STRUCTURAL BIOLOGY

ISSN: 2059-7983
journals.iucr.org/d

Structural and biochemical characterization of a glutathione transferase from the citrus canker pathogen *Xanthomonas*

Eduardo Hilario, Sawyer De Keyser and Li Fan

Acta Cryst. (2020). D**76**, 778–789

STRUCTURAL BIOLOGY

ISSN 2059-7983

Structural and biochemical characterization of a glutathione transferase from the citrus canker pathogen *Xanthomonas*

Eduardo Hilario, Sawyer De Keyser and Li Fan*

Department of Biochemistry, University of California-Riverside, Riverside, California, USA. *Correspondence e-mail: lifan@ucr.edu

Received 21 January 2020
Accepted 7 July 2020

Edited by A. Berghuis, McGill University, Canada

Keywords: glutathione *S*-transferases; GSTs; *Xanthomonas citri* subsp. *citri*; citrus canker; glutathione; bromoacetate; dehalogenases; *Xanthomonas*.

PDB references: XacGST, apo form, 6nxv; complex with glutathione, 6nv6

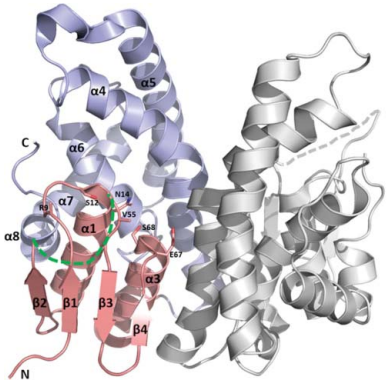
Supporting information: this article has supporting information at journals.iucr.org/d

The genus *Xanthomonas* comprises several cosmopolitan plant-pathogenic bacteria that affect more than 400 plant species, most of which are of economic interest. Citrus canker is a bacterial disease that affects citrus species, reducing fruit yield and quality, and is caused by the bacterium *Xanthomonas citri* subsp. *citri* (Xac). The Xac3819 gene, which has previously been reported to be important for citrus canker infection, encodes an uncharacterized glutathione *S*-transferase (GST) of 207 amino-acid residues in length (XacGST). Bacterial GSTs are implicated in a variety of metabolic processes such as protection against chemical and oxidative stresses. XacGST shares high sequence identity (45%) with the GstB dehalogenase from *Escherichia coli* O6:H1 strain CFT073 (EcGstB). Here, XacGST is reported to be able to conjugate glutathione (GSH) with bromoacetate with a K_m of 6.67 ± 0.77 m*M*, a k_{cat} of 42.69 ± 0.32 s^{-1} and a k_{cat}/K_m of 6.40 ± 0.72 mM^{-1} s^{-1} under a saturated GSH concentration (3.6 m*M*). These values are comparable to those previously reported for EcGstB. In addition, crystal structures of XacGST were determined in the apo form (PDB entry 6nxv) and in a GSH-bound complex (PDB entry 6nv6). XacGST has a canonical GST-like fold with a conserved serine residue (Ser12) at the GSH-binding site near the N-terminus, indicating XacGST to be a serine-type GST that probably belongs to the theta-class GSTs. GSH binding stabilizes a loop of about 20 residues containing a helix that is disordered in the apo XacGST structure.

1. Introduction

The genus *Xanthomonas* comprises a number of cosmopolitan plant-pathogenic bacteria that affect a wide range of hosts (Swings *et al.*, 1990). Among the phytopathogenic bacteria, the genus *Xanthomonas* comprises 20 species which together infect more than 400 plant species, including at least 124 monocotyledonous and 268 dicotyledonous plant species. *Xanthomonas* species have a narrow host range and are usually restricted to a specific plant genus or species. Most of these plants are economically important crops such as bananas, beans, cassava, cabbages, cotton, citrus, grapes, mangoes, peppers, rice, tomatoes and wheat (Leyns *et al.*, 1984; Mansfield *et al.*, 2012). Crop diseases impact both product quality and yield, resulting in economic, environmental and health issues worldwide (Strange & Scott, 2005). *Xanthomonas citri* subsp. *citri*, formerly named *Xanthomonas axonopodis* pv. *citri* (Xac), is the causal agent of citrus canker, which affects most commercial citrus cultivars. Typical symptoms include defoliation, fruit drop and reduced marketability of mature and symptomatic fruit.

The Ferro laboratory at FCAV-UNESP, Sao Paulo, Brazil has identified almost 80 Xac genes involved in adaptation, pathogenesis and virulence (Laia, 2007), and our laboratory



has been investigating the structure and function of these gene products (Hilario *et al.*, 2011, 2012). The open reading frame Xac3819 (UniProtKB accession No. Q8PG02) is considered to be one of the essential Xac genes required for plant–pathogen interactions. Xac3819 encodes a polypeptide of 207 amino-acid residues representing an uncharacterized and unclassified glutathione transferase (XacGST), historically also called glutathione *S*-transferase (EC 2.5.1.18; Ahmad *et al.*, 2015). The Xac genome contains 11 uncharacterized and unclassified GST genes (Xac0894, Xac1007, Xac1299, Xac1461, Xac1474, Xac2230, Xac2394, Xac2460, Xac3819, Xac4047 and Xac4352; da Silva *et al.*, 2002). The results from the Ferro laboratory suggest that among these Xac GSTs only Xac3819 is required for the citrus canker phenotype.

The induction of GST genes is an evolutionarily conserved response of cells to oxidative stress (Hayes *et al.*, 2005; Armstrong, 1991). GSTs are a ubiquitous family of multi-functional isozymes that are involved in cellular detoxification by catalyzing the conjugation of glutathione (GSH) to a large number of endogenous and exogenous compounds (Habig *et al.*, 1974). GSH is a tripeptide (γGlu-Cys-Gly) which protects living cells against oxidative stress as an essential compound for the detoxification of xenobiotic compounds (Habig *et al.*, 1974; Mannervik *et al.*, 1988): the conjugation of xenobiotics with GSH usually leads to the formation of less reactive products that are readily excreted (Hayes *et al.*, 2005; Armstrong, 1991). Overall, GSH is involved in the modulation of cell proliferation, the regulation of cell growth and cell death and the modulation of immune function, and participates in the metabolism of estrogens, leukotrienes and prostaglandins, the maturation of iron–sulfur clusters in diverse proteins, the transportation and storages of nitric oxide and the reduction of ribonucleotides to deoxy-ribonucleotides (Lu, 2013). GSH is the most important nonprotein thiol compound found in plants, playing a pivotal role in cellular signaling and defense in infected plants (Gullner *et al.*, 2018). GSTs conjugate GSH with xenobiotics by catalyzing the nucleophilic attack of GSH on electrophilic C, S or N atoms of nonpolar xenobiotic substrates to facilitate their inactivation, solubilization and excretion, and to prevent their interaction with crucial cellular proteins and nucleic acids (Ketterer & Meyer, 1989). GSTs have conferred the development of resistance against antibiotics, insecticides, herbicides, environmental pollutants, chemical carcinogens and other exogenous compounds in organisms ranging from prokaryotes to the most complex eukaryotes (Hayes *et al.*, 2005; Mannervik *et al.*, 1988; Hayes & Pulford, 1995; Pavlidi *et al.*, 2018). Bacterial GSTs are implicated in a variety of metabolic processes such as protection against chemical and oxidative stresses, antimicrobial drug resistance and the degradation of lignin (Gullner *et al.*, 2018; Allocati *et al.*, 2009, 2012).

GSTs are grouped into four superfamilies; cytosolic, mitochondrial (kappa-class enzymes), microsomal and bacterial fosfomycin-resistance protein (Mannervik *et al.*, 2005). The cytosolic GSTs are a large and diverse superfamily of enzymes. They have been subgrouped into 16 Greek letter-labeled classes (alpha, beta, chi, delta, epsilon, iota, lambda, mu, nu, theta, omega, pi, phi, sigma, tau and zeta) based on their chemical, physical and structural properties (Sheehan *et al.*, 2001; Hayes *et al.*, 2005; Oakley, 2011; Allocati *et al.*, 2012). Bacterial and yeast genome-sequencing projects have been useful for the identification of large numbers of GSTs of unknown function in microorganisms (with as many as 30 in a single organism; Vuilleumier & Pagni, 2002; Skopelitou, Dhavala *et al.*, 2012; Skopelitou, Muleta *et al.*, 2012). However, the majority of GST sequences deposited in public databases unfortunately have not been experimentally characterized or even assigned to one of the canonical classes (Atkinson & Babbitt, 2009). They were automatically annotated (owing to successive sequence depositions from genome-sequencing projects) using computational analysis by simple inference of similarity between the target sequences and sequences previously deposited in the database. Thus, it is possible that some of these uncharacterized putative GSTs may not be glutathione transferase enzymes, but rather proteins with sequence similarities to GSTs (Vuilleumier, 1997).

Despite the low amino-acid sequence identity among GST superfamily members, the overall three-dimensional protein fold is remarkably similar among the different GST classes (Sheehan *et al.*, 2001). GSTs typically consist of two domains named the N-terminal and C-terminal domains, with the enzyme active site structured in a cleft between these two domains. The GST active site can be divided into two subsites: the GSH-binding site (G-site) and the hydrophilic binding site (H-site). The N-terminal domain contains the G-site, which has a conserved cysteine, serine or tyrosine residue to activate the GSH cysteinyl side chain for nucleophilic attack of GSH on nonpolar compounds containing an electrophilic C, N or S atom (Hayes *et al.*, 2005). The nonpolar compound is bound at the nearby H-site in the C-terminal domain. Therefore, GSTs have also been classified into two groups based on their ability to activate the GSH cysteinyl side chain at the G-site: the tyrosine-type GSTs (Y-GSTs) and the serine/cysteine-type GSTs (S-GSTs and C-GSTs). Additional key residues from both the G-site and H-site are essential for GSH–substrate conjugation (Sheehan *et al.*, 2001). For example, a highly conserved Pro53 residue together with a conserved Asp155 residue are important in maintaining the protein in a catalytically competent structure (Wilce *et al.*, 1995; Wilce & Parker, 1994; Allocati *et al.*, 1999). Not surprisingly, the G-site amino-acid residues in the N-terminal domain are relatively conserved but the H-site residues show a high degree of variability in different GSTs with varying substrate specificities (Dirr *et al.*, 1994).

Microbial GSTs have been ignored for a long time since they display very low activity towards the standard GST substrates of mammalian and plant enzymes (Marsh *et al.*, 2008). Bacterial GST genes are often located within gene clusters involved in the degradation of various aromatic compounds, suggesting important roles in dehalogenation, detoxification and the mineralization of environmental pollutants and their degradation products (Marsh *et al.*, 2008; Fuenmayor *et al.*, 1998; Bartels *et al.*, 1999; Daubaras *et al.*,

1995; Hofer *et al.*, 1994; Vlieg *et al.*, 2000). However, very limited information is available about the role of GSTs during plant–pathogen interaction (Gullner *et al.*, 2018). There are no reports to date on the biochemical or structural characterization of any GSTs from the infamous plant-pathogen genus *Xanthomonas*. Here, we present for the first time crystal structures of XacGST (Xac3819) with and without GSH bound at the N-terminal catalytic site and characterize the kinetics of XacGST as a dehalogenase on bromoacetate. We hypothesize that XacGST is an essential GST that protects Xac against cellular oxidative stress during plant–pathogen interaction and might be an essential virulence factor in Xac pathogenesis acting against phytoalexins and reactive oxygen species accumulated at the infection interface.

2. Methods

2.1. XacGST DNA cloning

The oligonucleotide primers XAC3819-BamHI-FW (5′-cgc GGATCCATGCCCGCCATCACGATC-3′) and XAC3819-EcoRI-REV (5′-ccgGAATTCTCAAGTCAGCGCGGTCATC-3′) were used to amplify the XAC3819 (XacGST) open reading frame of 624 base pairs from the genomic DNA of *X. citri* subsp. *citri* strain 306 (NCBI txid190486). The underlined sequences correspond to a restriction-digestion site and the lower-case sequences are additional bases to enhance cleavage close to the end of DNA fragments by a specific restriction enzyme. The amplified DNA fragment was digested with BamHI and EcoRI, gel-extracted (0.8% agarose gel in 1× TAE buffer; Ultra-Sep gel extraction kit from Omega Biotek) and then ligated with a modified pET SUMO expression vector previously digested with the same enzymes. The ligation product was used to transform competent cells of *Escherichia coli* strain DH10B. Several recombinant colonies were selected from Luria–Bertani (LB) agar plates and amplified in 6 ml liquid LB medium. Recombinant plasmid samples were digested with BamHI/EcoRI to confirm the integration of the insert into the pET SUMO expression vector. A positive recombinant plasmid construct pET-SUMO-XacGST was DNA-sequenced and used to transform *E. coli* strain Rosetta(DE3)pLysS competent cells. All molecular biology steps were performed as described in Sambrook & Russell (2001).

2.2. XacGST protein expression and purification

A single colony of *E. coli* strain Rosetta(DE3)pLysS cells harboring the pET-SUMO-XacGST construct was incubated in 50 ml LB culture medium containing 35 mg ml^{-1} kanamycin and 35 mg ml^{-1} chloramphenicol for 15 h at 37°C with shaking at 225 rev min^{-1}. The overnight cell culture was diluted 100 times in 4 l fresh LB medium containing antibiotics and grown under the same conditions until the OD_{600} reached 0.6. At this point, isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to a final concentration of 0.2 m*M* to induce overexpression for 16 h at 25°C. The cells were then centrifuged at 5000*g* for 20 min at 4°C and the cell pellet was resuspended in 80 ml 100 m*M* Tris–HCl pH 8.0 containing 500 m*M* sodium chloride, 5% glycerol, 30 m*M* imidazole–HCl pH 8.0. Cell lysis was performed by sonication (Branson 450 Digital Sonifier) using 50% duty power in an ice–water bath. The crude extract was clarified by centrifugation (30 000*g*, 20 min, 4°C). Recombinant His_6-SUMO-XacGST protein was initially purified using a 10 ml Ni–NTA gravity-flow column at room temperature. The supernatant was loaded onto the affinity column and nonspecifically bound proteins were removed by washing with lysis buffer. Large amounts of soluble His_6-SUMO-XacGST were eluted in 100 ml 20 m*M* Tris–HCl pH 8.0 containing 200 m*M* sodium chloride, 5% glycerol and 500 m*M* imidazole–HCl pH 8.0. Ammonium sulfate was added to 40% saturation at 25°C and the mixture was centrifuged (30 000*g*, 20 min, 25°C). The pellet fraction containing large amounts of near-pure His_6-SUMO-XacGST was dissolved in 25 ml 25 m*M* Tris–HCl pH 8.0 containing 100 m*M* sodium chloride and 10% glycerol. Recombinant SUMO protease was added (in a 1:1000 ratio) and the mixture was incubated for 4 h at 25°C. The digestion product was loaded onto a 10 ml Ni–NTA gravity-flow column at 25°C and the flowthrough sample containing nontagged XacGST was collected and concentrated using 10 kDa Amicon Ultra-15 centrifugal filters (Merck Millipore). Concentrated aliquots at 15 mg ml^{-1} in 25 m*M* Tris–HCl pH 8.0 containing 300 m*M* sodium chloride, 5% glycerol and 1 m*M* GSH were loaded onto a HiPrep 16/60 Sephacryl S-100 HR column (GE Healthcare, catalog No. 17-1165-01) connected to a ÄKTA-purifier system (GE Healthcare). The peak fractions containing large amounts of pure XacGST were collected and concentrated to 10 mg ml^{-1}. Protein aliquots were prepared, flash-frozen and stored at −20°C prior to protein crystallization trials. All protein samples were stored at −20°C and samples of interest were analyzed using sodium dodecyl sulfate polyacrylamide gel electrophoresis (15% SDS–PAGE gel) stained with Coomassie Blue R-250 (Laemmli, 1970). The protein concentration was determined by the Bradford method using bovine serum albumin as a standard (Bradford & Williams, 1976).

2.3. XacGST protein crystallization and structure determination

To select an appropriate XacGST protein concentration for crystallization screening, we used the PCT Pre-Crystallization Test (Hampton Research, catalog No. HR2-140). The MCSG Crystallization Suite (MCSG-1–MCSG-4; Anatrace), Art Robbins Intelli-Plates 96-3 and an Art Robbins Phoenix dispensing system were used to evaluate XacGST protein crystallization at 25°C. Small protein crystals were initially observed in several conditions containing bis-Tris–HCl buffer, PEG 3350 and magnesium chloride. The best XacGST crystals were obtained by sitting-drop vapor diffusion at 25°C by mixing 2 µl protein sample at 7.5 mg ml^{-1} with an equal volume of reservoir solution consisting of 100 m*M* bis-Tris–HCl pH 7.0–7.25, 200 m*M* magnesium chloride, 25–27%(*w*/*v*) PEG 3350. Crystals appeared in 2–3 days and grew for up to

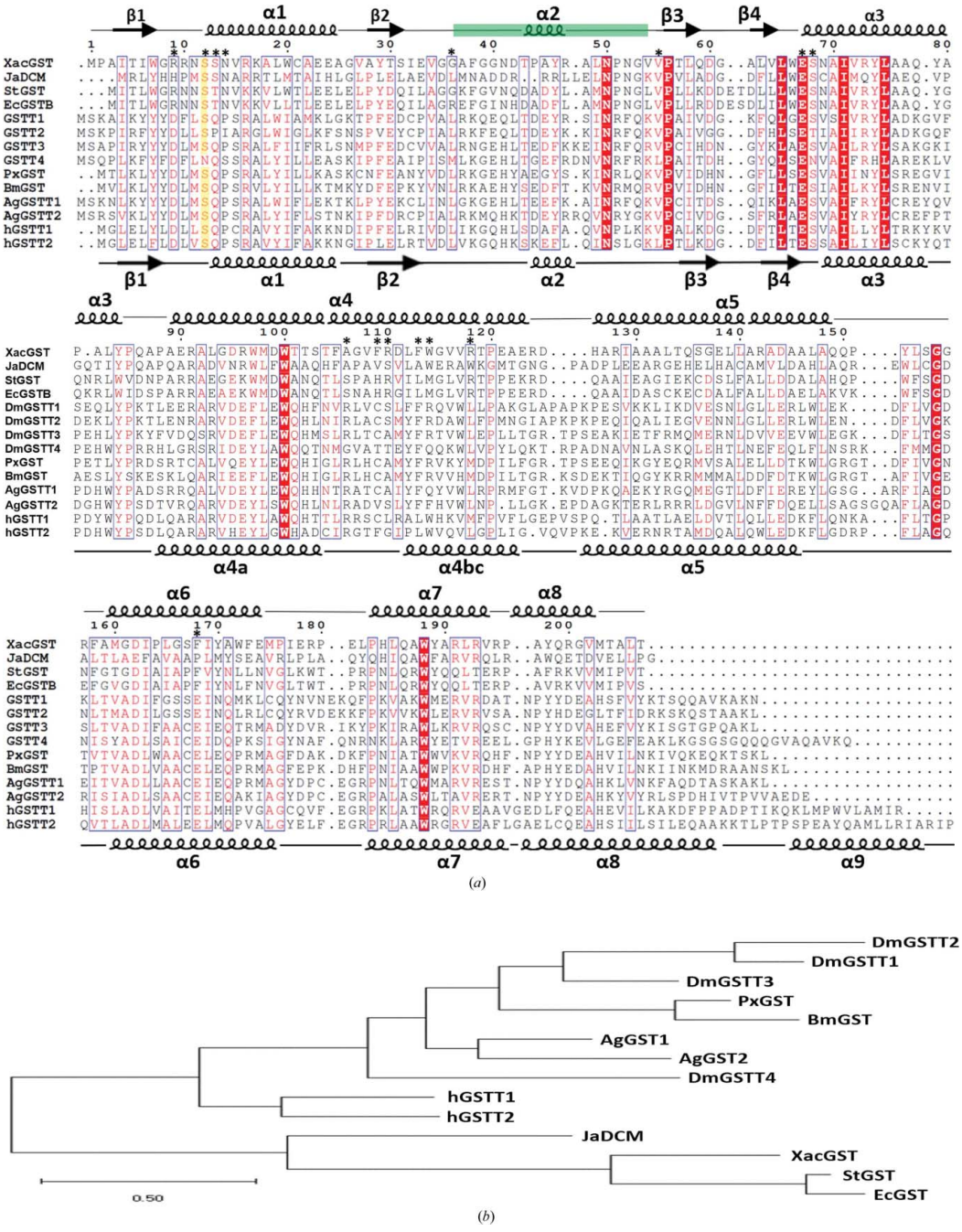


Figure 1
Multiple sequence alignment of XacGST with theta-class GST members and their evolutionary relationship. (*a*) Sequence alignment was performed using *ESPript* 3.0 (Robert & Gouet, 2014) utilizing the *ClustalW* algorithm (Thompson *et al.*, 1994). Sequence aliases with the corresponding UniProtKB accession numbers for classified theta-class GSTs are as follows: XacGST, *Xanthomonas citri* substr. *citri* strain 306, Q8PG02; JaDCM, *Janthinobacterium* sp. KBS0711 dichloromethane dehalogenase, A0A0M2WQ11; StGST, uncharacterized and unclassified YliJ from *Salmonella typhimurium*, Q8Z859; EcGSTB, unclassified GstB from *Escherichia coli* O6:H1 strain CFT073, P0ACA8; DmGSTT1, DmGSTT2, DmGSTT3 and DmGSTT4, *Drosophila melanogaster* GST T1, T2, T3 and T4, Q7K0B6, A1Z7X7, E1JJS1 and Q8MRM0, respectively; PxGST, *Papilio xuthus* GST, I4DNY4; BmGST, *Bombyx mori* GST, B0LB14; AgGSTT1 and AgGSTT2, *Aedes aegypti* GST T1 and T2, Q8MUQ1 and Q8MUQ2, respectively; hGSTT1 and hGSTT2, *Homo sapiens* GST theta-1 and theta-2, P30711 and P0CG29, respectively. Strictly conserved residues among GSTs are highlighted with a red background and highly homologous residues are shown in blue boxes. Strictly conserved class-dependent residues from the GSH catalytic site of the serine-type GSTs (theta class) are highlighted with a yellow background. The secondary-structural elements of XacGST (PDB entry 6nv6) and hGSTT2 (PDB entry 4mpg) defined by the crystal structure are indicated as black coils and arrows at the top and bottom of the sequence alignment, respectively. The unstructured loop region (residues 35–53) observed in the GSH-free monomers is represented as a green bar. The residues marked with asterisks in the XacGST sequence are related to GSH binding. (*b*) Phylogenetic analysis of the amino-acid sequences of theta-class GSTs. The scale bar is equal to 0.5 substitutions per site and the length of each branch is proportional to the average number of substitutions per site as indicated by the scale. The evolutionary history was inferred using the maximum-likelihood method based on the JTT matrix-based model (Jones *et al.*, 1992) and the evolutionary analyses were conducted in *MEGA X* (Kumar *et al.*, 2018)

6–7 days as rod-like or needle-like crystals. A single crystal was soaked in cryoprotectant consisting of its reservoir solution containing up to 30% DMSO, mounted on a nylon loop (Hampton Research, USA) and then flash-cooled in liquid nitrogen. Diffraction data were collected under gaseous nitrogen at −175°C (Cobra, Oxford Cryosystems) using X-rays at a wavelength of 1.5418 Å (Rigaku MicroMax-007 HF) and a Rigaku R-AXIS IV++ image-plate detector at the Department of Biochemistry at University of California-Riverside. *CrystalClear* 2.0 (Rigaku) was used during data collection. X-ray diffraction intensities were processed using *iMosflm* (Battye *et al.*, 2011) and scaled with *SCALA* (Evans, 2006), both implemented in the *CCP*4 package (Winn *et al.*, 2011). A molecular-replacement search model was initially generated with PDB entry 4kh7 (Enzyme Function Initiative, unpublished work) and *CHAINSAW* (Stein, 2008). Molecular replacement was performed in *Phaser* (McCoy *et al.*, 2007). Crystal structure refinement was performed by iterative cycles of manual inspection of $2F_o - F_c$ and $F_o - F_c$ electron-density maps with *Coot* (Emsley *et al.*, 2010) followed by structure refinement using the *Phenix* package (Liebschner *et al.*, 2019). The analysis of interacting residues at the protein interface, amino-acid interactions, inaccessible amino-acid residues, solvent-accessible residues and calculations of solvation energy and solvation-energy gain upon interface formation were performed using the *PDBePISA* (*Protein Interfaces, Surfaces and Assemblies*) service at the European Bioinformatics Institute (http://www.ebi.ac.uk/pdbe/prot_int/pistart.html; Krissinel & Henrick, 2007) using the coordinate files of XacGST with GSH bound (PDB entry 6nv6, chains *A* and *B*) and of YliJ from *Salmonella typhimurium* (PDB entry 4kh7, chains *A* and *B*).

2.4. XacGST dehalogenase activity

XacGST enzyme activity with the substrate bromoacetate was measured via a discontinuous spectrophotometric assay using 5,5′-dithiobis-(2-nitrobenzoic acid), also known as DTNB or Ellman's reagent. A 5 m*M* DTNB stock solution was freshly prepared in ethanol. A 200 m*M* bromoacetate and 50 m*M* GSH stock solution in 50 m*M* sodium phosphate pH 7.0 was prepared and the pH was adjusted with NaOH. Prior to the start of each experiment, a quartz cuvette containing 950 µl 200 m*M* Tris–HCl pH 8.0, 50 µl 5 m*M* DTNB was used to blank the spectrophotometer. For each uncatalyzed reaction, we used a final bromoacetate concentration ranging from 0 to 30.0 m*M* (0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 and 30 m*M*) and kept the final concentration of GSH at 3.6 m*M* in a 200 µl mixture buffered in 50 m*M* sodium phosphate pH 7.0. No XacGST was added to this mixture and bromoacetate was added last. Each reaction was incubated at room temperature for 2 min. After this, a 25 µl aliquot of uncatalyzed reaction mixture was combined with 925 µl 200 m*M* Tris–HCl pH 8.0 and 50 µl 5 m*M* DTNB. A 1.0 cm path-length quartz cuvette was used and absorbance measurements were recorded at 412 nm using a GE Healthcare GeneQuant 100 spectrophotometer; 412 nm is the λ_{max} for the 2-nitro-5-thiobenzoate anion (TNB^{2-}; ε = 13 600 M^{-1} cm^{-1}). All reactions were performed in triplicate. For the catalyzed reaction, XacGST was added to the reaction mixture. The final concentrations of bromoacetate, XacGST and GSH were 0–30 m*M*, 3.6 m*M* and 1.3 µ*M*, respectively. The absorbance measurements in both the uncatalyzed and the catalyzed experiments were used to calculate the initial and final concentration of GSH for each bromoacetate concentration according to Beer's law: $A_{412} = \varepsilon\ (M^{-1}\ cm^{-1}) \times L\ (cm) \times$ dilution factor $\times$ 1000. The GSH concentration after 2 min reaction (1–30 m*M* bromoacetate) was subtracted from the initial GSH concentration (0 m*M* bromoacetate). The corresponding value was divided by the reaction time (120 s) to obtain the rate ($\Delta C/t$). Finally, to obtain the adjusted reaction rate of GSH (in µ*M*), the rate of the uncatalyzed reaction was then subtracted from the rate of the catalyzed reaction and the final value was multiplied by 1000. A graph showing the concentration of bromoacetate (m*M*) versus the adjusted reaction rate (µ*M* s^{-1}) was plotted. A Lineweaver–Burk plot was created by plotting the inverse initial velocity ($1/V_0$; *x* axis) as a function of the inverse of the substrate concentration (1/[S]; *y* axis), where $1/V = \{K_m/(V_{max} \times [S])\} + (1/V_{max})$.

3. Results and discussion

3.1. XacGST is a serine-type GST of the theta class with dehalogenase activity

The XacGST sequence contains an open reading frame of 624 base pairs, which encodes a GST with 207 residues (Fig. 1). The deduced XacGST (Xac3819) amino-acid sequence shows the highest sequence identity of 47% to the unclassified StGST (YliJ) from *S. typhimurium* (UniProtKB accession code Q8Z859; PDB entry 4kh7) and 45% identity to GstB from *E. coli* O6:H1 strain CFT073 (EcGstB; UniProtKB accession code P0ACA8; no crystal structure available) (Fig. 1). In addition, the XacGST amino-acid sequence also shows 27%, 29%, 24% and 23% sequence identity to the insect theta-class *Drosophila melanogaster* GSTs T1, T2, T3 and T4 (DmGSTT1, DmGSTT2, DmGSTT3 and DmGSTT4; UniProtKB accession codes Q7K0B6, A1Z7X7, E1JJS1 and Q8MRM0, respectively), 27% identity to both *Aedes aegypti* GST T1 and T2 (AgGSTT1 and AgGSTT2; UniProtKB accession codes Q8MUQ1 and Q8MUQ2, respectively), 28% and 26% identity to *Papilio xuthus* PxGST and *Bombyx mori* BmGSTT (UniProtKB accession codes I4DNY4 and B0LB14, respectively) and 29% and 26% identitiy to the human theta-class GSTs hGSTT1 and hGSTT2, respectively (Fig. 1). The sequence conservation increases in the N-terminal region since this region contains the G-site and is evolutionarily conserved. Compared with the full-length XacGST, the first 85 N-terminal residues of XacGST have a much higher sequence identity to the bacterial StGST (60%), EcGstB (53%) and JaDCM (*Janthinobacterium* sp. KBS0711 dichloromethane dehalogenase; 41%), the insect AgGSTT2 (45%), AgGSTT1 (33%), DmGSTT1 (35%), DmGSTT2 (39%), DmGSTT3 (35%), DmGSTT4 (29%), PxGST (29%) and BmGST (30%),

and the human hGSTT1 (41%) and hGSTT2 (33%). All of these GSTs except DmGSTT4 contain a serine residue at the GSH-binding site (Fig. 1*a*). These results suggest that XacGST is likely to be a member of the theta-class GSTs since the theta-class GST members have previously been reported to have sequence identities of between 13% and 61% (Vuilleumier, 1997). We analyzed the hypothetical evolutionary history among these theta-class GSTs (Fig. 1*b*). It is evident from the phylogenetic tree that the theta class is a heterogenic group and that members from different kingdoms cannot be clustered together owing to the independent and parallel evolution of GST genes from ancient ancestors (Vuilleumier, 1997). The bacterial XacGST shares the same evolutionary clade with JaDCM, StGST and EcGstB, suggesting that they were very closely related to each other during the course of evolution and that they are are likely to have evolved from the same ancestral progenitor.

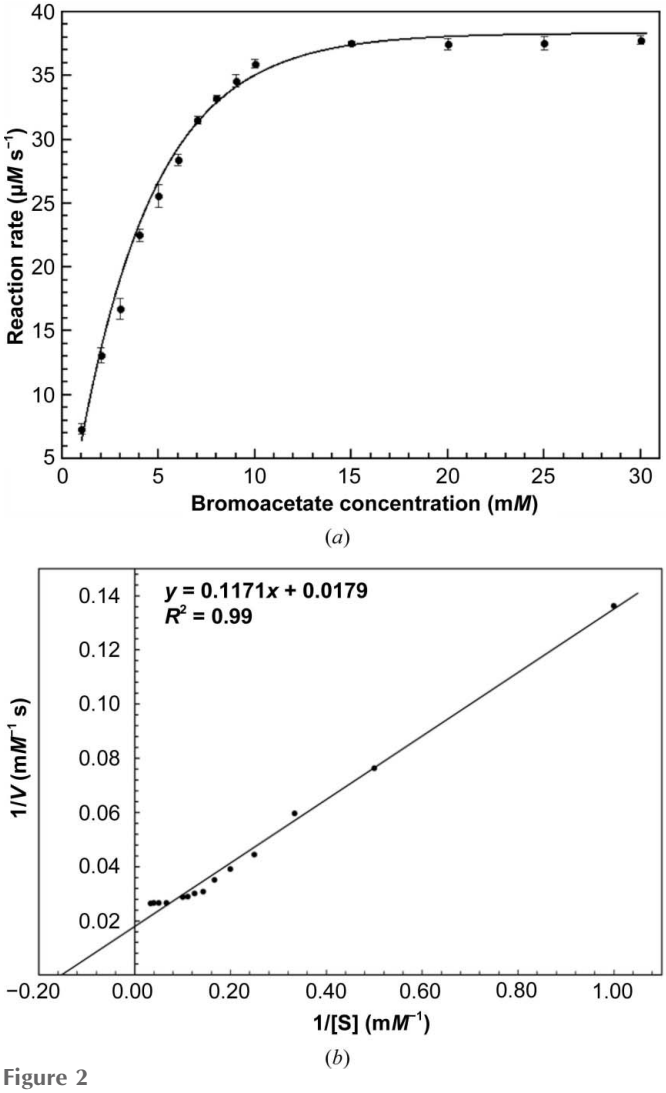


Figure 2
XacGST dehalogenase activity using bromoacetate as a substrate. (*a*) A plot of the rate of bromoacetate–GSH conjugation versus bromoacetate concentration. The experiments were carried out in triplicate. Error bars refer to the standard deviations. (*b*) Lineweaver–Burk plot for XacGST, where [S] represents the substrate concentration of bromoacetate and *V* represents the initial reaction velocity.

E. coli GstB has previously been reported to be involved in bromoacetate resistance via a dehalogenation mechanism (Desai & Miller, 2010) and was the first bacterial GST that was found to confer resistance to high concentrations of arsenate in solution (Chrysostomou *et al.*, 2015). *In vitro*, EcGstB has been shown to conjugate GSH with both arsenate (Chrysostomou *et al.*, 2015) and bromoacetate (Desai & Miller, 2010). Since XacGST shares high sequence similarity with EcGstB, we tested whether XacGST acts a dehalogenase on bromoacetate using recombinant XacGST expressed and purified from *E. coli*. Kinetic analysis (Fig. 2) revealed a maximal velocity (V_{max}) of 37.96 ± 0.28 m*M* s^{-1}, a Michaelis constant (K_m) of 6.67 ± 0.77 m*M*, a turnover number (k_{cat}) of 42.69 ± 0.32 s^{-1} and a catalytic efficiency (k_{cat}/K_m) of 6.40 ± 0.72 mM^{-1} s^{-1} under a saturating glutathione concentration (3.6 m*M* GSH) in 200 m*M* Tris–HCl pH 8.0. These results are comparable to those reported for EcGstB under the same conditions by Moore (2017): a K_m of 3.50 m*M*, a k_{cat} of 33.40 s^{-1} and a k_{cat}/K_m of 10.00 mM^{-1} s^{-1} under a saturating glutathione concentration (3.6 m*M* GSH) in 200 m*M* Tris–HCl pH 8.0. Similar results were also reported under similar conditions: a K_m of 5.00 m*M*, a k_{cat} of 27 s^{-1} and a k_{cat}/K_m of 5.40 mM^{-1} s^{-1} under a saturating glutathione (1.5 m*M*) concentration in 100 m*M* HEPES–NaOH pH 7.0 (Desai & Miller, 2010).

3.2. The crystal structure of XacGST and its complex with GSH

To further characterize XacGST, we determined crystal structures of XacGST in the absence and presence of GSH by the molecular-replacement method using a modified model of StGST (PDB entry 4kh7) as the template. The XacGST crystal structure without GSH (PDB entry 6nxv) is shown in Figs. 3(*a*) and 3(*b*) and that with GSH bound (PDB entry 6nv6) is shown in Figs. 3(*c*) and 3(*d*). Data-processing and structure-refinement statistics are summarized in Table 1. Under both conditions, XacGST crystallized in space group $P2_12_12_1$ with two independent homodimers (four molecules) in the asymmetric unit (Figs. 3*a*–3*d*). The unit-cell parameters were *a* = 87.16, *b* = 105.50, *c* = 124.32 Å for apo XacGST (PDB entry 6nxv) and *a* = 89.12, *b* = 106.51, *c* = 125.23 Å for the GSH-bound complex (PDB entry 6nv6). The GSH binding does not affect the crystal packing. The two subunits of the XacGST homodimer in both crystal structures are related by twofold symmetry, as observed in most GST crystal structures.

The XacGST structure displays the canonical GST fold with two distinct domains: an N-terminal domain composed of a mixture of β-strands and α-helices and an all-α-helical C-terminal domain. The N-terminal domain (residues 1–85) has a thioredoxin-like structure mainly consisting of a four-stranded β-sheet, with one side having two α-helices (α1 and α3) that interact with the C-terminal domain (Figs. 3*b* and 3*d*). The N-terminal domain begins with the first strand (β1; residues 4–7), followed by the first α-helix (α1; residues 12–26), the second β-strand (β2; residues 29–32) and a flexible loop region (residues 35–53) connecting strand 2 to strand 3. The second α-helix connects to the third strand (β3; residues

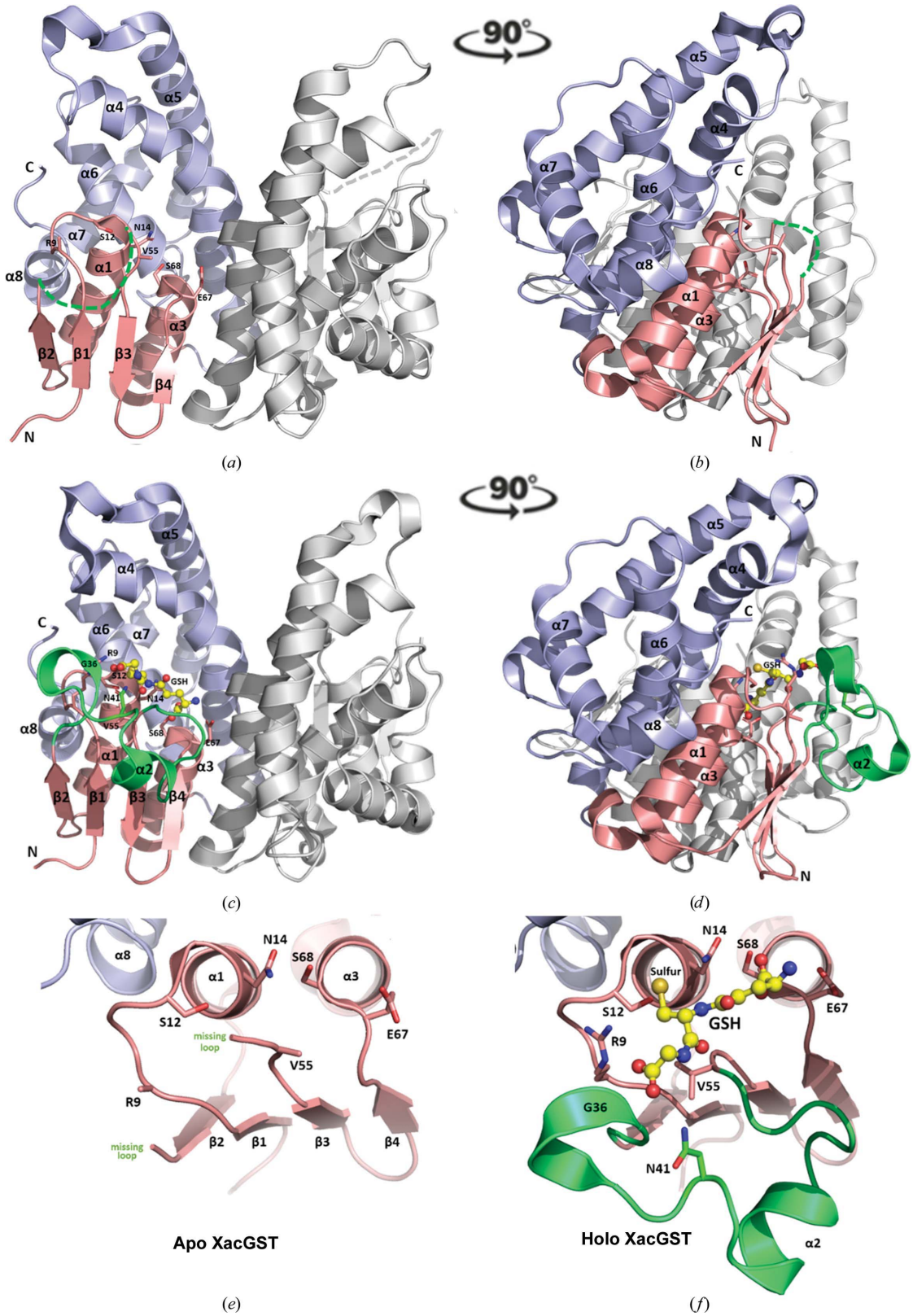


Figure 3
Crystal structures of XacGST and its complex with GSH. (*a*, *b*) XacGST dimer (PDB entry 6nxv, chains *A* and *B*) showing an N-terminal thioredoxin-like domain (*β*1–*α*1–*β*2–*α*2–*β*3–*β*4–*α*3 topology, salmon) and a C-terminal all-*α*-helical domain (*α*4–*α*8 topology, light blue). The partner monomer is colored light gray. (*c*, *d*) XacGST dimer (PDB entry 6nv6, chains *A* and *B*) containing GSH bound in one enzyme active site (PDB entry 6nv6, chain *B*). (*e*, *f*) Enlarged views of the flexible loop region of apo XacGST (*e*) and the XacGST–GSH complex (*f*). The loop between strands 2 and 3 contains helix *α*2 in the GSH-bound monomer and is represented by green or gray dashed lines in GSH-free monomers, where it is disordered. The bound GSH molecule is shown as a ball-and-stick representation and is colored according to atom type: C in yellow, O in red, N in blue and S in orange. Key residues interacting with GSH are highlighted with sticks and labeled. This figure was produced with *PyMOL* (DeLano, 2002) and manually modified.

Table 1
X-ray diffraction data-processing and model-refinement statistics.

Values in parentheses are for the highest resolution shell.

	XacGST	XacGST–GSH
PDB code	6nxv	6nv6
Data collection		
X-ray source	MicroMax-007 HF	MicroMax-007 HF
Wavelength (Å)	1.5418	1.5418
Space group	$P2_12_12_1$ [No. 19]	$P2_12_12_1$ [No. 19]
a, b, c (Å)	87.16, 105.50, 124.32	89.12, 106.51, 125.23
α, β, γ (°)	90.00, 90.00, 90.00	90.00, 90.00, 90.00
Rotation (°)	135	125.5
No. of images	270	249
No. of reflections	159325 (23027)	171185 (23962)
No. of unique reflections	30033 (4282)	35328 (5096)
Resolution range (Å)	30.00–2.75 (2.90–2.75)	30.00–2.65 (2.79–2.65)
R_{merge}† (%)	12.1 (68.6)	10.9 (63.9)
Mean $I/\sigma(I)$	8.9 (2.0)	10.7 (2.0)
Completeness (%)	98.8 (98.1)	99.9 (100.0)
Multiplicity	5.3 (5.4)	4.8 (4.7)
Mosaicity (°)	0.98	0.82
$CC_{1/2}$	0.994 (0.752)	0.993 (0.632)
Refinement statistics		
Resolution (Å)	29.81–2.75	30.00–2.65
Total reflections	29941	35272
R_{work}/R_{free}‡ (%)	19.37/24.13	17.69/22.59
No. of non-H atoms		
Protein	5814	6042
Cl^- ion	24	48
GSH	—	20
DMS	16	—
Water	245	397
Average B factors (Å^2)		
Protein	46.41	44.80
Cl^- ion	72.25	70.58
GSH	—	57.76
DMS	84.55	—
Water	48.20	44.27
R.m.s. deviatons from ideal geometry		
Bond lengths (Å)	0.004	0.006
Bond angles (°)	0.643	0.870
Ramachandran plot		
Favored regions	708 [95.16%]	743 [97.25%]
Allowed regions	36 [4.84%]	21 [2.75%]
Disallowed	0 [0.00%]	0 [0.00%]

† $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ and $\langle I(hkl)\rangle$ are the observed individual and mean intensities of a reflection with indices hkl, respectively, $\sum_i$ is the sum over i measurements of a reflection with indices hkl and $\sum_{hkl}$ is the sum over all reflections. ‡ $R = \sum_{hkl} ||F_{obs}| - |F_{calc}|| / \sum_{hkl} |F_{obs}|$. R_{free} is the R value calculated for 5% of the data set that was not included in the refinement.

57–60), followed by a fourth strand (β4; residues 63–65) and a third α-helix (α3; residues 67–80). Strands 1, 2 and 4 are parallel to each other, while strand 3 is antiparallel. A small loop (residues 84–87) links the N-terminal and C-terminal domains.

Owing to weak electron density, the flexible loop region (residues 35–53) connecting strand 2 to strand 3 could not be traced at a map contour level scaled at 1.0σ in any monomer of the apo XacGST structure (PDB entry 6nxv; Figs. 3*a* and 3*b*; dashed loop colored green for chain *A* or gray for chain *B*). In the four XacGST molecules in the asymmetric unit, apo XacGST (PDB entry 6nxv) has the following missing residues: 37–51 in chain *A*, 34–53 in chain *B*, 37–52 in chain *C* and 35–51 in chain *D*. Fortunately, crystals prepared from XacGST co-purified with GSH allowed us to trace this flexible loop, which contains a short helix (α2; residues 43–48) and a GSH molecule bound in the G-site (PDB entry 6nv6, chain *B*) as shown in Figs. 3(*c*), 3(*d*) and 3(*f*). To our surprise, only one of the four XacGST molecules in the asymmetric unit had a bound GSH, which is important to stabilize this flexible loop since the other three XacGST monomers without bound GSH in the same crystal have a disordered loop: residues 36–53 in chain *A*, 38–53 in chain *C* and 37–53 in chain *D* are missing.

The C-terminal domain (residues 84–207) contains five α-helices (α4–α8). The fourth α-helix (α4; residues 87–119) and fifth α-helix (α5; residues 125–150) together form an up–down arrangement. An eight-residue loop region (residues 151–158) connects the fifth α-helix to the sixth α-helix (α6; residues 159–175), which runs back near-parallel to the fourth α-helix. A seven-residue loop (residues 175–181) connects the sixth α-helix to the seventh α-helix (α7; residues 182–194), which lies in a perpendicular plane to all of the other α-helices, followed by a short eighth α-helix (α8; residues 195–202), which runs almost parallel to the first α-helix.

The N- and C-terminal domains of apo XacGST are almost identical to those of the XacGST–GSH complex. A structural superposition of C^α atoms of the apo XacGST (PDB entry 6nxv, chains *A* and *B*) and GSH-bound XacGST (PDB entry 6nv6, chains *A* and *B*) dimers yielded a root-mean-square deviation (r.m.s.d.) of 0.392 Å for 379 of 397 aligned amino-acid residues, indicating that the structures are virtually identical except for the loop region between strands 2 and 3 (residues 36–53). The XacGST dimers in both structures have approximate dimensions of 60 × 50 × 50 Å. XacGST dimerization occurs through an intersubunit interaction between helix α3 in the N-terminal domain of the first monomer and helix α4 in the C-terminal domain of the partner monomer, and vice versa. We used the *PISA* server to analyze the amino-acid residues involved in dimerization, salt bridges and hydrogen bonds and the inaccessible residues buried in the interior of XacGST with bound GSH. The residues involved in these interactions are summarized in Table 2. The solvent-accessible surface areas of monomers *A* and *B* (PDB entry 6nv6, chains *A* and *B*) are 1091.7 of 9102.4 Å^2 (12.0%) and 1107.2 of 9544.8 Å^2, respectively. The solvation energy and its gain upon dimerization for monomer *A* are −187.1 and −2.3 kcal mol^{-1} (1.2%) and those for monomer *B* are −194.5 and −2.4 kcal mol^{-1} (1.2%), respectively.

3.3. The interactions of XacGST with GSH

The crystal structure of XacGST in complex with GSH showed strong electron density for a GSH molecule in an extended conformation bound in the G-site via noncovalent interactions including electrostatic interactions and hydrogen bonds (Figs. 3*f* and 4*a*). The bound GSH is partially exposed to the solvent and is orientated almost parallel to the helical axis of α4 (along residues 104–114), above the beginning of α1 (residues 11–13) and at the end of α3 (residues 67–68). The G-site identified in XacGST includes amino-acid residues Arg9, Ser12, Asn14, Gly36, Val55, Glu67 and Ser68. The most important and conserved residue in theta-class GSTs is the serine residue involved in transferase activity (Ser12 in

XacGST), which is at a hydrogen-bond distance from the thiol group of the bound GSH (Sheehan *et al.*, 2001). It is well known that GSTs can decrease the pK_a of the sulfhydryl group of the bound GSH. The ionization of GSH to give a thiolate

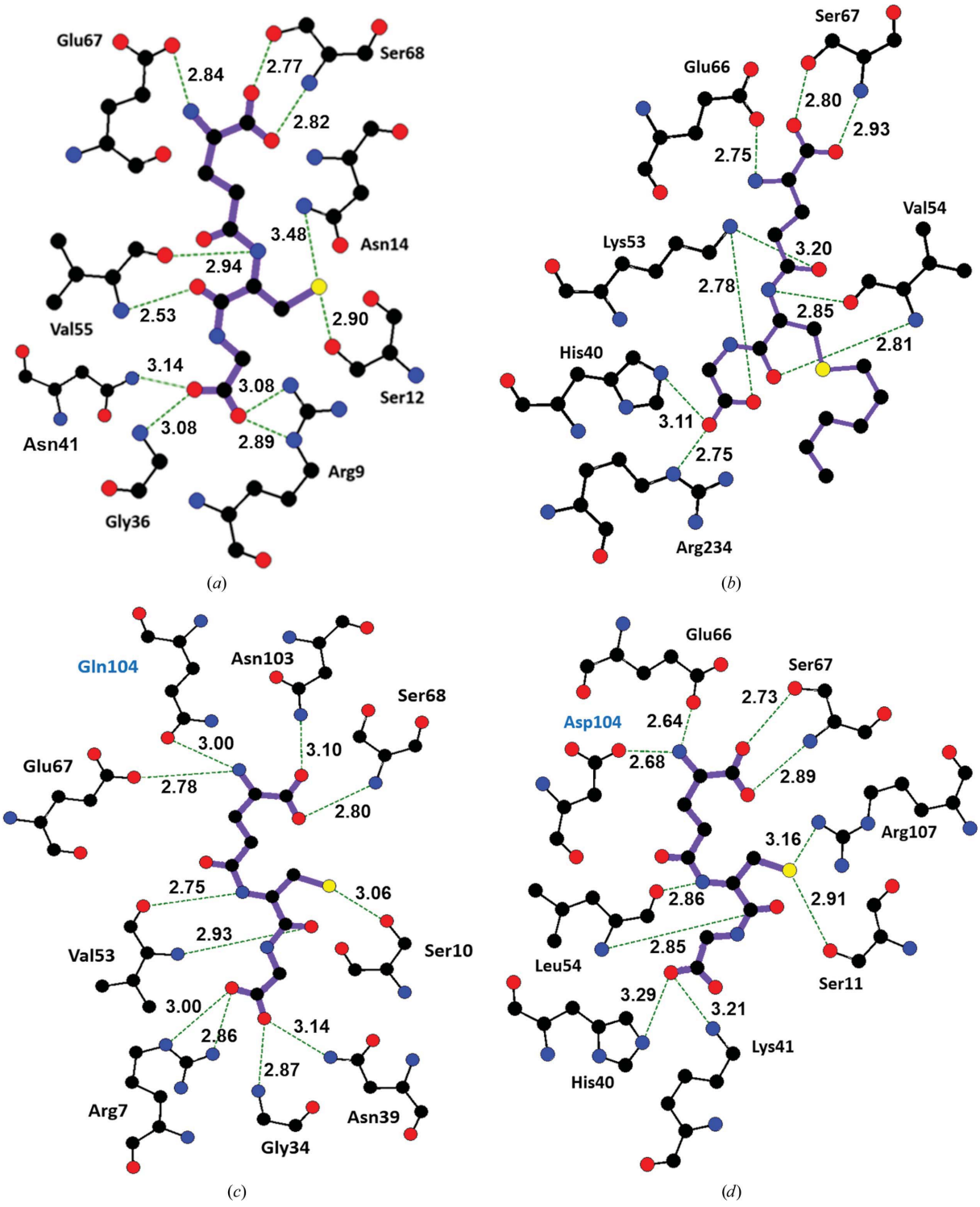


Figure 4
Schematic interactions of GSH with key residues in the G-site of XacGST and other theta-class GSTs. Representations of residues that make hydrogen bonds (green dashed lines with corresponding bond distances) to GSH bound in the G-site are shown. Residues involved in GSH binding are shown as stick models and labeled with residue identities. C, N, O and S atoms are shown as small black, blue, red and yellow solid spheres, respectively. GSH and protein atoms are shown as sticks colored magenta and black, respectively. Residues labeled in blue correspond to residues of the partner monomer in the dimer. (*a*) XacGST with GSH bound (PDB entry 6nv6, chain *B*). (*b*) StGST from *S. typhimurium* with GSH bound (YliJ; PDB entry 4kh7). (*c*, *d*) The theta-class human hGSTT1 mutant W234R (PDB entry 2c3q) and hGSTT2 (PDB entry 4mpg) in complex with *S*-hexylglutathione and GSH, respectively. Theses figures were prepared with *LigPlot* 2.1 (Wallace *et al.*, 1995) and manually modified (amino-acid residue and hydrogen-bond distance labels).

Table 2
PISA analysis of the XacGST model (PDB entry 6nv6, chains *A* and *B*).

	Chain *A*	Chain *B*
Dimeric interface	Leu63, Val64, Leu65, Trp66, Glu67, Asn69, Ala70, Arg73, Tyr74, Ala77, Gln78, Tyr84, Gln86, Pro88, Ala89, Arg91, Ala92, Leu93, Asp95, Arg96, Trp97, Asn99, Trp100, Thr101, Thr102, Ser103, Thr104, Arg142, Ala145	Pro51, Asn52, Leu63, Val64, Leu65, Trp66, Glu67, Asn69, Ala70, Arg73, Tyr74, Ala77, Gln78, Tyr84, Gln86, Pro88, Ala89, Arg91, Ala92, Leu93, Asp95, Arg96, Asp99, Trp100, Thr102, Ser103, Thr104
Inaccessible residues buried in the interior of the protein structure	Ala18, Trp20, Cys21, Pro56, Leu58, Ile71, Val72, Leu75, Ala76, Met98, Thr101, Phe105, Ser136, Leu140, Ala143, Asp162, Ile163, Pro164, Gly166, Ser167, Ile169, Leu185, Trp188, Tyr198, Val202	Arg16, Ala18, Trp20, Cys21, Ala22, Asn50, Thr57, Leu58, Ile71, Val72, Leu75, Ala76, Met98, Phe105, Leu112, Ser136, Leu140, Ala143, Asp162, Ile163, Pro164, Leu165, Gly166, Ser167, Ile169, Ala171, Leu185, Trp188, Leu192, Tyr198, Val202
Salt bridges	Arg73*A*–Asp95*B*, Arg73*A*–Asp99*B*, Aps95*A*–Arg73*B*, Asp99*A*–Arg73*B*	Asp95*B*–Arg73*A*, Asp99*B*–Arg73*A*, Arg73*B*–Asp95*A*, Arg73*B*–Asp99*A*
Hydrogen bonds	Val64*A*–Arg96*B*, Glu67*A*–Ser103*B*, Ala70*A*–Asp99*B*, Arg73*A*–Asp95*B*, Arg73*A*–Asp99*B*, Tyr84*A*–Tyr84*B*, Arg91*A*–Tyr84*B*, Asp95*A*–Arg73*B*, Arg96*A*–Val64*B*, Asp99*A*–Arg73*B*, Trp100*A*–Asn52*B*, Ser103*A*–Glu67*B*, Arg142*A*–Asn52*B*	Arg96*B*–Val64*A*, Ser103*B*–Glu67*A*, Asp99*B*–Ala70*A*, Asp95*B*–Arg73*A*, Asp99*B*–Arg73*A*, Tyr84*B*–Tyr84*A*, Tyr84*B*–Arg91*A*, Arg73*B*–Asp95*A*, Val64*B*–Arg96*A*, Arg73*B*–Asp99*A*, Asn52*B*–Trp100*A*, Glu67*B*–Ser103*A*, Asn52*B*–Arg142*A*

Table 3
The interactions of GSH with conserved amino-acid residues in the G-site of GSTs.

Bound GSH	XacGST	StGST	hGSTT1	hGSTT2
Amine from glutamyl moiety	Glu67	Glu67	Glu66	Glu66
Carboxylate O atoms from glutamyl moiety	Ser68	Ser68	Ser67	Ser67
S atom from cysteinyl moiety	Ser12, Asn14	Ser10	—	Ser11, Arg107
Amide and carbonyl from cysteinyl moiety	Val55	Val53	—	Leu54
Carboxylate O atoms from glycyl moiety	Arg9, Gly36, Asn41	Arg7, Gly34, Asn39	His40, Arg234	His40, Lys41

anion (GS^-) is the most important initial chemical step. Many *in vitro* and *in silico* studies have demonstrated that binding of GSH to GSTs decreases the pK_a of the –SH group from 9.2 to about 6.2 in the enzyme G-site (Perperopoulou *et al.*, 2018). Lower values of pK_a lead to ionization of the thiol group, promoting the creation of a stronger nucleophile that can readily react at near-neutral pH with the electrophilic substrate bound to the enzyme in the H-site (Perperopoulou *et al.*, 2018).

Hydrogen-bonding interactions in the G-site appear to be conserved among XacGST (Fig. 4*a*), the unclassified StGST (Fig. 4*b*) and the theta-class human hGSTT1 (Fig. 4*c*) and hGSTTT2 (Fig. 4*d*). Two-dimensional analysis of the hydrogen bonds between the bound GSH (Figs. 4*a*, 4*b* and 4*d*) and specific residues revealed similarities in the G-site, as summarized in Table 3. The human hGSTT1 mutant W234R (Fig. 4*c*) was crystallized in complex with *S*-hexylglutathione, instead of GSH, preventing a more detailed comparison of the related residues within the G-site (dashes in Table 3). The carboxylate O atoms of the GSH glycyl moiety make hydrogen bonds to the main-chain amide of the Gly36 residue and the side chains of the Arg9 and Asn41 residues in XacGST (Fig. 4*a*). The carbonyl O atom of the GSH cysteinyl moiety makes a hydrogen bond to the main-chain amide of Val55, preceding the *cis*-Pro56 residue located at the beginning of the third strand and at the base of the GSH-binding site. The *cis*-Pro56 residue in XacGST is conserved among different GSTs classes and is essential for proper folding and maintenance of the G-site conformation (Wilce *et al.*, 1995). The cysteinyl S atom of GSH points downwards into the core of the structure and is at a hydrogen-bonding distance (2.90 Å) from the hydroxyl group of Ser12 in XacGST (Fig. 4*a*), which was previously reported to be an essential residue for lowering the pK_a of the thiol group of GSH (Caccuri *et al.*, 1997). The amide and carboxyl group of the GSH cysteinyl moiety form hydrogen bonds to the backbone carbonyl group and amide carbonyl group of Val55. The carboxylate O atoms of the GSH glutamyl moiety make hydrogen bonds to the main-chain amide N atom and hydroxyl group of the conserved Ser68 (Fig. 4*a*). Glu67 in XacGST forms a hydrogen bond to the GSH glutamyl amine group in the GSH-bound XacGST structure (Fig. 4*a*). The same is observed for the equivalent residue Glu66 in hGSTT1 and hGSTT2 as well as for Glu67 in the bacterial StGST structure (PDB entry 4kh7). Hydrophobic contacts in the G-site with GSH bound are limited to Asn14, Gly35, Val54 and Pro56.

In summary, we have demonstrated that XacGST has *in vitro* dehalogenase activity to conjugate GSH with bromoacetate. Structural and sequence analyses suggest that XacGST is a member of the theta-class GSTs.

Acknowledgements

We would like to thank Dr Jesus Aparecido Ferro (FCAV, Jaboticabal-SP, Brazil) and his laboratory staff for providing us with the Xac genomic DNA sample for amplification of the open reading frame Xac3819 (XacGST). Author contributions are as follows. EH conducted all experiments, analyzed the results, performed crystal structure refinement, prepared all figures and tables, and wrote the first complete manuscript draft. SK contributed to the preparation and growth of crystals suitable for X-ray diffraction data collection. EH and LF

conceived the idea for the project, performed the experimental design, analyzed the data and wrote the paper. All authors discussed the data and read the manuscript.

Funding information

This work was partially supported by a grant from the National Institute of Food and Agriculture (NIFA) (grant CA-R-BCH-5051-H to LF) and the National Institutes of Health (grant R01GM108893 to LF).

References

Ahmad, S., Thulasingam, M., Palombo, I., Daley, D. O., Johnson, K. A., Morgenstern, R., Haeggström, J. Z. & Rinaldo-Matthis, A. (2015). *Biochim. Biophys. Acta*, **1854**, 1365–1371.

Allocati, N., Casalone, E., Masulli, M., Ceccarelli, I., Carletti, E., Parker, M. W. & Di Ilio, C. (1999). *FEBS Lett.* **445**, 347–350.

Allocati, N., Federici, L., Masulli, M. & Di Ilio, C. (2009). *FEBS J.* **276**, 58–75.

Allocati, N., Federici, L., Masulli, M. & Di Ilio, C. (2012). *Biochimie*, **94**, 588–596.

Armstrong, R. N. (1991). *Chem. Res. Toxicol.* **4**, 131–140.

Atkinson, H. J. & Babbitt, P. C. (2009). *Biochemistry*, **48**, 11108–11116.

Bartels, F., Backhaus, S., Moore, E. R. B., Timmis, K. N. & Hofer, B. (1999). *Microbiology*, **145**, 2821–2834.

Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. W. (2011). *Acta Cryst.* D**67**, 271–281.

Bradford, M. M. & Williams, W. L. (1976). *Fed. Proc.* **35**, 274.

Caccuri, A. M., Antonini, G., Nicotra, M., Battistoni, A., Lo Bello, M., Board, P. G., Parker, M. W. & Ricci, G. (1997). *J. Biol. Chem.* **272**, 29681–29686.

Chrysostomou, C., Quandt, E. M., Marshall, N. M., Stone, E. & Georgiou, G. (2015). *ACS Chem. Biol.* **10**, 875–882.

Daubaras, D. L., Hershberger, C. D., Kitano, K. & Chakrabarty, A. M. (1995). *Appl. Environ. Microbiol.* **61**, 1279–1289.

DeLano, W. L. (2002). *PyMOL*. https://pymol.org.

Desai, K. K. & Miller, B. G. (2010). *Proc. Natl Acad. Sci. USA*, **107**, 17968–17973.

Dirr, H., Reinemer, P. & Huber, R. (1994). *Eur. J. Biochem.* **220**, 645–661.

Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.

Evans, P. (2006). *Acta Cryst.* D**62**, 72–82.

Fuenmayor, S. L., Wild, M., Boyes, A. L. & Williams, P. A. (1998). *J. Bacteriol.* **180**, 2522–2530.

Gullner, G., Komives, T., Kiraly, L. & Schroder, P. (2018). *Front. Plant Sci.* **9**, 1836.

Habig, W. H., Pabst, M. J. & Jakoby, W. B. (1974). *J. Biol. Chem.* **249**, 7130–7139.

Hayes, J. D., Flanagan, J. U. & Jowsey, I. R. (2005). *Annu. Rev. Pharmacol. Toxicol.* **45**, 51–88.

Hayes, J. D. & Pulford, D. J. (1995). *Crit. Rev. Biochem. Mol. Biol.* **30**, 445–520.

Hilario, E., Li, Y., Niks, D. & Fan, L. (2012). *Acta Cryst.* D**68**, 846–853.

Hilario, E., Martin, F. J. M., Bertolini, M. C. & Fan, L. (2011). *J. Mol. Biol.* **408**, 74–86.

Hofer, B., Backhaus, S. & Timmis, K. N. (1994). *Gene*, **144**, 9–16.

Hylckama Vlieg, J. E. T. van, Leemhuis, H., Spelberg, J. H. L. & Janssen, D. B. (2000). *J. Bacteriol.* **182**, 1956–1963.

Jones, D. T., Taylor, W. R. & Thornton, J. M. (1992). *Comput. Appl. Biosci.* **8**, 275–282.

Ketterer, B. & Meyer, D. J. (1989). *Mutat. Res.* **214**, 33–40.

Krissinel, E. & Henrick, K. (2007). *J. Mol. Biol.* **372**, 774–797.

Kumar, S., Stecher, G., Li, M., Knyaz, C. & Tamura, K. (2018). *Mol. Biol. Evol.* **35**, 1547–1549.

Laemmli, U. K. (1970). *Nature*, **227**, 680–685.

Laia, M. L. (2007). Thesis. Universidade Estadual Paulista (UNESP), Brazil.

Leyns, F., De Cleene, M., Swings, J. G. & De Ley, J. (1984). *Bot. Rev.* **50**, 308–356.

Liebschner, D., Afonine, P. V., Baker, M. L., Bunkóczi, G., Chen, V. B., Croll, T. I., Hintze, B., Hung, L.-W., Jain, S., McCoy, A. J., Moriarty, N. W., Oeffner, R. D., Poon, B. K., Prisant, M. G., Read, R. J., Richardson, J. S., Richardson, D. C., Sammito, M. D., Sobolev, O. V., Stockwell, D. H., Terwilliger, T. C., Urzhumtsev, A. G., Videau, L. L., Williams, C. J. & Adams, P. D. (2019). *Acta Cryst.* D**75**, 861–877.

Lu, S. C. (2013). *Biochim. Biophys. Acta*, **1830**, 3143–3153.

Mannervik, B., Board, P. G., Hayes, J. D., Listowsky, I. & Pearson, W. R. (2005). *Method Enzymol.* **401**, 1–8.

Mannervik, B., Danielson, U. H. & Ketterer, B. (1988). *Crit. Rev. Biochem.* **23**, 283–337.

Mansfield, J., Genin, S., Magori, S., Citovsky, V., Sriariyanum, M., Ronald, P., Dow, M., Verdier, V., Beer, S. V., Machado, M. A., Toth, I., Salmond, G. & Foster, G. D. (2012). *Mol. Plant Pathol.* **13**, 614–629.

Marsh, M., Shoemark, D. K., Jacob, A., Robinson, C., Cahill, B., Zhou, N. Y., Williams, P. A. & Hadfield, A. T. (2008). *J. Mol. Biol.* **384**, 165–177.

McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.

Moore, J. M. (2017). Thesis. Youngstown State University, Ohio, USA.

Oakley, A. (2011). *Drug Metab. Rev.* **43**, 138–151.

Pavlidi, N., Vontas, J. & Van Leeuwen, T. (2018). *Curr. Opin. Insect Sci.* **27**, 97–102.

Perperopoulou, F., Pouliou, F. & Labrou, N. E. (2018). *Crit. Rev. Biotechnol.* **38**, 511–528.

Robert, X. & Gouet, P. (2014). *Nucleic Acids Res.* **42**, W320–W324.

Sambrook, J. & Russell, D. W. (2001). *Molecular Cloning: A Laboratory Manual.* New York: Cold Spring Harbor Laboratory Press.

Sheehan, D., Meade, G., Foley, V. M. & Dowd, C. A. (2001). *Biochem. J.* **360**, 1–16.

Silva, A. C. R. da, Ferro, J. A., Reinach, F. C., Farah, C. S., Furlan, L. R., Quaggio, R. B., Monteiro-Vitorello, C. B., Sluys, M. A. V., Almeida, N. F., Alves, L. M. C., do Amaral, A. M., Bertolini, M. C., Camargo, L. E. A., Camarotte, G., Cannavan, F., Cardozo, J., Chambergo, F., Ciapina, L. P., Cicarelli, R. M. B., Coutinho, L. L., Cursino-Santos, J. R., El-Dorry, H., Faria, J. B., Ferreira, A. J. S., Ferreira, R. C. C., Ferro, M. I. T., Formighieri, E. F., Franco, M. C., Greggio, C. C., Gruber, A., Katsuyama, A. M., Kishi, L. T., Leite, R. P., Lemos, E. G. M., Lemos, M. V. F., Locali, E. C., Machado, M. A., Madeira, A. M. B. N., Martinez-Rossi, N. M., Martins, E. C., Meidanis, J., Menck, C. F. M., Miyaki, C. Y., Moon, D. H., Moreira, L. M., Novo, M. T. M., Okura, V. K., Oliveira, M. C., Oliveira, V. R., Pereira, H. A., Rossi, A., Sena, J. A. D., Silva, C., de Souza, R. F., Spinola, L. A. F., Takita, M. A., Tamura, R. E., Teixeira, E. C., Tezza, R. I. D., Trindade dos Santos, M., Truffi, D., Tsai, S. M., White, F. F., Setubal, J. C. & Kitajima, J. P. (2002). *Nature*, **417**, 459–463.

Skopelitou, K., Dhavala, P., Papageorgiou, A. C. & Labrou, N. E. (2012). *PLoS One*, **7**, e34263.

Skopelitou, K., Muleta, A. W., Pavli, O., Skaracis, G. N., Flemetakis, E., Papageorgiou, A. C. & Labrou, N. E. (2012). *Funct. Integr. Genomics*, **12**, 157–172.

Stein, N. (2008). *J. Appl. Cryst.* **41**, 641–643.

Strange, R. N. & Scott, P. R. (2005). *Annu. Rev. Phytopathol.* **43**, 83–116.

Swings, J., Van Den Mooter, M., Vauterin, L., Hoste, B., Gillis, M., Mew, T. W. & Kersters, K. (1990). *Int. J. Syst. Bacteriol.* **40**, 309–311.

Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). *Nucleic Acids Res.* **22**, 4673–4680.
Vuilleumier, S. (1997). *J. Bacteriol.* **179**, 1431–1441.
Vuilleumier, S. & Pagni, M. (2002). *Appl. Microbiol. Biotechnol.* **58**, 138–146.
Wallace, A. C., Laskowski, R. A. & Thornton, J. M. (1995). *Protein Eng. Des. Sel.* **8**, 127–134.
Wilce, M. C. J., Board, P. G., Feil, S. C. & Parker, M. W. (1995). *EMBO J.* **14**, 2133–2143.
Wilce, M. C. J. & Parker, M. W. (1994). *Biochim. Biophys. Acta*, **1205**, 1–18.
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. 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). *Acta Cryst.* D**67**, 235–242.