

Structure of an insect δ -class glutathione *S*-transferase from a DDT-resistant strain of the malaria vector *Anopheles gambiae*

Liqing Chen,^{a*} Pamela R. Hall,^{a†}
 Xiaoyin E. Zhou,^a Hilary
 Ranson,^b Janet Hemingway^b and
 Edward J. Meehan^a

^aLaboratory for Structural Biology, Department of Chemistry, Graduate Programs of Biotechnology, Chemistry and Materials Science, University of Alabama in Huntsville, Huntsville, AL 35899, USA, and ^bThe Liverpool School of Tropical Medicine, Pembroke Place, Liverpool L3 5QA, England

† Present address: Department of Molecular Cardiology and Center for Structural Biology, Lerner Research Institute, Cleveland Clinic Foundation, OH 44195, USA.

Correspondence e-mail: chenlq@email.uah.edu

Glutathione *S*-transferases (GSTs) are a major family of detoxification enzymes which possess a wide range of substrate specificities. Most organisms possess many GSTs belonging to multiple classes. Interest in GSTs in insects is focused on their role in insecticide resistance; many resistant insects have elevated levels of GST activity. In the malaria vector *Anopheles gambiae*, elevated GST levels are associated with resistance to the organochlorine insecticide DDT [1,1,1-trichloro-2,2-bis-(*p*-chlorophenyl)ethane]. This mosquito is the source of an insect GST, agGSTd1-6, which metabolizes DDT and is inhibited by a number of pyrethroid insecticides. The crystal structure of agGSTd1-6 in complex with its inhibitor *S*-hexyl glutathione has been determined and refined at 2.0 Å resolution. The structure adopts a classical GST fold and is similar to those of other insect δ -class GSTs, implying a common conjugation mechanism. A structure-based model for the binding of DDT to agGSTd1-6 reveals two subpockets in the hydrophobic binding site (H-site), each accommodating one planar *p*-chlorophenyl ring.

Received 27 June 2003
 Accepted 15 September 2003

PDB Reference: agGSTd1-6,
 1pn9, r1pn9sf.

1. Introduction

Glutathione *S*-transferases (GSTs; EC 2.5.1.18) are a major family of detoxification enzymes that among other reactions conjugate glutathione (GSH; γ -glutamyl-cysteinyl-glycine) to xenobiotic compounds (*e.g.* drugs, herbicides, insecticides) with electrophilic centers, converting them from reactive lipophilic molecules into water-soluble non-reactive conjugates that may easily be excreted (Hayes & Pulford, 1995). Most organisms possess multiple GSTs belonging to two or more classes with differing catalytic activities to accommodate the wide range of substrate specificities. Mammalian GSTs have been classified into eight soluble classes (α , μ , π , θ , σ , ζ , κ and ω) and a microsomal class (Mannervik, 1985; DeJong *et al.*, 1988; Meyer *et al.*, 1991; Pemble *et al.*, 1996; Board *et al.*, 1997, 2000). In plants, GSTs are grouped into five classes (θ , λ , ζ , ϕ and τ ; Edwards *et al.*, 2000; Dixon *et al.*, 2002). Insect GSTs were recently classified into six classes (δ , ε , σ , θ , ω and ζ) by comparative analysis of the *Drosophila melanogaster* and *Anopheles gambiae* genomes from the three formerly recognized (Fournier *et al.*, 1992; Syvanen *et al.*, 1994; Ranson *et al.*, 2002; Ding *et al.*, 2003). The δ - and ε -class GSTs have been implicated in detoxification, particularly in conferring resistance towards various insecticides (Hemingway, 2000; Prapanthadara *et al.*, 2000; Ranson *et al.*, 2001; Ortelli *et al.*, 2003).

Interest in insect GSTs is focused on their role in insecticide resistance. Elevated levels of GST activity have been detected

in strains of insects resistant to organophosphates (Fournier *et al.*, 1992), organochlorines (Grant & Hammock, 1992) and pyrethroid insecticides (Kostaropoulos *et al.*, 2001; Vontas *et al.*, 2001). GSTs from the important malaria vector *A. gambiae* are of particular interest because of their involvement in resistance to the organochlorine insecticide DDT [1,1,1-trichloro-2,2-bis-(*p*-chlorophenyl)ethane]. In *A. gambiae*, an increased rate of DDT dehydrochlorination in the resistant strain is associated with quantitative increases in multiple GST enzymes (Prapanthadara *et al.*, 1993). This mosquito is the source of an insect δ -class GST, agGSTd1-6, the product of one of four alternative transcripts from the agGSTd1 gene, which metabolizes DDT and binds to a number of pyrethroid insecticides (Ranson *et al.*, 1997). The latter characteristic has allowed this recombinant enzyme to be used to monitor pyrethroid concentrations on insecticide-impregnated bednets (Enayati *et al.*, 2001). The potential role of agGSTd1-6 in pesticide detoxification makes it an attractive target for structural analysis. Malaria kills several million people each year (Phillips, 2001). Preventive measures have focused on control of the mosquito vectors using insecticides such as DDT (Trigg & Kondrachine, 1998). However, the advent of DDT-resistant strains of mosquitos has decreased the effectiveness of this control measure. Structural studies of insect GSTs from insecticide-resistant strains will help in the understanding of the mechanisms of resistance to important pesticides and guide the design of novel inhibitors to overcome insecticide resistance. To this end, we have determined the crystal structure of agGSTd1-6 from *A. gambiae* at 2.0 Å resolution.

2. Materials and methods

2.1. Protein preparation, crystallization and data collection

Cloning of agGSTd1-6 from a DDT-resistant *A. gambiae* strain has been described previously (Ranson *et al.*, 1997). Detailed protocols for agGSTd1-6 expression, purification, crystallization and preliminary X-ray analysis have also been reported (Roberts *et al.*, 2001). Our initial screening of crystallization conditions produced five different crystal forms for agGSTd1-6. The crystal form used in the structure determination was one of the two primitive orthorhombic forms. The crystals were grown at room temperature (298 K) by the hanging-drop vapor-diffusion method with 100 mM Tris–HCl buffer pH 7.5, 30% PEG 4000, 10% 2-propanol as the crystallization solution. Crystals formed in space group $P2_12_12_1$, with unit-cell parameters $a = 50.2$, $b = 89.5$, $c = 100.0$ Å, and contain two monomers in the asymmetric unit. Preliminary data collection was carried out with an in-house rotating-anode X-ray source. The programs *DENZO* and *SCALEPACK* were used for data processing and analysis (Otwinowski & Minor, 1996). Final diffraction data to 2.0 Å resolution were collected at the 19ID beamline of the APS at Argonne National Laboratory, USA. Data-collection statistics are summarized in Table 1.

Table 1

Summary of data-collection, phasing and refinement statistics.

Values in parentheses are for the highest resolution shell.

Data collection	
Space group	$P2_12_12_1$
Unit-cell parameters (Å)	$a = 50.2$, $b = 89.5$, $c = 100.0$
Temperature (K)	93
Wavelength (Å)	0.979
Crystal-to-detector distance (mm)	200
Oscillation range per frame (°)	0.5
Exposure time (s)	10
Resolution (Å)	2.0
Total No. of reflections	
Unique	26280
Measured	118522
Completeness (%)	84.3 (87.3)
R_{merge}	0.074 (0.195)
$I/\sigma(I)$ (average)	17.4 (7.4)
Phasing (molecular replacement)	
Search model	dmGST21 dimer
Cross-rotation†	
Peak 1	$\theta_1 = 171.0$, $\theta_2 = 41.4$, $\theta_3 = 42.7^\circ$; Height = 0.100
Peak 2	$\theta_1 = 11.5$, $\theta_2 = 57.1$, $\theta_3 = 251.6^\circ$; Height = 0.039
Translation†	
Peak 1	$\theta_1 = 171.0$, $\theta_2 = 41.4$, $\theta_3 = 42.6^\circ$; $T_x = 8.43$, $T_y = 36.64$, $T_z = 8.16$; Monitor = 0.388, Packing = 0.63
Peak 2	$\theta_1 = 12.2$, $\theta_2 = 57.9$, $\theta_3 = 249.9^\circ$; $T_x = 17.6$, $T_y = 33.8$, $T_z = 42.6$; Monitor = 0.091, Packing = 0.61
Initial R factor	0.462
Refinement	
Resolution range (Å)	20–2.0
No. of reflections	24652
No. of atoms	
Protein	3310
Ligand	52
Water	233
R factors (%)	
R_{work}	20.9 (21.9)
R_{free}	25.4 (27.1)
R.m.s.d.s	
Bond length (Å)	0.006
Bond angle (°)	1.2

† Only the highest two peaks are listed.

2.2. Structure determination and refinement

The structure of agGSTd1-6 was determined by the molecular-replacement method. The search model was a dimer of a δ -class GST from *D. melanogaster*, dmGST21, which has 57% sequence identity with that of agGSTd1-6 (Wang & Rose, 2000). Molecular-replacement calculations and structure refinement were carried out using the *CNS* program package (Brünger *et al.*, 1998). The graphics program *O* was used in model building (Jones *et al.*, 1991). Both the cross-rotation function and subsequent translation-function searches yielded one solution that was much higher than the next highest peak (Table 1). This solution was used as the initial model for agGSTd1-6 and gave an R factor of 0.46 ($R_{\text{free}} = 0.46$) using 20–2.0 Å data. Subsequent rigid-body, energy minimization with simulated annealing and restrained individual B -factor refinements lowered the R factor to 0.346 ($R_{\text{free}} = 0.393$). The composite omit map was calculated in order to reduce the model bias. This omit map was used to guide the

model building of the agGSTd1-6 structure, including fitting of the inhibitor *S*-hexyl glutathione (GTX). Refinement proceeded through several cycles in combination with manual rebuilding. After adding solvent molecules, the refinement converged at an *R* factor of 20.9% ($R_{\text{free}} = 25.4\%$). The final model contains all residues of monomers *A* and *B*, one *S*-hexyl glutathione per monomer and 233 water molecules. The refinement statistics are summarized in Table 1.

2.3. Protein-fold analysis

Secondary-structure elements were defined by their hydrogen-bonding patterns. The *DALI* algorithm was used to search for structural homologues of agGSTd1-6 and was also used for structure-based sequence alignment of agGSTd1-6 with other GSTs (Holm & Sander, 1993, 1998). Ribbon and stereo diagrams were prepared using the program *MOLSCRIPT* (Kraulis, 1991).

3. Results and discussion

3.1. Overall structure of agGSTd1-6

The refined model of agGSTd1-6 in complex with its inhibitor *S*-hexylglutathione (GTX) had good overall geometry,

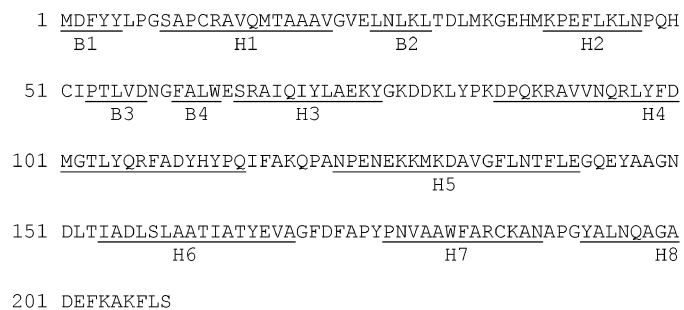


Figure 1

Sequence of agGSTd1-6 and assignment of its secondary-structure elements. The secondary-structure elements are underlined and labeled (α -helices starting with H and β -strands with B).

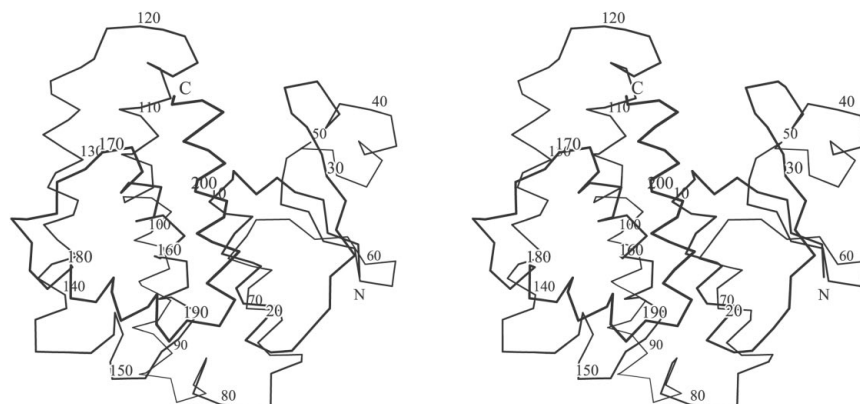


Figure 2

Stereoview of the C^α trace of agGSTd1-6. Every tenth residue is labelled. The N- and C-termini are also labeled.

with r.m.s deviations for bond lengths and angles of 0.006 Å and 1.2°, respectively (Table 1). The Ramachandran plot statistics showed that 90.4% of the dihedral angles were found to be in the most favored regions, 8.0% in the additional allowed regions and only 1.6% (six residues) in the generously allowed regions. None of the non-glycine residues were in disallowed regions. The average *B* values for the protein and GTX atoms were 30.6 and 25.4 Å², respectively.

There are two agGSTd1-6 monomers in the crystal asymmetric unit, with each monomer consisting of 209 residues (Fig. 1). The r.m.s. deviation of the 209 C^α atoms between the two monomers is 0.70 Å. agGSTd1-6 adopts the canonical GST fold (Figs. 2 and 3), containing eight α -helices (H1–H8) and four β -strands (B1–B4), and the structure can be divided into two distinct domains plus a short hinge loop, namely an N-terminal domain (residues 1–78), a linker (residues 79–85) and a C-terminal domain (residues 86–209). The N-terminal domain consists of a central four-stranded mixed β -sheet flanked on one side by helices H1 (residues 9–22) and H3 (residues 64–76) and on the other by helix H2 (residues 40–47). These secondary-structural elements are arranged in a $\beta\alpha\beta\alpha\beta\alpha$ motif in which the β -strand B3 (residues 53–57) is antiparallel with respect to the other three β -strands B1 (residues 1–5), B2 (residues 26–30) and B4 (residues 60–63). The mixed β -sheet adopts a ‘–1 +2 +1’ topology. Pro53 at the start of the β -strand B3 is in the *cis* conformation. Equivalent *cis*-residues have been found in all GST structures so far determined (Armstrong, 1997). This proline residue appears to be critical for the correct formation of the active site. The C-terminal domain has an all- α fold with a bundle of five α -helices [H4 (residues 86–115), H5 (residues 123–142), H6 (residues 154–169), H7 (residues 177–189) and H8 (residues 193–209)]. Helix H4 is slightly bent at position Gly102.

3.2. Active-site structure

One molecule of GTX is bound in the active site of each monomer (Figs. 3 and 4). The active site is located in a deep cleft formed at the interface of the two domains (Fig. 3). The inhibitor molecule GTX sits tightly inside the active-site pocket formed by residues Leu6, Ser9, Ala10, Pro11, Leu33, Met34, His38, His50, Ile52, Glu64, Ser65, Arg66, Tyr105, Phe108, Tyr113, Ile116, Phe117, Phe203 and Phe207 (Fig. 4). The active site can be divided into two subsites, the glutathione (GSH) binding site (G-site) and the hydrophobic binding site (H-site). The G-site is mainly hydrophilic and is polar in nature. The GSH moiety of the inhibitor GTX lies in this site, with its γ -glutamyl region forming hydrogen bonds with the side chain of Glu64, the main-chain amide and the hydroxyl group of Ser65 and the side chain of Arg66. Its cysteinyl moiety forms two hydrogen bonds, one to the main-chain carbonyl of Ile52 and the other to the amide N atom of Ile52. The

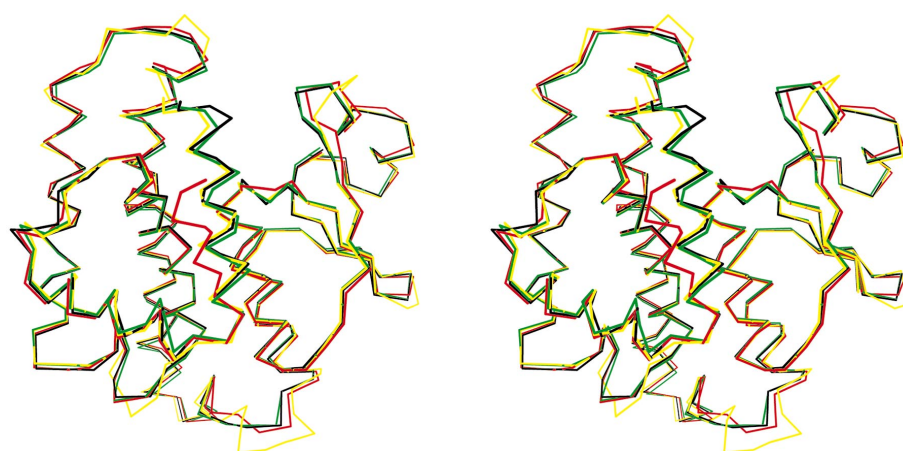
Table 2

Structural alignment statistics of agGSTd1-6 with other GSTs.

Produced using the *DALI* algorithm (Holm & Sander, 1993, 1998). agGSTd1-6, *A. gambiae*; adGSTd1-3 and adGSTd1-4, *A. dirus* species B; dmGST21 and dmGST2, *D. melanogaster*; lcGST, *L. cuprina*; hGSTt2-2, human θ .

Protein	PDB code	Z†	RMSD‡	LALI§	LSEQ2¶	%IDE††	Reference
agGSTd1-6	1pn9	40.8	0.0	209	209	100	This paper
adGSTd1-3	1jlv	36.1	0.6	207	207	82	Oakley <i>et al.</i> (2001)
dmGST21	n/a	34.4	1.0	209	213	57	Wang & Rose (2000)
adGSTd1-4	1jlw	34.2	1.0	208	217	69	Oakley <i>et al.</i> (2001)
lcGST	n/a	31.5	1.0	200	201	68	Wilce <i>et al.</i> (1995)
hGSTt2-2	3ljr	24.6	1.8	206	244	33	Rossjohn <i>et al.</i> (1998)
dmGST2	1m0u	18.4	2.6	187	203	15	Agianian <i>et al.</i> (2003)

† Z score, strength of structural similarity in standard deviations above that expected. ‡ Positional root-mean-square deviation of superimposed C α atoms in Å. § Total number of equivalent residues. ¶ Length of the entire chain of the equivalent structure. †† Percentage of sequence identity over equivalent positions.

**Figure 7**

Stereoview of the superimposed structures of agGSTd1-6 (black), adGSTd1-3 (green), adGSTd1-4 (yellow) and lcGST (red).

This research was supported in part by NSF-EPSCoR, NASA and a generous gift from an anonymous donor to the Laboratory for Structural Biology, University of Alabama in Huntsville. HR is supported by the Royal Society. We would like to thank Dr Randy Alkire of SBC-CAT for hosting our synchrotron data collection at APS beamline 19ID. Use of the Argonne National Laboratory Structural Biology Center beamlines at the Advanced Photon Source was supported by the US Department of Energy, Office of Energy Research under contract No. W-31-109-ENG-38. We are grateful to Drs Bi-Cheng Wang and John Rose for the coordinates of dmGST21, and to Dr Michael Parker for the coordinates of lcGST. We also want to thank Joyce Looger for her technical assistance with computer programs.

References

- Agianian, B., Tucker, P. A., Schouten, A., Leonard, K., Bullard, B. & Gros, P. (2003). *J. Mol. Biol.* **326**, 151–165.
 Armstrong, R. N. (1997). *Chem. Res. Toxicol.* **10**, 2–18.
 Board, P. G., Baker, R. T., Chelvanayagam, G. & Jermin, L. S. (1997). *Biochem. J.* **328**, 929–939.
 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., Chrnyk, B. A., Perregaux, D. E., Gabel, C. A., Geoghegan, K. F. & Pandit, J. (2000). *J. Biol. Chem.* **275**, 24798–24806.
 Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Gross-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst. D* **54**, 905–921.
 Clark, A. G. & Shamaan, N. A. (1984). *Pestic. Biochem. Physiol.* **22**, 249–261.
 DeJong, J. L., Morgenstern, R., Jornvall, H., DePierre, J. W. & Tu, C.-P. D. (1988). *J. Biol. Chem.* **263**, 8430–8436.
 Ding, Y., Ortelli, F., Rossiter, L. C., Hemingway, J. & Ranson, H. (2003). *BMC Genomics*, **4**, 35–50.
 Dixon, D. P., Laphorn, A. & Edwards, K. (2002). *Gene Biol.* **3**, 3004.1–3004.10.
 Edwards, R., Dixon, D. P. & Walbot, V. (2000). *Trends Plant Sci.* **5**, 193–198.
 Enayati, A. A., Vontas, J. G., Small, G. J., McCarroll, L. & Hemingway, J. (2001). *Med. Vet. Entomol.* **15**, 58–63.
 Fournier, D., Bride, J. M., Poire, M., Berge, J. B. & Plapp, F. W. (1992). *J. Biol. Chem.* **267**, 1840–1845.
 Grant, D. F. & Hammock, B. D. (1992). *Mol. Gen. Genet.* **234**, 169–176.
 Hayes, J. D. & Pulford, D. J. (1995). *Crit. Rev. Biochem. Mol. Biol.* **30**, 445–600.
 Hemingway, J. (2000). *Insect Biochem. Mol. Biol.* **30**, 1009–1015.
 Holm, L. & Sander, C. (1993). *J. Mol. Biol.* **233**, 123–138.
 Holm, L. & Sander, C. (1998). *Proteins Struct. Funct. Genet.* **33**, 88–96.
 Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst. A* **47**, 110–119.
 Kostaropoulos, I., Papadopoulos, A. I., Metaxakis, A., Boukouvala, E. & Papadopolou-Mourkidou, E. (2001). *Insect Mol. Biol.* **31**, 313–319.
 Kraulis, P. J. (1991). *J. Appl. Cryst.* **24**, 946–950.
 Mannervik, B. (1985). *Adv. Enzymol.* **57**, 357–417.
 Meyer, D. J., Coles, B., Pemble, S. E., Gilmore, K. S., Fraser, G. M. & Ketterer, B. (1991). *Biochem. J.* **274**, 409–414.
 Oakley, A. J., Harnnoi, T., Udomsinprasert, R., Jirajaroenrat, K., Ketterman, A. J. & Wilce, M. C. (2001). *Protein Sci.* **10**, 2176–2185.
 Ortelli, F., Rossiter, L. C., Vontas, J., Ranson, H. & Hemingway, J. (2003). *Biochem. J.* **373**, 957–963.
 Otwinowski, Z. & Minor, W. (1996). *Methods Enzymol.* **276**, 307–326.
 Pemble, S. E., Wardle, A. F. & Tayler, J. B. (1996). *Biochem. J.* **319**, 749–754.
 Phillips, R. S. (2001). *Clin. Microbiol. Rev.* **14**, 208–226.
 Prapanthadara, L., Hemingway, J. & Ketterman, A. J. (1993). *Pestic. Biochem. Physiol.* **47**, 119–133.
 Prapanthadara, L., Promtet, N., Koottathep, S., Somboon, P. & Ketterman, A. J. (2000). *Insect Biochem. Mol. Biol.* **30**, 395–403.
 Ranson, H., Claudianos, C., Ortelli, F., Abgrall, C., Hemingway, J., Sharakhova, M. V., Unger, M. F., Collins, F. H. & Feyereisen, R. (2002). *Science*, **298**, 179–181.
 Ranson, H., Prapanthadara, L. & Hemingway, J. (1997). *Biochem. J.* **324**, 97–102.
 Ranson, H., Rossiter, L., Ortelli, F., Jensen, B., Wang, X., Roth, C. W., Collins, F. H. & Hemingway, J. (2001). *Biochem. J.* **359**, 295–304.
 Roberts, P. H., Zhou, X., Holmes, A. M., Ranson, H., Small, G.,

- Hemingway, J., Ng, J. D., Chen, L. & Meehan, E. J. (2001). *Acta Cryst.* **D57**, 134–136.
- Rossjohn, J., McKinstry, W. J., Oakley, A. J., Verger, D., Flanagan, J., Chelvanayagam, G., Tan, K. L., Board, P. G. & Parker, M. W. (1998). *Structure*, **6**, 309–322.
- Sheehan, D., Meade, G., Foley, V. M. & Dowd, C. A. (2001). *Biochem. J.* **360**, 1–16.
- Syvanen, M., Zhou, Z. H. & Wang, J. Y. (1994). *Mol. Gen. Genet.* **245**, 25–31.
- Trigg, P. I. & Kondrachine, A. V. (1998). *Bull. World Health Organ.* **76**, 11–16.
- Vontas, J. G., Small, G. J. & Hemingway, J. (2001). *Biochem. J.* **357**, 65–72.
- Wang, B. C. & Rose, J. (2000). Personal communication.
- Wilce, M. C., Board, P. G., Feil, S. C. & Parker, M. W. (1995). *EMBO J.* **14**, 2133–2143.