5	2C3N_A	Ser
<a href="#">2C3Q</a>	Thêta	human	-	GTX	2005	1,85	2C3N_A	Ser
<a href="#">2C3T</a>	Thêta	human	-	-	2005	2,4	2C3N_A	Ser
<a href="#">2LJR</a>	Thêta	human	-	-	1998	3,2	1LJR_A	Ser
<a href="#">3LJR</a>	Thêta	human	-	GGC	1998	3,3	1LJR_A	Ser
<a href="#">4F0B</a>	Ure2p	phach	GDS	-	2013	1,45	4F0B_A	?
<a href="#">4F0C</a>	Ure2p	phach	GDS	-	2013	1,9	4F0C_A	?
<a href="#">1G6W</a>	Ure2p	sacce	-	-	2000	2,5	1G6Y_B	?
<a href="#">1G6Y</a>	Ure2p	sacce	-	-	2000	2,8	1G6Y_B	?
<a href="#">1HQO</a>	Ure2p	sacce	-	-	2000	2,3	1G6Y_B	?
<a href="#">1JZR</a>	Ure2p	sacce	GTT	-	2001	2,9	1G6Y_B	?
<a href="#">1K0A</a>	Ure2p	sacce	GTT	GTX	2001	2,5	1G6Y_B	?
<a href="#">1K0B</a>	Ure2p	sacce	GTT	-	2001	2,5	1G6Y_B	?
<a href="#">1K0C</a>	Ure2p	sacce	GTT	GTB	2001	2,5	1G6Y_B	?
<a href="#">1K0D</a>	Ure2p	sacce	GTT	-	2001	2,2	1G6Y_B	?
<a href="#">4L8E</a>	Ure2p	xenne	GSH	-	2013	1,7	4L8E_A	?
<a href="#">3R3E</a>	Xi	coli	GSH	-	2011	2,21	4G0K_A	Cys
<a href="#">4G0I</a>	Xi	coli	-	-	2012	2,05	4G0K_A	Cys
<a href="#">4G0K</a>	Xi	coli	-	OVS	2012	2,56	4G0K_A	Cys
<a href="#">4G0L</a>	Xi	coli	GSH	-	2012	2,62	4G0K_A	Cys

Code PDB	Classe	Organisme	Ligand1	Ligand2	Année	Résol. (Å)	Modèle retenu	Résidu catalyt.
<a href="#">3M1G</a>	Xi	corgl	-	-	2010	2,1	3M1G_A	Cys
<a href="#">3PPU</a>	Xi	phach	GTT	-	2010	2,3	3PPU_A	Cys
<a href="#">4FQU</a>	Xi	sphch	-	-	2012	3	4FQU_A	Cys
<a href="#">4IGJ</a>	Zeta	anade	-	-	2012	1,48	4IGJ_B	Ser
<a href="#">1E6B</a>	Zeta	arath	-	-	2000	1,65	1E6B_A	Ser
<a href="#">3LG6</a>	Zeta	cocim	-	-	2010	2,2	3N5O_A	Ser
<a href="#">3N5O</a>	Zeta	cocim	GTT	-	2010	1,85	3N5O_A	Ser
<a href="#">1FW1</a>	Zeta	human	GSH	-	2000	1,9	1FW1_A	Ser
<a href="#">3NIV</a>	Zeta	legpn	-	-	2010	2,3	3NIV_C	Ser
<a href="#">2CZ2</a>	Zeta	mouse	GSH	-	2005	1,4	2CZ2_A	Ser
<a href="#">2CZ3</a>	Zeta	mouse	-	-	2005	2,3	2CZ2_A	Ser
<a href="#">2JL4</a>	Zeta	ralstonia	GSH	-	2008	2,3	2V6K_A	Ser
<a href="#">2V6K</a>	Zeta	ralstonia	-	TGG	2007	1,3	2V6K_A	Ser



## **F Références bibliographiques**

- A**CHILONU, I., FANUCCHI, S., CROSS, M., FERNANDES, M. & DIRR, H. W. (2012). "Role of Individual Histidines in the pH-Dependent Global Stability of Human Chloride Intracellular Channel 1". *Biochemistry* **51**, pp. 995-1004.
- ADAMS, P. D., AFONINE, P. V., BUNKÓCZI, G., CHEN, V. B., DAVIS, I. W., ECHOLS, N., HEAD, J. J., HUNG, L. W., KAPRAL, G. J., GROSSE-KUNSTLEVE, R. W., McCOY, A. J., MORIARTY, N. W., OEFFNER, R., READ, R. J., RICHARDSON, D. C., RICHARDSON, J. S., TERWILLIGER, T. C. & ZWART, P. H. (2010). "PHENIX: A comprehensive Python-based system for macromolecular structure solution". *Acta D.* **66**, pp. 213-221.
- ADLER, V., YIN, Z., FUCHS, S. Y., BENEZRA, M., ROSARIO, L., TEW, K. D. PINCUS, M. R., SARDANA, M., HENDERSON, C. J., WOLF, C. R., DAVIS, R. J. & RONAI, Z. (1999). "Regulation of JNK signaling by GSTp". *EMBO J.* **18**, pp.1321-1334.
- ALLOCATI, N., CASALONE, E., MASULLI, M., CECCARELLI, I., CARLETTI, E., PARKER, M. W. & DI ILIO, C. (1999). "Functional analysis of the evolutionarily conserved proline 53 residue in *Proteus mirabilis* glutathione S-transferase B1-1". *FEBS Lett.* **445**, pp. 347-350.
- ALLOCATI, N., FEDERICI, L., MASULLI, M. & DI ILIO, C. (2009). "Glutathione transferases in bacteria", *FEBS Journal.* **276**, pp. 58-75.
- ALLOCATI, N., FEDERICI, L., MASULLI, M. & DI ILIO, C. (2012). "Distribution of glutathione transferases in Gram-positive bacteria and Archaea". *Biochimie* **94**, pp. 588-596.
- ALTSCHUL, S.F., GISH, W., MILLER, W., MYERS, E.W. & LIPMAN, D.J. (1990). "Basic local alignment search tool." *J. Mol. Biol.* **215**, pp. 403-410.
- ANDREWS, C. J., SKIPSEY, M., TOWNSON, J. K., MORRIS, C., JEPSON, I. & EDWARDS, R. (1997). "Glutathione transferase activity towards herbicides used selectively in soybean". *Pestic. Sci.* **51**, pp. 213-222.
- ARCA, P., HARDISSON, C. & SUAREZ, J. E. (1997). "Purification of a glutathione S-transferase that mediates fosfomycin resistance in bacteria". *Antimicrob. Agents Chemother.* **34**, pp. 844-848.
- ARMSTRONG, R. N. (1991). "Glutathione S-transferases: reaction mechanism, structure, and function". *Chem. Res. Toxicol.* **4**, pp. 131-140.
- ARMSTRONG, R. N. (1997). "Structure, Catalytic Mechanism, and Evolution of the Glutathione Transferases". *Chem. Res. Toxicol.* **10**, pp. 2-18.

- ARMSTRONG, R. (1998). "Mechanistic imperatives for the evolution of glutathione transferases". *Curr. Opin. Chem. Biol.* **2**, pp. 618-623.
- ASOJO, O. A., HOMMA, K., SEDLACEK, M., NGAMELUE, M., GOUD, G. N., ZHAN, B., DEUMIC, V., ASOJO, O. & HOTEZ, P. J. (2007). "X-ray structures of Na-GST-1 and Na-GST-2 two glutathione s-transferase from the human hookworm *Necator americanus*". *BMC Struct. Biol.* **7**.
- ATKINSON, H. J. & BABBITT, P. C. (2009). "Glutathione transferases are structural and functional outliers in the thioredoxin fold". *Biochemistry* **48**, pp. 11108-11116.
- AI, M., ZHOU, J. M. & PERRETT, S. (2004). "The yeast prion protein Ure2 shows glutathione peroxidase activity in both native and fibrillar forms". *J. Biol. Chem.* **279**, pp. 50025-50030.
- B**AI, M., ZHOU, J. M. & PERRETT, S. (2004). "The yeast prion protein Ure2 shows glutathione peroxidase activity in both native and fibrillar forms". *J. Biol. Chem.* **279**, pp. 50025-50030.
- BARR, D. P. & AUST, S. D. (1994). "Mechanisms the white rot fungi use to degrade pollutants". *Environ. Sci. Technol.* **28**, pp. 79-87.
- BERRY, K. L. & HOBERT, O. (2006). "Mapping functional domains of chloride intracellular channel (CLIC) proteins in vivo". *J. Mol. Biol.* **359**, pp. 1316-1333.
- BICO, P., ERHARDT, J., KAPLAN, W. & DIRR, H. (1995). "Porcine class pi glutathione S-transferase: anionic ligand binding and conformational analysis". *Biochim. Biophys. Acta* **1247**, pp. 225-230.
- BOARD, P. G., BAKER, R. T., CHELVANAYAGAM, G. & JERMIN, L. S. (1997). "Zeta, a novel class of glutathione transferases in a range of species from plants to humans". *Biochem. J.* **328**, pp. 929-935.
- BOARD, P. G., COGGAN, M., CHELVANAYAGAM, G., EASTEAL, S., JERMIIN, L. S., SCHULTE, G. K. DANLEY, D. E., HOTH, L. R., GRIFFOR, M. C., KAMATH, A. V., ROSNER, M. H., CHRUNYK, B. A., PERREGAUX, D. E., GABEL, C. A., GEOGHEGAN, K. F. & PANDIT, J. (2000). "Identification, characterization and crystal structure of the omega class glutathione transferases". *J. Biol. Chem.* **275**, pp. 24798-24806.
- BOARD, P. G. & MENON, D. (2013). "Glutathione transferases, regulators of cellular metabolism and physiology". *Biochim. Biophys. Acta* **1830**, pp. 3267-3288.
- BOUSSET, L., BELRHALI, H., JANIN, J., MELKI, R., MORERA, S. (2001). "Structure of the globular region of the prion protein Ure2 from the yeast *Saccharomyces cerevisiae*". *Structure* **9**, pp. 39-46.
- BRESELL, A., WEINANDER, R., LUNDQVIST, G., RAZA, H., SHIMOJI, M., SUN, T.-H., BALK, L., WIKLUND, R., ERIKSSON, J., JANSSON, C., PERSSON, B., JAKOBSSON, P.-J. & MORGENSTERN, R. (2005). "Bioinformatic and enzymatic characterization of the MAPEG superfamily". *FEBS J.* **272**, pp.1688-1703.

- BROCK, J., BOARD, P. G. & OAKLEY, A. (2013). "Structural Insights into Omega-Class Glutathione Transferases: A Snapshot of Enzyme Reduction and Identification of a Non-Catalytic Ligandin Site". *PLoS ONE* **8**, pp. 1-10.
- BURDSALL, H. H. JR (1981). "The taxonomy of *sporotrichum pruinosum* and *sporotrichum pulverulentum/phanerochaete chrysosporium*". *Mycologia* **73**, pp. 675-680.
- BURDSALL, H. H., (1985). "A contribution to the taxonomy of the genus *phanerochaete*". *Mycologia Memoirs* **10**, pp. 61-63.
- BUSENLEHNER, L. S., ALANDER, J., JEGERSCOHL, C., HOLM, P. J., BHAKAT, P., HEBERT, H., MORGENSTERN, R. & ARMSTRONG, R. N. (2007). "Location of substrate binding sites within the integral membrane protein microsomal glutathione transferase-1". *Biochemistry* **46**, pp. 2812-2822.
- CACCURI, A. M., ACETO, A., PIEMONTE, F., DI ILIO, C., ROSATO, N. & FEDERICI, G. (1990). "Interaction of hemin with placental glutathione transferase". *Eur. J. Biochem.* **189**, pp. 493-497.
- CACCURI, A. M., ANTONINI, G., ALLOCATI, N., ILIO, C. D., DE MARIA, F., INNOCENTI, F., PARKER, M. W., MASULLI, M., BELLO, M. L., TURELLA, P., FEDERICI, G. & RICCIA, G. (2002). "GSTB1-1 from *Proteus mirabilis*: A snapshot of an enzyme in the evolutionary pathway from a redox enzyme to a conjugating enzyme". *J. Biol. Chem.* **277**, pp. 18777-18784.
- CHA, C.-J., KIM, S.-J., KIM, Y.-H., STINGLEY, R. & CERNIGLIA, C.E. (2002) "Molecular cloning, expression and characterization of a novel class glutathione S-transferase from the fungus *Cunninghamella elegans*". *Biochem. J.* **368**, pp. 589-595.
- CHEN, B.Y., MA, X. X., GUO, P. C., TAN, X., LI, W. F., YANG, J. P., ZHANG, N. N., CHEN, Y., XIA, Q. & ZHOU, C. Z. (2011). "Structure-guided activity restoration of the silkworm glutathione transferase Omega GSTO3-3". *J. Mol. Biol.* **412**, pp. 204-211.
- CLARK, A. G., SMITH, J. N., & SPEIR, T. W. (1973). "Cross-specificity in some vertebrate and insect glutathione-transferases with methyl parathion (dimethyl p-nitrophenyl phosphorothionate), 1-chloro-2,4-dinitro-benzene and S-crotonyl-N-acetylcysteamine as substrates". *Biochem. J.* **135**, pp. 385-392.
- COLE, C., BARBER, J. D. & BARTON, G. J. (2008). "The Jpred 3 secondary structure prediction server." *Nucleic Acids Res.* **36**, pp.197-201.
- COWTAN, K. (2006). "The Buccaneer software for automated model building. 1. Tracing protein chains". *Acta D.* **62**, pp. 1002-1011.
- DI ILIO, C., ACETO, A., PICCOLOMINI, R., ALLOCATI, N., FARAONE, A., CELLINI, L., RAVAGNAN, G. & FEDERICI, G. (1988). "Purification and characterization of three forms of glutathione transferase from *Proteus mirabilis*". *Biochem. J.* **255**, pp. 971-975.

DIRR, H., REINEMER, P. & HUBER, R. (1994). "X-ray crystal structures of cytosolic glutathione S-transferases - implications for protein architecture, substrate recognition and catalytic function". *Eur. J. Biochem.* **220**, pp. 645-661.

DIXON, D. P., DAVIS, B. G. & EDWARDS, R. (2002). "Functional divergence in the glutathione transferase superfamily in plants—Identification of two classes with putative functions in redox homeostasis in *Arabidopsis thaliana*". *J. Biol. Chem.* **277**, pp. 30859-30869.

DIXON, D. P., LAPTHORN, A., MADESIS, P., MUDD, E. A., DAY, A. & EDWARDS, R. (2008). "Binding and glutathione conjugation of porphyrinogens by plant glutathione transferases". *J. Biol. Chem.* **283**, pp. 20268-20276.

DIXON, D. P. & EDWARDS, R. (2010). "Roles for Stress-inducible Lambda Glutathione Transferases in flavonoid metabolism in plants as identified by ligand fishing". *J. Biol. Chem.* **285**, pp. 36322-36329.

DRAGANI, B., STERBERG, G., MELINO, S., PETRUZZELLI, R., MANNERVIK, B. & ACETO, A. (1997). "The Conserved N-capping Box in the hydrophobic core of Glutathione S-Transferase P1-1 Is Essential for Refolding". *J. Biol. Chem.* **272**, pp. 25518-25523.

EMSLEY, P., LOHKAMP, B., SCOTT, W. G. & COWTAN, K. (2010). "Features and development of Coot". *Acta D.* **66**, pp. 486-501.

EVANS, G. & PETTIFER, R. F. (2001). "CHOOCH: a program for deriving anomalous-scattering factors from X-ray fluorescence spectra". *J. Appl. Cryst.* **34**, pp. 82-86.

EVANS, P. (2006). "Scaling and assessment of data quality". *Acta D.* **62**, pp. 72-82.

FERGUSON, A. D., MCKEEVER, B. M., XU, S., WISNIEWSKI, D., MILLER, D. K., YAMIN, T.-T., SPENCER, R. H., CHU, L., UJJAINWALLA, F., CUNNINGHAM, B. R., EVANS, J. F. & BECKER, J. W. (2007). "Crystal structure of inhibitor bound human 5-lipoxygenase-activating protein". *Science* **317**, pp. 510-512.

FEDERICI, L., MASULLI, M., BONIVENTO, D., DI MATTEO, A., GIANNI, S., FAVALORO, B., DI ILIO, C. & ALLOCATI, N. (2007). "Role of Ser11 in the stabilization of the structure of *Ochrobactrum anthropi*". *Biochem. J.* **403**, pp. 267-274.

FELSENSTEIN, J. (1989). « PHYLIP - Phylogeny Inference Package (Version 3.2). » *Cladistics* **5**, pp. 164-166.

FROVA, C. (2006). "Glutathione transferases in the genomics era: new insights and perspectives". *Biomol. Eng.* **23**, pp. 149-169.

FYFE, P. K., WESTROP, G. D., SILVA, A. M., COOMBS, G. H. & HUNTER, W. N. (2012). "Leishmania TDR1 structure, a unique trimeric glutathione transferase capable of deglutathionylation and antimonial prodrug activation". *PNAS* **109**, pp. 11693-11698.

GERLT, J. A., ALLEN, K. N., ALMO, S. C., ARMSTRONG, R. N., BABBITT, P. C., CRONAN, J. E., DUNAWAY-MARIANO, D., IMKER, H. J., JACOBSON, M. P., MINOR, W., POULTER, C. D., RAUSHEL, F. M., SALI, A., SHOICHET, B. K. & SWEEDLER, J. V. (2011). "The Enzyme Function Initiative". *Biochemistry* **50**, pp. 9950-9962.

GILLE, C. & FROMMEL, C. (2001). "STRAP: editor for Structural Alignments of Proteins". *Bioinformatics*. **17**, pp. 377-378.

GREEN, A.R., HAYES, R.P., XUN, L., KANG, C. (2012). "Structural understanding of GSH-dependent reduction mechanism of glutathionyl-hydroquinone reductases." *J. Biol. Chem.* **287**, pp. 35838-35848.

HABIG, W. H., PABST, M. J. & JAKOBY, W. B. (1974). "Glutathione S-transferases. The first enzymatic step in mercapturic acid formation". *J. Biol. Chem.* **249**, pp. 7130-7139.

HANSEN, A. M., GU, Y., LI, M., ANDRYKOVITCH, M., WAUGH, D. S., JIN, D. J. & Ji, X. (2005). "Structural basis for the function of stringent starvation protein a as a transcription factor". *J. Biol. Chem.* **280**, pp. 17380-17391.

HANSSON, L. O., BOLTON-GROB, R., MASSOUD, T. & MANNERVIK, B. (1999). "Evolution of differential substrate specificities in Mu class glutathione transferases probed by DNA shuffling". *J. Mol. Biol.* **287**, pp. 265-276.

HARROP, S. J., DEMAERE, M. Z., FAIRLIE, W. D., REZTSOVA, T., VALENZUELA, S. M., MAZZANTI, M., TONINI, R., QIU, M. R., JANKOVA, L., WARTON, K., BAUSKIN, A. R., WU, W. M., PANKHURST, S., CAMPBELL, T. J., BREIT, S. N. & CURMI, P. M. (2001). "Crystal structure of a soluble form of the intracellular chloride ion channel CLIC1 (NCC27) at 1.4 Å resolution". *J. Biol. Chem.* **276**, pp. 44993-45000.

HATAKKA, A., (2001). "Biodegradation of lignin" in "Lignin humic substances and coal". *Biopolymers* **1**, pp. 129-145, Ed. Wiley-VCH.

HAYES, J. D. & MANTLE, T. J. (1986). "Use of immunoblot techniques to discriminate between the glutathione S-transferase Y<sub>f</sub>, Y<sub>k</sub>, Y<sub>a</sub>, Y<sub>n</sub>/Y<sub>b</sub> and Y<sub>c</sub> subunits and to study their distribution in extra-hepatic tissues. Evidence for three immunochemically distinct groups of transferases in the rat". *Biochem. J.* **233**, pp. 779-788.

HAYES, J. D., FLANAGAN, J. D. & JOWSEY, I. R. (2005). "Glutathione transferases". *Annu. Rev. Pharmacol. Toxicol.* **45**, pp. 51-88.

HEGAZY, U. M., MANNERVIK, B. & STENBERG, G. (2004). "Functional role of the lock and key motif at the subunit interface of glutathione transferase p1-1". *J. Biol. Chem.* **279**, pp. 9586-9596.

HOLM, P. J., BHAKAT, P., JEGERSCHOLD, C., GYOBU, N., MITSUOKA, K., FUJIYOSHI, Y., MORGENSTERN, R. & HEBERT, H. (2006). "Structural basis for detoxification and oxidative stress protection in membranes". *J. Mol. Biol.* **360**, pp. 934-45.

- HOLM, L. & ROSENSTROM, P. (2010). "Dali server: Conservation mapping in 3D". *Nucl. Ac. Res.* **38**, pp.545-549.
- HOWE, K., BATEMAN, A. & DURBIN, R. (2002). QuickTree: building huge Neighbour-Joining trees of protein sequences. *Bioinformatics* **18**, pp. 1546-1547.
- HUANG, Y. C., MISQUITTA, S., BLOND, S. Y., ADAMS, E. & COLMAN, R. F. (2008). "Catalytically active monomer of glutathione S-transferase pi and key residues involved in the electrostatic interaction between subunits". *J. Biol. Chem.* **283**, pp. 32880-32888.
- JEPPESEN, M. G., ORTIZ, P., SHEPARD, W., KINZY, T. G., NYBORG, J. & ANDERSEN, G. R. (2003). "The crystal structure of the glutathione S-transferase-like domain of elongation factor 1B gamma from *Saccharomyces cerevisiae*". *J. Biol. Chem.* **278**, pp. 47190-47198.
- JI, X., JOHNSON, W. W., SESAY, M. A., DICKERT, L., PRASAD, S. M., AMMON, H. L., ARMSTRONG, R. N. & GILLILAND, G. L. (1994). "Structure and function of the xenobiotic substrate binding site of a glutathione S-transferase as revealed by X-ray crystallographic analysis of product complexes with the diastereomers of 9-(S-glutathionyl)-10-hydroxy-9,10-dihydrophenanthrene". *Biochemistry* **33**, pp. 1043-1052.
- JI, X., ZHANG, P., ARMSTRONG, R. N. & GILLILAND, G. L. (1992). "The three-dimensional structure of a glutathione S-transferase from the Mu gene class. Structural analysis of the binary complex of isoenzyme 3-3 and glutathione at 2.2-Å resolution". *Biochemistry* **31**, pp. 10169-10184.
- JI, X., VON ROSENVINGE, E. C., JOHNSON, W. W., TOMAREV, S. I., PIATIGORSKY, J., ARMSTRONG, R. N. & GILLILAND, G. L. (1995). "Three-dimensional structure, catalytic properties, and evolution of a sigma class glutathione transferase from squid, a progenitor of the lens S-crystallins of cephalopods". *Biochemistry* **34**, pp. 5317-5328.
- JI, X., VON ROSENVINGE, E. C., JOHNSON, W. W., ARMSTRONG, R. N. & GILLILAND, G. L. (1996). "Location of a potential transport binding site in a sigma class glutathione transferase by x-ray crystallography". *PNAS* **93**, pp. 8208-8213.
- JI, X., TORDOVA, M., O'DONNELL, R., PARSONS, J. F., HAYDEN, J. B., GILLILAND, G. L. & ZIMNIAK, P. (1997). "Structure and function of the xenobiotic substrate-binding site and location of a potential non-substrate binding site in a class pi glutathione S-transferase". *Biochemistry* **36**, pp. 9690-9702.
- JOHANSSON, A. S. & MANNERVIK, B. (2001). "Human glutathione transferase A3-3, a highly efficient catalyst of double-bond isomerization in the biosynthetic pathway of steroid hormones". *J. Biol. Chem.* **276**, pp. 33061-33065.
- KABSCH, W. (2010). "XDS". *Acta D.* **66**, pp. 125-132.

- KANAOKA, Y., AGO, H., INAGAKI, E., NANAYAMA, T., MIYANO, M., KIKUNO, R. FUJII, Y., EGUCHI, N., TOH, H., URADE, Y., & HAYAISHI, O. (1997). "Cloning and crystal structure of hematopoietic prostaglandin D synthase". *Cell* **90**, pp. 1085-1095.
- KATAYAMA, Y., NISHIKAWA, S., NAKAMURA, M., YANO, K., YAMASAKI, M., MOROHOSHI, N. & HARAGUCHI, T. (1987). "Cloning and expression of *Pseudomonas paucimobilis* SYK-6 genes involved in the degradation of vanillate and protocatechuate in *P. putida*". *Mokuzai Gakkaishi* **33**, pp. 77-79.
- KONG, K. H., TAKASU, K., INOUE, H. & TAKAHASHI, K. (1992). "Tyrosine-7 in human class-pi glutathione-S-transferase is important for lowering the pKa of the thiol-group of glutathione in the enzyme-glutathione complex". *BBRC*. **184**, pp. 194-197.
- KOSLOFF, M., HAN, G. W., SRI KRISHNA, S., SCHWARZENBACHER, R., FASNACHT, M., ELSLIGER, M. A., ABDUBEK, P., AGARWALLA, S., AMBING, E., ASTAKHOVA, T., AXELROD, H. L., CANAVES, J. M., CARLTON, D., CHIU, H. J., CLAYTON, T., DIDONATO, M., DUAN, L., FEUERHELM, J., GRITTINI, C., GRZECHNIK, S. K., HALE, J., HAMPTON, E., HAUGEN, J., JAROSZEWSKI, L., JIN, K. K., JOHNSON, H., KLOCK, H. E., KNUTH, M. W., KOESEMA, E., KREUSCH, A., KUHN, P., LEVIN, I., MCMULLAN, D., MILLER, M. D., MORSE, A. T., MOY, K., NIGOGHOSSIAN, E., OKACH, L., OOMMACHEN, S., PAGE, R., PAULSEN, J., QUIJANO, K., REYES, R., RIFE, C. L., SIMS, E., SPRAGGON, G., SRIDHAR, V., STEVENS, R. C., VAN DEN BEDEM, H., VELASQUEZ, J., WHITE, A., WOLF, G., XU, Q., HODGSON, K. O., WOOLEY, J., DEACON, A. M., GODZIK, A., LESLEY, S. A. & WILSON, I. A. (2006). "Comparative structural analysis of a novel glutathione S-transferase (Atu5508) from *Agrobacterium tumefaciens* at 2.0 Å resolution". *Proteins* **65**, 527-537.
- L**ABORDE, E. (2010). "Glutathione transferases as mediators of signaling pathways involved in cell proliferation and cell death". *Cell Death Differ.* **17**, pp.1373-1380.
- LADNER, J. E., PARSONS, J. F., RIFE, C. L., GILLILAND, G. L. & ARMSTRONG, R. N. (2004). "Parallel evolutionary pathways for glutathione transferases: structure and mechanism of the mitochondrial class kappa enzyme rGSTK1-1". *Biochemistry* **43**, pp. 352-361.
- LEDERER, B. & BOGER, P. (2003). "Binding and protection of porphyrins by glutathione S-transferases of *Zea mays*". *Biochim. Biophys. Acta* **1621**, pp. 226-233.
- LIM, C. E., MATTHAEI, K. I., BLACKBURN, A. C., DAVIS, R. P., DAHLSTROM, J. E., KOINA, M. E., ANDERS, M. W. & BOARD, P. G. (2004). "Mice deficient in glutathione transferase zeta/maleylacetoacetateisomérase exhibit a range of pathological changes and elevated expression of alpha, mu and pi class glutathione transferases". *Am. J. Pathol.* **165**, pp. 679-693.
- LITWACK, G., KETTERER, B. & ARIAS, I. M. (1971). "Ligandin: a hepatic protein which binds steroids, bilirubin, carcinogens and a number of exogenous organic anions". *Nature* **234**, pp. 466-467.

LITTLER, D. R., BROWN, L. J., BREIT, S. N., PERRAKIS, A. & CURMI, P. M. (2010). "Structure of human CLIC3 at 2 Ang. resolution". *Proteins* **78**, pp. 1594-1600.

**M**A, X.-X., JIANG, Y.-L., HE, Y.-X., BAO, R., CHEN, Y. & ZHOU, C.-Z. (2009). "Structures of yeast glutathione-S-transferase Gtt2 reveal a new catalytic type of GST family". *EMBO Rep.* **10**, pp. 1320-1326.

MACKINNON, S. S., MALEVANETS, A. & WODAK, S. J. (2013). "Intertwined associations in structures of homooligomeric proteins". *Structure* **21**, pp. 638-49.

MANNERVIK, B., AWASTHI, Y. C., BOARD, P. G., HAYES, J. D., DI ILIO, C., KETTERER, B., LISTOWSKY, I., MORGENSTERN, R., MURAMATSU, M., PEARSON, W. R., PICKETT, C. B., SATO, K., WIDERSTEN, M. & WOLF, C. R. (1992). "Nomenclature for human glutathione transferases [1]". *Biochem. J.* **282**, pp. 305-306.

MANNERVIK, B., BOARD, P. G., HAYES, J. D., LISTOWSKY, I., PEARSON, W. R. & HELMUT SIES AND LESTER, P. (2005). "Nomenclature for Mammalian Soluble Glutathione Transferases". *Methods in Enzymology, Vol.* **401**, pp. 1-8.

MARTINEZ MOLINA, D., WETTERHOLM, A., KOHL, A., MCCARTHY, A. A., NIEGOWSKI, D., OHLSON, E., HAMMARBERG, T., ESHAGHI, S., HAEGGSTROM, J. Z. & NORDLUND, P. (2007). "Structural basis for synthesis of inflammatory mediators by human leukotriene C4 synthase". *Nature* **448**, pp. 613-6.

MASAI, E., KATAYAMA, Y., NISHIKAWA, S., YAMASAKI, M., MOROHOSHI, N. & HARAGUCHI, T. (1989). "Detection and localization of a new enzyme catalyzing the beta-aryl ether cleavage in the soil bacterium (*Pseudomonas paucimobilis* SYK-6)". *FEBS Lett.* **249**, pp. 348-352.

MASAI, E., KATAYAMA, Y., KAWAI, S., NISHIKAWA, S., YAMASAKI, M. & MOROHOSHI, N. (1991). "Cloning and sequencing of the gene for a *Pseudomonas paucimobilis* enzyme that cleaves beta-aryl ether". *J. Bacteriol.* **173**, pp. 7950-7955.

MASAI, E., KATAYAMA, Y., KUBOTA, S., KAWAI, S., YAMASAKI, M. & MOROHOSHI, N. (1993). "A bacterial enzyme degrading the model lignin compound beta-etherase is a member of the glutathione- S-transferase superfamily". *FEBS Lett.* **323**, pp. 135-140.

MASAI, E., ICHIMURA, A., SATO, Y., MIYAUCHI, K., KATAYAMA, Y. & FUKUDA, M. (2003). "Roles of the enantioselective glutathione S-transferases in cleavage of beta-aryl ether". *J. Bacteriol.* **185**, pp. 1768-1775.

MASAI, E., KATAYAMA, Y. & FUKUDA, M. (2007). "Genetic and Biochemical investigations on bacterial catabolic pathways for lignin-derived aromatic compounds". *Biosci. Biotechnol. Biochem.* **71**, pp. 1-15.

McCARTHY, D. L., NAVARRETE, S., WILLET, W. S., BABBITT, P. C. & COPLEY, S. D. (1996). "Exploration of the relationship between tetrachlorohydroquinone dehalogenase and the glutathione S-transferase superfamily". *Biochemistry* **35**, pp.14634-14642.



- McCOY, A. J., GROSSE-KUNSTLEVE, R. W., ADAMS, P. D., WINN, M. D., STORONI, L. C. & READ, R. J. (2007). "Phaser crystallographic software". *J. Appl. Cryst.* **40**, pp. 658-674.
- McGOLDRICK, S., O'SULLIVAN, S. M. & SHEEHAN, D. (2005). "Glutathione transferase-like proteins encoded in genomes of yeasts and fungi: insights into evolution of a multifunctional protein superfamily". *FEMS Microbiol. Lett.* **242**, pp. 1-12.
- McTIGUE, M. A., WILLIAMS, D. R. & TAINER, J. A. (1995). "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.* **246**, pp. 21-27.
- MEYER, D. J., COLES, B., PEMBLE, S. E., GILMORE, K. S. & FRASER, G. M. (1991). "Theta, a new class of glutathione S-transferases purified from rat and man". *Biochem. J.* **274**, pp. 409-414.
- MOREL, M., NGADIN, A. A., DROUX, M., JACQUOT, J.-P. & GELHAYE, E. (2009). "The fungal glutathione S-transferase system. Evidence of new classes in the wood-degrading basidiomycete *Phanerochaete chrysosporium*". *Cell. Mol. Life Sci.* **66**, pp. 3711-3725.
- MORGENSTERN, R., GUTHENBERG, C. & DEPIERRE, J. W. (1982). "Microsomal glutathione transferase. Purification, initial characterization and demonstration that it is not identical to the cytosolic glutathione transferases A, B and C". *Eur. J. Biochem.* **128**, pp. 243-248.
- MURZIN, A. G., BRENNER, S. A., HUBBARD, T. & CHOTHIA, C. (1995). SCOP: "A structural classification of proteins database for the investigation of sequences and structures". *J. Mol. Biol.* **247**, pp. 536-540.
- N**ISHIDA, M., HARADA, S., NOGUCHI, S., SATOW, Y., INOUE, H. & TAKAHASHI, K. (1998). "Three-dimensional structure of *Escherichia coli* glutathione S-transferase complexed with glutathione sulfonate: catalytic roles of Cys10 and His106". *J. Mol. Biol.* **281**, pp. 135-147.
- O**AKLEY, A. J., LOBELLO, M., BATTISTONI, A., RICCI, G., ROSSJOHN, J., VILLAR, H. O. & PARKER, M. W. (1997). "The structures of human glutathione transferase P1-1 in complex with glutathione and various inhibitors at high resolution". *J. Mol. Biol.* **274**, pp. 84-100.
- OAKLEY, A. J., LO BELLO, M., NUCCETELLI, M., MAZZETTI, A. P. & PARKER, M. W. (1999). "The ligandin (non-substrate) binding site of human pi class glutathione transferase is located in the electrophile binding site (H-site)". *J. Mol. Biol.* **291**, pp. 913-926.
- OAKLEY, A. J., HARNNOI, T., UDOMSINPRASERT, R., JIRAJAROENRAT, K., KETTERMAN, A. J. & WILCE, M. C. J. (2001). "The crystal structures of glutathione S-transferases isozymes 1-3 and 1-4 from *Anopheles dirus* species B". *Prot. Sci.* **10**, 2176-2185.

- OAKLEY, A. J. (2005). "Glutathione transferases: new functions". *Curr. Opin. Struct. Biol.* **15**, pp. 716-723.
- OAKLEY, A. (2011). "Glutathione transferases: a structural perspective". *Drug Metabolism Reviews* **43**, pp. 138-151.
- OTSUKA, Y., SONOKI, T., IKEDA S., KAJITA, S., NAKAMURA, M. & KATAYAMA, Y. (2003). "Detection and characterization of a novel extracellular fungal enzyme that catalyzes the specific and hydrolytic cleavage of lignin guaiacylglycerol  $\beta$ -aryl ether linkages". *Eur. J. Biochem.* **270**, pp. 2353-2362.
- ÖZTETIK, E. (2008). "A Tale of Plant Glutathione S-Transferases: Since 1970". *Bot. Rev.* **74**, pp. 419-437.
- PARK, A. K., MOON, J. H., JANG, E. H., PARK, H., AHN, I. Y., LEE, K. S. & CHI, Y. M. (2013). "The structure of a shellfish specific GST class glutathioneS-transferase from antarctic bivalve *Laternula elliptica* reveals novel active site architecture". *Proteins* **81**, pp. 531-537.
- PEMBLE, S. E., WARDLE, A. F. & TAYLOR, J. B. (1996). "Glutathione S-transferase class Kappa: characterization by the cloning of rat mitochondrial GST and identification of a human homologue". *Biochem. J.* **319**, pp. 749-754.
- PEREZ, J. MUNOZ-DORADO, J. DE LA RUBIA, T. & MARTINEZ, J. (2002). "Biodegradation and biological treatments of cellulose, hemicellulose and lignin: an overview". *Int. Microbiol.* **5**, pp. 53-63.
- PERRAKIS, A., MORRIS, R. & LAMZIN, V. S. (1999). "Automated protein model building combined with iterative structure refinement". *Nat. Struct. Biol.* **6**, pp. 458-463.
- POLEKHINA, G., BOARD, P. G., BLACKBURN, A. C. & PARKER, M. W. (2001). "Crystal structure of maleylacetoacetate isomérase/glutathione transferase zeta reveals the molecular basis for its remarkable catalytic promiscuity". *Biochemistry* **40**, pp. 1567-1576.
- PRADE, L., HUBER, R., MANOHARAN, T. H., FAHL, W. E. & REUTER, W. (1997). "Structures of class pi glutathione S-transferase from human placenta in complex with substrate, transition-state analogue, and inhibitor". *Structure* **5**, pp. 1287-1295.
- REINEMER, P., DIRR, H. W., LADENSTEIN, R., HUBER, R., LO BELLO, M., FEDERICI, G. & PARKER, M. W. (1992). "Three-dimensional structure of class pi glutathione S-transferase from human placenta in complex with S-hexylglutathione at 2.8 Å resolution". *J. Mol. Biol.* **227**, pp. 214-226.
- REINEMER, P., PRADE, L., HOF, P., NEUEFEIND, T., HUBER, R., ZETTL, R., PALME, K., SCHELL, J., KOELLN, I., BARTUNIK, H. D. & BIESELER, B. (1996). "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.* **255**, pp. 289-309.

- ROSSJOHN, J., MCKINSTRY, W. J., OAKLEY, A. J., VERGER, D., FLANAGAN, J., CHELVANAYAGAM, G., TAN, K.-L., BOARD, P. G. & PARKER, M. W. (1998a). "Human theta class glutathione transferase: the crystal structure reveals a sulfate-binding pocket within a buried active site". *Structure* **6**, pp. 309-322.
- ROSSJOHN, J., POLEKHINA, G., FEIL, S. C., ALLOCATI, N., MASULLI, M., DE ILLIO, C. & Parker, M. W. (1998b). "A mixed disulfide bond in bacterial glutathione transferase: functional and evolutionary implications". *Structure* **6**, pp. 721-734.
- RICHARDSON, J. S. & RICHARDSON, D. C. (1988) "Amino acid preferences for specific locations at the ends of alpha helices". *Science* **240**, pp. 1648-1652.
- SATO, S., FELTUS, F. A., IYER, P. & TIEN, M. (2009). "The first genome-level transcriptome of the wood-degrading fungus *Phanerochaete chrysosporium* grown on red oak". *Curr. Genet.* **55**, pp. 273-286.
- SCHULLER, D. J., LIU, Q., KRIKSUNOV, I. A., CAMPBELL, A. M., BARRETT, J., BROPHY, P. M. & HAO, Q. (2005). "Crystal structure of a new class of glutathione transferase from the model human hookworm nematode *Heligmosomoides polygyrus*". *Proteins* **61**, pp. 1024-1031.
- SHEEHAN, D., MEADE, G., FOLEY, V. M. & DOWD, C. A. (2001). "Structure, function and evolution of glutathione transferases : implications for classification of non-mammalian members of an ancient enzyme superfamily", *Biochem. J.* **360**, pp. 1-16.
- SHOULDICE, S. R., HERAS, B., WALDEN, P. M., TOTSIKA ,M., SCHEMBRI, M. A. & MARTIN, J. L. (2011). "Structure and function of DsbA, a key bacterial oxidative folding catalyst". *Antioxid. Redox Signal.* **14**, 1729-1760.
- SINGH, D. & CHEN, S. (2008). "The white-rot fungus *Phanerochaete chrysosporium*: conditions for the production of lignin-degrading enzymes". *Appl Microbiol. Biotechnol.* **81**, pp. 399-417.
- SINNING, I., KLEYWEGT, G. J., COWAN, S. W., REINEMER, P., DIRR, H. W., HUBER, R., GILLILAND, G. L., ARMSTRONG, R. N., JI, X., BOARD, P. G., OLIN, B., MANNERVIK B. & JONES, T. A. (1993). "Structure determination and refinement of human alpha class glutathione transferase A1-1, and a comparison with the mu and pi class enzymes". *J. Mol. Biol.* **232**, pp. 192-212.
- STOURMAN, N. V., BRANCH, M. C., SCHAAB, M. R., HARP, J. M., LADNER, J. E. & ARMSTRONG, R. N. (2011). "Structure and function of YghU, a Nu-class glutathione transferase related to YfcG from escherichia coli". *Biochemistry* **50**, pp. 1274-1281.
- SKOPELITOU, K., DHAVALA, P., PAPAGEORGIU, A. C. & LABROU, N. E. (2012). "A glutathione transferase from *Agrobacterium tumefaciens* reveals a novel class of bacterial GST superfamily". *PLoS ONE* **7**, e34263.

- THOM, R., CUMMINS, I., DIXON, D. P., EDWARDS, R., COLE, D. J. & LAPTHORN, A. J. (2002). "Structure of a tau class glutathione S-transferase from wheat active in herbicide detoxification". *Biochemistry* **41**, pp. 7008-7020.
- THUILLIER, A., NGADIN, A. A., THION, C., BILLARD, P., JACQUOT, J.-P., GELHAYE, E. & MOREL M. (2011). "Functional Diversification of Fungal Glutathione Transferases from the Ure2p Class". *Int. J. Evol. Biol.* **2011**, 938308, pp. 1-9.
- THUILLIER, A., RORET, T., FAVIER, F., GELHAYE, E., JACQUOT, J. P., DIDIERJEAN, C. & MOREL-ROUHIER, M. (2013). "Atypical features of a Ure2p glutathione transferase from *Phanerochaete chrysosporium*". *FEBS Lett.* **587**, pp. 2125-30.
- TILLMAN, R. W., SIEGEL, M. R. & LONG, J. W. (1973). "Mechanism and fate of the fungicide chlorothalonil 2,4,5,6,-tetra-chlorothalonitrile in biological systems. Part 1: reactions with cells and subcellular components of *Saccharomyces pastorianus*. Pestic". *Biochem. Physiol.* **3**, pp. 160-167.
- TIPPING, E. & KETTERER, B. (1981). "The influence of soluble binding proteins on lipophile transport and metabolism in hepatocytes". *Biochem. J.* **195**, pp. 441-452.
- TOCHEVA, E. I., FORTIN, P. D., ELTIS, L. D. & MURPHY, M. E. P. (2006). "Structures of ternary complexes of BphK, a bacterial glutathione S-transferase that reductively dechlorinates polychlorinated biphenyl metabolites". *J. Biol. Chem.* **281**, pp. 30933-30940.
- TOMAREV, S. T., ZINOVIEVA, R. D., GUO K. & PIATIGORSKY J. (1993). "Squid glutathione S-transferase". *J. Biol. Chem.* **268**, pp. 4534-4542.
- UDOMSINPRASERT, R., PONGJAROENKIT, S., WONGSANTICHON, J., OAKLEY, A. J., PRAPANTHADARA, L. A., WILCE, M. C. & KETTERMAN, A. J. (2005). "Identification, characterization and structure of a new Delta class glutathione transferase isoenzyme". *Biochem. J.* **388**, pp. 763-71.
- UMLAND, T. C., TAYLOR, K. L., RHEE, S., WICKNER, R. B. & DAVIES, D. R. (2001). "The crystal structure of the nitrogen regulation fragment of the yeast prion protein Ure2p". *PNAS* **98**, pp. 1459-1464.
- VAGIN, A. & TEPLYAKOV, A. (2010). "Molecular replacement with MOLREP". *Acta D.* **66**, pp. 22-25.
- VINCE, R., DALUGE, S. & WADD, W. B. (1971). "Inhibition of glyoxalase I by S-substituted glutathiones". *J. Med. Chem.* **14**, 402-404.
- WADINGTON, M. C., LADNER, J. E., STOURMAN, N. V., HARP, J. M. & ARMSTRONG, R. N. (2009). "Analysis of the structure and function of YfcG from *Escherichia coli* reveals an efficient and unique disulfide bond reductase". *Biochemistry* **48**, pp. 6559-6561.

WHITBREAD, A. K., MASOUNI, A., TETLOW, N., SCHMUCK, E., COGGAN, M. & BOARD, P. G. (2005). "Characterization of the Omega Class of Glutathione Transferases". *Methods in Enzymology* **401**, pp. 78-99.

WIKTELIUS, E. & STENBERG, G. (2007). "Novel class of glutathione transferases from cyanobacteria exhibit high catalytic activities towards naturally occurring isothiocyanates". *Biochem. J.* **406**, pp. 115-123.

WILCE, M. C. J. & PARKER, M. W. (1994). "Structure and function of glutathione S-transferases". *Biochim. Biophys. Acta.* **1205**, pp.1-18.

WILCE, M. C., BOARD, P. G., FEIL, S. C. & PARKER, M. W. (1995). "Crystal structure of a theta-class glutathione transferase". *EMBO J.* **14**, pp. 2133-2143.

WINAYANUWATTIKUN, P. & KETTERMAN, A. J. (2004). "Catalytic and structural contributions for glutathione binding residues in a Delta class glutathione S-transferase". *Biochem. J.* **382**, pp. 751-757.

WINAYANUWATTIKUN, P. & KETTERMAN, A. J. (2005). "An electron-sharing network involved in the catalytic mechanism is functionally conserved in different glutathione transferase classes". *J. Biol. Chem.* **280**, pp. 31776-31782.

WINAYANUWATTIKUN, P. & KETTERMAN, A. J. (2007). "Glutamate-64, a newly identified residue of the functionally conserved electron-sharing network contributes to catalysis and structural integrity of glutathione transferases". *Biochem. J.* **402**, pp. 339-348.

WINN, M. D., BALLARD, C. C., COWTAN, K. D., DODSON, E. J., EMSLEY, P., EVANS, P. R., KEEGAN, R. M., KRISINEL, E. B., LESLIE, A. G. W., McCOY, A., MCNICHOLAS, S. J., MURSHUDOV, G. N., PANNU, N. S., POTTERTON, E. A., POWELL, H. R., READ, R. J., VAGIN, A. & WILSON, K. S. (2011). "Overview of the CCP4 suite and current developments". *Acta D.* **67**, pp. 235-242.

XIA, B., VLAMIS-GARDIKAS, A., HOLMGREN, A., WRIGHT, P. E. & DYSON, H. J. (2001). "Solution structure of *Escherichia coli* glutaredoxin-2 shows similarity to mammalian glutathione-S-transferases". *J. Mol. Biol.* **310**, pp. 907-918.

XUN, L., BELCHIK, S. M., XUN, R., HUANG, Y., ZHOU, H., SANCHEZ, E., KANG, C. & BOARD, P. G. (2010). "S-Glutathionyl-(chloro)hydroquinone reductases: A novel class of glutathione transferases". *Biochem. J.* **428**, pp. 419-427.

YAMADA, T., KOMOTO, J., WATANABE, K., OHMIYA, Y. & TAKUSAGAWA, F. (2005). "Crystal structure and possible catalytic mechanism of microsomal prostaglandin E synthase type 2 (mPGES-2)". *J. Mol. Biol.* **348**, pp. 1163-1176.

ZHANG, Z. R. & PERRETT, S. (2009). "Novel glutaredoxin activity of the yeast prion protein Zure2 reveals a native-like dimer within fibrils". *J. Biol. Chem.* **284**, pp. 14058-14067.

- ZHANG, Z. R., BAI, M., WANG, X. Y., ZHOU, J. M. & PERRETT, S. (2008). "Restoration" of Glutathione Transferase Activity By Single-site Mutation of The Yeast Prion Protein Ure2". *J. Biol. Chem.* **384**, pp. 641-651.
- ZHOU, H., BROCK, J., CASAROTTO, M. G., OAKLEY A. J. & BOARD, P. G. (2011). "Novel folding and stability defects cause a deficiency of human glutathione transferase omega 1". *J. Biol. Chem.* **286**, pp. 4271-4279.
- ZHOU, H., BROCK, J., LIU, D., BOARD, P. G. & OAKLEY A. J. (2012). "Structural insights into the dehydroascorbate reductase activity of human omega-class glutathione transferases". *J. Mol. Biol.* **420**, pp. 190-203.