Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

X-ray structure of glutathione *S*-transferase from *Schistosoma japonicum* in a new crystal form reveals flexibility of the substrate-binding site

Arne Christian Rufer,[a,b]* Lars Thiebach,[b] Kristin Baer,[b] Helmut W. Klein[b] and Michael Hennig[a]

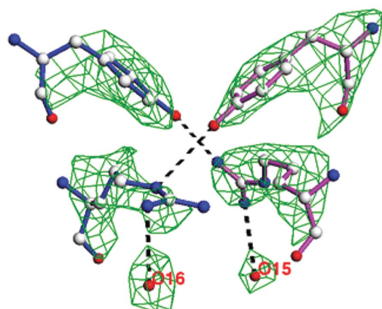
[a]F. Hoffmann–La Roche AG, Pharma Research Discovery Chemistry, 4070 Basel, Switzerland, and [b]University of Cologne, Institute for Biochemistry, 50674 Cologne, Germany

Correspondence e-mail: arne.rufer@roche.com

Received 22 November 2004
Accepted 11 February 2005
Online 24 February 2005

PDB Reference: glutathione *S*-transferase, 1y6e, r1y6esf.

The crystal structure of the 26 kDa glutathione *S*-transferase from *Schistosoma japonicum* (*Sj*GST) was determined at 3 Å resolution in the new space group $P2_12_12_1$. The structure of orthorhombic *Sj*GST reveals unique features of the ligand-binding site and dimer interface when compared with previously reported structures. *Sj*GST is recognized as the major detoxification enzyme of *S. japonicum*, a pathogenic helminth causing schistosomiasis. As resistance against the established inhibitor of *Sj*GST, praziquantel, has been reported these results might prove to be valuable for the development of novel drugs.



© 2005 International Union of Crystallography
All rights reserved

1. Introduction

A fusion protein of the constitutively dimeric 26 kDa glutathione *S*-transferase from *Schistosoma japonicum* (*Sj*GST) with the full-length intracellular domain of the human insulin receptor facilitates investigation of receptor tyrosine kinase activation mediated by dimerization (Baer *et al.*, 2001). As the structures of both the core kinase domain of the human insulin receptor (IRK; Hubbard *et al.*, 1994) and *Sj*GST (McTigue *et al.*, 1995) are known, we attempted to crystallize the corresponding fusion protein, *Sj*GST-IRK, by means of carrier-protein-driven crystallization (Carter *et al.*, 1994; Lim *et al.*, 1994; Zhan *et al.*, 2001; Smyth *et al.*, 2003). The construct *Sj*GST-IRK was shown to be stable during purification and kinase activity was demonstrated in both auto- and substrate-phosphorylation assays (Baer, personal communication).

Crystallization experiments with the fusion protein yielded crystals that were shown to consist of *Sj*GST only. The fusion protein is obviously cleaved under the crystallization conditions and the released *Sj*GST crystallizes in a new crystal form ($P2_12_12_1$), showing distinctively novel features of the dimer interface and ligand-binding site. Coordinates and structure factors have been deposited in the PDB (PDB code 1y6e).

2. Experimental methods

Cloning of the construct GST-IRK will be described elsewhere (Baer *et al.*, manuscript in preparation). A PCR product comprising the coding region of IRK (residues Val966–Lys1271, C969S, Y972F) was cloned into pAc-G2T, resulting in a fusion protein containing a thrombin-cleavage site as a linker between GST and IRK. Expression in Sf9 cells and purification was performed according to Baer *et al.* (2001), with the exception that the crude lysate was incubated for 30 min at 277 K after adding Triton X-100 in order to solubilize *Sj*GST-IRK and Tris-(2-carboxyethyl)-phosphine hydrochloride (TCEP) was used as a reducing agent during purification and subsequent crystallization. 10^9 cells from a 1 l fermentation typically yielded 10 mg of GST-IRK, which was found to be electrophoretically pure (Fig. 1).

Prior to crystallization, GST-IRK was concentrated to 10 mg ml^{-1} in 20 m*M* Tris–HCl pH 7.5, 150 m*M* NaCl, 2 m*M* TCEP in a 30 kDa concentrator (Amicon, Millipore). Crystallization trials were set up directly after purification at 294 K with the modified microbatch method (D'Arcy *et al.*, 2003) and crystals appeared within 7 d with 0.1 *M* Bis-Tris pH 5.5, 25%(*w*/*v*) PEG 3350 (Index Screen condition No. 42; Hampton Research) and 0.1 *M* sodium acetate pH 5.5,

Table 1
Data-collection and refinement statistics.

Values in parentheses are for the highest resolution shell.

Beamline	X06SA, SLS
Wavelength (Å)	0.90006
Resolution (Å)	30.0–3.0 (3.14–3.0)
Space group	$P2_12_12_1$
Unit-cell parameters (Å)	a = 161.22, b = 50.75, c = 57.50
V_M† (Å^3 Da^{-1})	2.2
Solvent content (%)	44.8
Total reflections	114066
Unique reflections	9910
Average redundancy	11.5 (8.6)
$I/\sigma(I)$	8.1 (5.5)
Completeness (%)	99.6 (99.4)
Wilson B (Å^2)	43.2
R_{merge}‡	12.3 (43.7)
$\langle B \rangle$ (Å^2)	36.7
$\sigma(B)$ (Å^2)	14.8
R_{cryst} (%)	21.1
R_{free} (%)	27.9
R.m.s.d. bond lengths (Å)	0.008
R.m.s.d. bond angles (°)	1.44
Ramachandran plot (%)	
Most preferred regions	85.9
Allowed regions	12.8
Generously allowed regions	1.3
Disallowed regions	0.0

† Matthews (1968). ‡ $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I(hkl)$.

36%(w/v) PEG MME 5000 (Stura Footprint Screen 2 condition No. 20; Molecular Dimensions Ltd) as precipitant solutions. Crystals grew to a maximum diameter of 15 µm within 7 d and were flash-frozen in liquid nitrogen directly from the screening plates.

Data were collected at the Swiss Light Source beamline X06SA at 100 K using a 165 mm MAR CCD detector with a 260 mm crystal-to-detector distance and an oscillation range of 1°. 180 frames were collected and processed with *DENZO* and *SCALEPACK* (Otwinowski, 1993). Despite their small size, the crystals diffracted to 3.0 Å resolution and analysis of the systematic abscences revealed that they belong to space group $P2_12_12_1$. A noncrystallographic twofold axis generates one GST dimer per asymmetric unit. The crystallographic parameters are summarized in Table 1.

A molecular-replacement solution was found with the program *MOLREP* (Vagin & Teplyakov, 1997; Collaborative Computational Project, Number 4, 1994) using the modified PDB entry 1dug (Ware *et al.*, 1999), *i.e.* an unligated dimer of GST comprising residues 1–217. Following rigid-body refinement in *REFMAC* (Murshudov *et al.*, 1997; Collaborative Computational Project, Number 4, 1994), the structure of orthorhombic GST was refined by iterative conjugate-gradient minimization with subsequent grouped B-factor refinement in *CNX* (Brünger, 1992) and manual model building with *MOLOC* (Gerber, 1992; refinement statistics in Table 1). The two carboxy-terminal amino acids Pro217 and Lys218 have been omitted from refinement because no clear electron density was visible for these residues. Based on the same argument, six residues (*i.e.* SDLVPR) of the linker peptide encoded by pAc-G2T that were still attached at the carboxy-terminus of the *Sj*GST moiety according to mass spectrometry and amino-terminal sequencing (data not shown) were also left out from refinement. Finally, 16 water molecules were added manually based on electron density and hydrogen bonding.

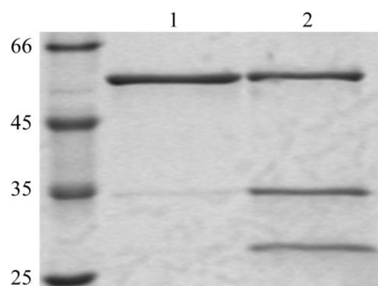


Figure 1
SDS–PAGE. Lane 1 shows the pooled elution fractions of *Sj*GST-IRK from a gel-filtration column. In lane 2 the same preparation is depicted after storage on ice for 10 d. The molecular weights (kDa) of the marker proteins are indicated.

Table 2
26 kDa *Sj*GST (EC 2.5.1.18) structures in PDB.

		Unit-cell parameters				
PDB entry	Space group	a (Å)	b (Å)	c (Å)	Apo/ligand	Fusion protein
1dug	$P4_12_12$	105.78	105.78	137.23	l	Yes
1b8x	$P4_32_12$	93.40	93.40	57.60	a	Yes
1bg5	$P4_32_12$	92.17	92.17	57.57	a	Yes
1gne	$P4_32_12$	94.74	94.74	58.13	l	Yes
1ua5	$P4_32_12$	92.53	92.53	57.66	l	No
1gta	$P6_322$	125.20	125.20	70.20	a	No
1gtb	$P6_322$	123.80	123.80	70.20	l	No
1m99	$P6_322$	115.07	115.07	78.28	l	No
1m9a	$P6_322$	114.99	114.99	78.35	l	No
1m9b	$P6_322$	116.57	116.57	78.75	l	No

3. Results and discussion

Analysis of the X-ray data resulted in determination of the space group as orthorhombic $P2_12_12_1$, with unit-cell parameters a = 161.22, b = 50.75, c = 57.50 Å. The calculated volume of the asymmetric unit, 117 622.3 Å^3, was obviously incompatible with the presence of the intact fusion protein with a molecular weight of 61 kDa. Consequently, we concluded that only a fragment of *Sj*GST-IRK had crystallized. Although the fusion protein had been purified to homogeneity, it was realised that the protein spontaneously degraded to *Sj*GST and IRK during prolonged storage on ice (*i.e.* within 10 d), independent of the purification protocol employed (Fig. 1). Addition of protease inhibitors (RoComplete, Roche; diisopropyl fluorophosphate, Fluka) to the storage buffer did not prevent degradation and no proteases could be detected using a sensitive spectroscopic assay with resorufin-labelled casein as substrate (Universal Protease Substrate assay, Roche). We assume that spontaneous autocatalytic cleavage as described for nucleolin (Chen *et al.*, 1991), which also undergoes degradation independent of both exogenous and endogenous protease activity, could be the reason for the degradation. Replacement of the pAc-G2T linker with a thrombin-cleavage site by a more rigid connection according to Smyth *et al.* (2003) was not pursued as the introduction of a (Gly-Ala)$_5$ linker resulted in significant loss of activity in autophosphorylation assays indicating suboptimal alignment of the kinase domains.

Molecular-replacement trials with a high-resolution structure of *Sj*GST (PDB code 1dug) were performed which yielded marked peaks for both the rotation and translation functions. Given the crystal packing and electron density, the crystals were determined to consist of *Sj*GST crystallized in space group $P2_12_12_1$ (Table 1). To our knowledge these are the first orthorhombic crystals of *Sj*GST, as exclusively tetragonal and hexagonal crystal forms have been reported so far (Table 2; for a review, see Zhan *et al.*, 2001). For the deposited *Sj*GST structures no correlation exists between space group and crystallization conditions in terms of protein fusion or presence of ligand (Table 2). Inspection of the crystal packing in orthorhombic *Sj*GST leads to the conclusion that the solvent-channel cavity around the carboxy-termini could potentially accommodate covalently linked peptides but no larger protein domain for carrier-

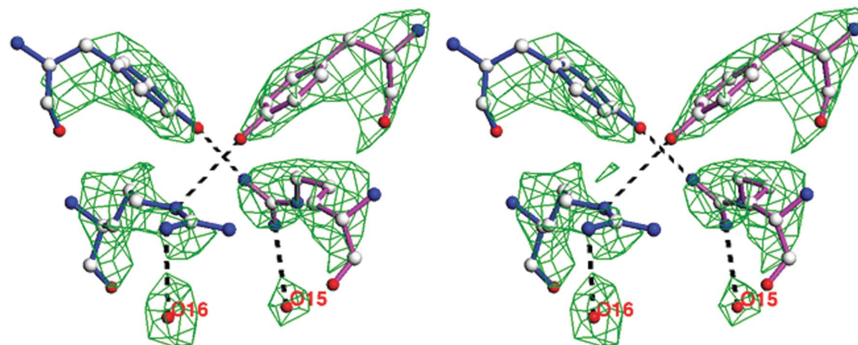


Figure 2
Stereoview of the reciprocal dimer contact between Tyr103 and Arg107 of chains *A* (blue) and *B* (magenta). The distance of the hydroxy O atom of Tyr103*A* to the terminal amino N atom of Arg107*B* is 3.3 Å and N^{ε} of Arg107*A* is 3.5 Å away from the hydroxy O atom of Tyr103*B*. Residues Arg107 of chains *A* and *B* each hydrogen bond (3.3 and 3.1 Å distance to W16 and W15, respectively) to water molecules.

protein-driven crystallization. Although the overall structure of orthorhombic *Sj*GST is similar to previously reported *Sj*GST structures solved in different space groups (r.m.s.d.s for main- and side-chain atoms are 0.51 and 1.34 Å for chain *A* and 0.86 and 1.47 Å for chain *B* using residues 1–217 of 1dug as reference), a unique contact in the dimer interface was identified in the present structure.

The $F_o - F_c$ simulated-annealing omit electron-density map contoured at 4σ clearly indicates that the side chains of Tyr103 are rotated by approximately 100° around the $C^{\alpha}-C^{\beta}$ bond and the positions of the side chains of Arg107 are considerably different compared with all previously published *Sj*GST structures (Fig. 2). The phenyl hydroxy group of Tyr103 in chain *A* is at a hydrogen-bonding distance from the guanidinium group of Arg107 of chain *B* (and *vice versa*), thereby establishing a reciprocal dimer contact. Interestingly, residue Tyr103 of *Sj*GST is a critical part of the hydrophobic binding site (H-site) for endogenous substrates, xenobiotics and the only effective anti-schistosomial drug praziquantel, which additionally contacts Arg107 (McTigue *et al.*, 1995; Cardoso *et al.*, 2003; Hu *et al.*, 2004). In the structure presented here, Tyr103 partially occupies the praziquantel-binding site. Although the structure of *Sj*GST with bound praziquantel and its corresponding apo-structure (PDB codes 1gtb and 1gta, respectively; McTigue *et al.*, 1995) imply a preformed binding site for the drug, our data reveal a conformational flexibility for the critical residues Tyr103 and Arg107. This, together with the absence of a homologous tyrosine residue in mammalian GSTs (McTigue *et al.*, 1995), might support the design of specific drugs targeting the 26 kDa GST of *Schistosoma* spp., which have already been shown to develop resistance against praziquantel (Hu *et al.*, 2004, and references therein). The need for novel highly potent drugs is also emphasized by the finding that an estimated 250 million people are infected with *Schistosoma* spp.

The authors would like to thank the staff of X06SA at SLS, Villigen, Switzerland for support during data collection, A. Friedlein (F. Hoffmann–La Roche Ltd, Basel) for MS analysis, M. Gompert (University of Cologne) for maintaining Sf9 cell cultures and Anke Klose (University of Cologne) for constructing the linker mutant.

References

Baer, K., Al-Hasani, H., Parvaresch, S., Corona, T., Rufer, A., Nolle, V., Bergschneider, E. & Klein, H. W. (2001). *Biochemistry*, **40**, 14268–14278.
Brünger, A. T. (1992). *X-PLOR Manual Version 3.1*. New Haven, CT, USA: Yale University Press.
Cardoso, R. M., Daniels, D. S., Bruns, C. M. & Tainer, J. A. (2003). *Proteins*, **51**, 137–146.
Carter, D. C., Rüker, F., Ho, J. X., Lim, K., Keeling, K., Gilliland, G. & Ji, X. (1994). *Protein Pept. Lett.* **1**, 175–178.
Chen, C. M., Chiang, S. Y. & Yeh, N. H. (1991). *J. Biol. Chem.* **266**, 7754–7758.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
D'Arcy, A., Mac Sweeney, A., Stihle, M. & Haber, A. (2003). *Acta Cryst.* D**59**, 396–399.
Gerber, P. R. (1992). *Biopolymers*, **32**, 1003–1017.
Lim, K., Ho, J. X., Keeling, K., Gilliland, G. L., Ji, X., Rüker, F. & Carter, D. C. (1994). *Protein Sci.* **3**, 2233–2244.
Hu, W., Brindley, P. J., McManus, D. P., Feng, Z. & Han, Z. G. (2004). *Trends Mol. Med.* **10**, 217–225.
Hubbard, S. R., Wei, L., Ellis, L. & Hendrickson, W. (1994). *Nature (London)*, **372**, 746–753.
McTigue, M. A., Williams, D. R. & Tainer, J. A. (1995). *J. Mol. Biol.* **246**, 21–27.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.
Otwinowski, Z. (1993). *Proceedings of the CCP4 Study Weekend. Data Collection and Processing*, edited by L. Sawyer, N. Isaacs & S. Bailey, pp. 56–62. Warrington: Daresbury Laboratory.
Smyth, D. R., Mrozkiewicz, M. K., McGrath, W. J., Listwan, P. & Kobe, B. (2003). *Protein Sci.* **12**, 1313–1322.
Vagin, A. & Teplyakov, A. (1997). *J. Appl. Cryst.* **30**, 1022–1025.
Ware, S., Donahue, J. P., Hawiger, J. & Anderson, W. F. (1999). *Protein Sci.* **12**, 2663–2671.
Zhan, Y., Song, X. & Zhou, G. W. (2001). *Gene*, **281**, 1–9.