Acta Crystallographica Section F
Structural Biology Communications
ISSN 2053-230X

Expression, crystallization and preliminary X-ray diffraction analysis of thioredoxin glutathione reductase from *Schistosoma japonicum* in complex with FAD

Yongdong Li,[a,b] Qunfeng Wu,[a] Yun Peng,[a] Fuyan Huang,[a] Xun Li,[a] Lin Chen,[a] Dashuang Shi,[a,c] Xiaonong Zhou[b]* and Xiaolin Fan[a]*

[a]Key Laboratory of Organo-Pharmaceutical Chemistry, Jiangxi Province, Chemistry and Chemical Engineering College, Gannan Normal University, Ganzhou 341000, People's Republic of China, [b]National Institute of Parasitic Diseases, Chinese Center for Disease Control and Prevention, Shanghai, People's Republic of China, and [c]Center for Genetic Medicine Research and Department of Integrative Systems Biology, Children's National Medical Center, The George Washington University, Washington, DC 20010, USA

Correspondence e-mail: xiaonongzhou1962@gmail.com, vanxl@gnnu.edu.cn

Received 23 September 2013
Accepted 7 December 2013

Thioredoxin glutathione reductase from *Schistosoma japonicum* (SjTGR), a multifunctional enzyme, plays a vital role in antioxidant pathways and is considered to be a potential drug target for the development of antischistosomal chemotherapy. In this study, two constructs of a truncated form of SjTGR without the last two residues (Sec597–Gly598) were cloned, overexpressed and purified using wild-type and codon-optimized genes. Only SjTGR from the wild-type gene was found to form a complex with flavin adenine dinucleotide (FAD), which could be crystallized in the orthorhombic space group $P2_12_12_1$, with unit-cell parameters a = 84.185, b = 86.47, c = 183.164 Å, at 295 K using the hanging-drop vapour-diffusion method. One dimer was present in the crystallographic asymmetric unit and the calculated Matthews coefficient (V_M) and solvent content were 2.6 Å^3 Da^{-1} and 52.8%, respectively. Structural determination of SjTGR is in progress using the molecular-replacement method.

1. Introduction

Thioredoxin glutathione reductase (TGR) from *Schistosoma japonicum* (SjTGR), a unique multifunctional enzyme, performs the reduction of GSSG (oxidized glutathione) and Trx (thioredoxin) in *S. japonicum* redox systems, which is normally carried out by thioredoxin reductase (TrxR) and glutathione reductase (GR) in mammals (Alger & Williams, 2002; Johansson *et al.*, 2005; Song *et al.*, 2012). Platyhelminth parasites possess a unique redox system without canonical enzymes in comparison with free-living flatworms. Hence, TGR plays a key role in parasite survival and is normally regarded as a drug target for platyhelminth infections (Otero *et al.*, 2010). Research results have indicated that the multifunctional enzyme TGR meets all of the major criteria for a drug target in the anti-oxidant pathway (Kuntz *et al.*, 2007; Han *et al.*, 2012). TGR from schistosomes has been used as an essential drug target for the development of antischistosomal chemotherapy (Song *et al.*, 2012; Kuntz *et al.*, 2007).

The efficiency of recombinant protein production in an *Escherichia coli* system may be diminished by codon-usage bias (Plotkin & Kudla, 2011; Ikemura, 1985; Drummond & Wilke, 2008). Ways to overcome this problem include targeted mutagenesis to change rare codons or the addition of co-expressing genes encoding the rare codon tRNAs in *E. coli* (Gustafsson *et al.*, 2004; Brinkmann *et al.*, 1989; Springer & Sligar, 1987; Kleber-Janke & Becker, 2000). However, many genes show little bias in codon usage and reflect a lack of selection for mRNA translation. For instance, the expression of *Neurospora* FREQUENCY (FRQ) protein with an optimized sequence shows non-optimal codon usage for its biological function. Recombinant FRQ expression using the native gene leads to an optimal protein structure and function. However, the expression of the optimized gene leads to conformational changes and functional disorder (Zhou *et al.*, 2013).

In this study, a codon-optimized SjTGR (coSjTGR) gene was synthesized directly and the gene for wild-type SjTGR (wtSjTGR) was amplified from an *S. japonicum* cDNA library. We cloned and expressed both SjTGR genes in an *E. coli* system and purified both recombinant SjTGR proteins (wtSjTGR and coSjTGR) for crystallization. Interestingly, the expression and crystallization of wtSjTGR and coSjTGR differed significantly. Only wtSjTGR was found to form

a complex with flavin adenine dinucleotide (FAD) and to produce crystals suitable for structure determination.

2. Materials and methods

2.1. Codon optimization and gene cloning

Codon optimization was executed using *Optimizer* as created by Puigbò *et al.* (2007, 2008). We selected the codon usage of predicted highly expressed genes and the 'one amino acid–one codon' method. All codons of the query sequence were optimized and substituted by the synonymous codon used most frequently in the reference set (Puigbò *et al.*, 2007, 2008). The optimized SjTGR gene was then directly synthesized by chemical methods, and vector construction was completed by Sangon Biotech. The gene of wtSjTGR was amplified *via* the polymerase chain reaction from a female *S. japonicum* cDNA library. Sequence information (GenBank AY814219.1) was used to synthesize or to design the forward and reverse oligonucleotide primers. The primer sequences, which contained *Nde*I and *Bam*HI restriction sites (bold), were 5′-GTGCCGCGCGGCAGC–**CATATG**CCTCCGATTGAT-3′ and 5′-ACGGAGCTCGAATTC–**GGATCC**TCAGCAACCG-3′, in which the first 15 bases were designed for subsequent one-step cloning (ligation-independent cloning). The wtSjTGR gene (1791 bp) was cloned into expression

```
SjTGRaa.seq     M  P  P  I  D  G  T  S  Q  W  L  Q  R  T  I  E  S  A  A  V
wtSjTGR.seq     ATGCCTCCGATTGATGGAACATCCCAGTGGTTGCAGAGGACTATCGAATCAGCGGCGGTA
coSjTGR.seq     ATGCCGCCGATCGACGGTACCTCTCAGTGGCTGCAGCGTACCATCGAATCTGCGGCGGTT
                ***** ***** ** ** ** ** ****** ***** * ** ******** ********

SjTGRaa.seq     I  V  F  S  K  T  T  C  P  F  C  K  K  L  K  D  V  L  A  E
wtSjTGR.seq     ATCGTCTTTAGCAAAACAACTTGTCCATTTTGCAAAAAGCTAAAGGATGTTTTAGCTGAA
coSjTGR.seq     ATCGTTTTCTCTAAAACCACCTGCCCGTTCTGCAAAAAACTGAAAGACGTTCTGGCGGAA
                ***** **    ***** ** ** ** ** ******** ** ** ** *** * ** ***

SjTGRaa.seq     A  K  I  K  H  A  T  I  E  L  D  Q  L  S  N  G  S  V  I  Q
wtSjTGR.seq     GCAAAGATTAAACACGCTACAATTGAACTGGATCAATTATCCAATGGTTCGGTTATTCAA
coSjTGR.seq     GCGAAAATCAAACACGCGACCATCGAACTGGACCAGCTGTCTAACGGTTCTGTTATCCAG
                ** ** ** ******** ** ** ******** **  * ** ** ***** ***** **

SjTGRaa.seq     K  A  L  S  N  F  S  K  I  E  T  V  P  Q  M  F  V  R  G  K
wtSjTGR.seq     AAGGCATTATCTAACTTCTCTAAAATTGAAACAGTCCCGCAAATGTTTGTTAGAGGCAAG
coSjTGR.seq     AAAGCGCTGTCTAACTTCTCTAAAATCGAAACCGTTCCGCAGATGTTCGTTCGTGGTAAA
                ** **  * ****************** ***** ** ***** ***** *** * ** **

SjTGRaa.seq     F  I  G  D  S  K  A  V  L  N  Y  H  N  N  N  Q  L  Q  A  I
wtSjTGR.seq     TTCATTGGCGATTCTAAAGCAGTACTTAATTACCACAATAATAATCAATTGCAGGCGATC
coSjTGR.seq     TTCATCGGTGACTCTAAAGCGGTTCTGAACTACCACAACAACAACCAGCTGCAGGCGATC
                ***** ** ** ******** ** ** ** ******** ** ** **  ***********

SjTGRaa.seq     V  N  E  N  K  Y  D  Y  D  L  I  I  I  G  G  G  S  G  G  L
wtSjTGR.seq     GTCAACGAAAATAAGTATGACTATGATCTGATAATCATCGGTGGAGGATCTGGTGGACTC
coSjTGR.seq     GTTAACGAAAACAAATACGACTACGACCTGATCATCATCGGTGGTGGTTCTGGTGGTCTG
                ** ******** ** ** ***** ** ***** *********** ** ******** **

SjTGRaa.seq     A  A  G  K  E  A  A  K  Y  G  A  K  I  A  V  L  D  Y  V  E
wtSjTGR.seq     GCTGCTGGAAAGGAGGCAGCCAAATACGGCGCAAAGATAGCTGTTCTGGATTATGTAGAA
coSjTGR.seq     GCGGCGGGTAAAGAAGCGGCGAAATACGGTGCGAAAATCGCGGTTCTGGACTACGTTGAA
                ** ** ** ** ** ** ** ******** ** ** ** ** ******** ** ** ***

SjTGRaa.seq     P  T  P  M  G  T  T  W  G  L  G  G  T  C  V  N  V  G  C  I
wtSjTGR.seq     CCGACTCCAATGGGTACTACTTGGGGATTAGGTGGAACCTGTGTTAACGTTGGATGTATC
coSjTGR.seq     CCGACCCCGATGGGTACCACCTGGGGTCTGGGTGGTACCTGCGTTAACGTTGGTTGCATC
                ***** ** ******** ** *****  * ***** ***** *********** ** ***

SjTGRaa.seq     P  K  K  L  M  H  Q  A  G  L  L  S  H  S  L  E  D  A  Q  H
wtSjTGR.seq     CCTAAAAAATTAATGCACCAAGCTGGACTCTTAAGTCATTCTTTGGAAGATGCCCAACAT
coSjTGR.seq     CCGAAAAAACTGATGCACCAGGCGGGTCTGCTGTCTCACTCTCTGGAAGACGCGCAGCAC
                ** ****** * ******** ** ** **  *   *** *** ******* ** ** **

SjTGRaa.seq     F  G  W  S  L  D  K  S  E  I  S  H  D  W  S  T  M  V  E  G
wtSjTGR.seq     TTCGGTTGGAGCTTGGATAAATCAGAAATTTCCCATGATTGGTCAACTATGGTTGAAGGA
coSjTGR.seq     TTCGGTTGGTCTCTGGACAAATCTGAAATCTCTCACGACTGGTCTACCATGGTTGAAGGT
                *********    **** ***** ***** ** ** ** ***** ** ***********

SjTGRaa.seq     V  Q  S  H  I  G  S  L  N  W  G  Y  K  V  S  L  R  D  N  A
wtSjTGR.seq     GTTCAGAGTCACATCGGTTCTTTAAATTGGGGCTATAAAGTTTCACTAAGAGATAATGCG
coSjTGR.seq     GTTCAGTCTCACATCGGTTCTCTGAACTGGGGTTACAAAGTTTCTCTGCGTGACAACGCG
                ******  ************* * ** ***** ** ******** **  * ** ** ***

SjTGRaa.seq     V  T  Y  L  N  A  R  G  M  L  L  S  P  H  E  V  Q  I  T  E
wtSjTGR.seq     GTTACGTATCTTAATGCTCGTGGGATGCTATTAAGTCCTCATGAGGTTCAGATTACAGAA
coSjTGR.seq     GTTACCTACCTGAACGCGCGTGGTATGCTGCTGTCTCCGCACGAAGTTCAGATCACCGAA
                ***** ** ** ** ***** *****  *   *** ** ** ******** ** ***

SjTGRaa.seq     K  N  K  K  V  S  T  I  T  G  N  K  I  I  L  A  T  V  E  R
wtSjTGR.seq     AAGAATAAAAAAGTATCCACAATAACTGGAAATAAAATCATCTTAGCTACTGTCGAGCGT
coSjTGR.seq     AAAAACAAAAAAGTTTCTACCATCACCGGTAACAAAATCATCCTGGCGACCGTTGAACGT
                ** ** ******** ** ** ** ** ** ** ********* * ** ** ** ** ***

SjTGRaa.seq     P  K  Y  P  E  I  P  G  A  I  E  Y  G  I  T  S  D  D  L  F
wtSjTGR.seq     CCAAAATACCCAGAAATACCTGGAGCAATCGAATATGGGATTACAAGTGATGATTTGTTT
coSjTGR.seq     CCGAAATACCCGGAAATCCCGGGTGCGATCGAATACGGTATCACCTCTGACGACCTGTTC
                ** ******** ***** ** ** ** ******** ** ** **   *** **  ****
```

Figure 1
The wtSjTGR and coSjTGR genes. Asterisks indicate codon changes.

vector pET-28a with an N-terminal His_6 tag and a thrombin recognition site (MGSS**HHHHHH**SSG**LVPRGS**H). The resulting positive plasmid was confirmed using DNA sequencing and transformed into *E. coli* BL21 (DE3) strain.

2.2. Expression and purification

Protein expression was induced with 0.2 m*M* IPTG at room temperature (about 295 K). After overnight incubation, the cells from 1 l of culture were harvested by centrifugation at 10 000*g* for 15 min (Eppendorf 5804 R) and suspended in 20 ml Ni-affinity lysis buffer [300 m*M* NaCl, 50 m*M* NaH_2PO_4 pH 7.5, 10%(*v*/*v*) glycerol, 5 m*M* β-mercaptoethanol]. Cells were disrupted by sonication for 4 × 40 s on ice at maximum power. Cellular debris was removed by centrifugation for 30 min at 12 000*g* and passage through a 0.45 µm filter.

The protein was purified using a HisTrap Ni-affinity column (GE Healthcare) followed by a HiTrap DEAE column (GE Healthcare) using an ÄKTA FPLC system. The soluble fraction was loaded onto a 5 ml HisTrap Ni-affinity column previously equilibrated with lysis buffer. The column was washed with lysis buffer containing 50 m*M* imidazole and the protein was eluted using lysis buffer containing 250 m*M* imidazole. The protein was then dialyzed into a buffer consisting of 50 m*M* NaCl, 20 m*M* Tris–HCl pH 8.0, 2 m*M* EDTA, 5 m*M* β-mercaptoethanol and loaded onto a DEAE HiTrap column previously equilibrated with the dialysis buffer. The final pure protein was eluted using the linear gradient method using a buffer consisting of 500 m*M* NaCl, 20 m*M* Tris–HCl pH 8.0, 2 m*M* EDTA. The recombinant protein was soluble and the yield was about 20 mg per litre of culture. Protein purity was assessed by SDS–PAGE followed by Coomassie Blue staining, and protein concentration was determined by the Bradford assay with bovine serum albumin as a standard (Bradford, 1976). The protein was concentrated to about 20 mg ml^{-1} using an Amicon Y5 membrane concentrator (Millipore) and stored at 193 K until use.

2.3. Crystallization

Crystallization trials using both SjTGR samples (wtSjTGR and coSjTGR) were performed at 295 K using the sitting-drop vapour-diffusion method in CrystalEX (Corning) 96-well plates with PEG/Ion 2 from Hampton Research. wtSjTGR crystals were observed in two conditions [0.1 *M* sodium formate pH 7.0, 12% PEG 3350 and 8%(*v*/*v*) Tacsimate pH 6.0, 20% PEG 3350] after 3 d and the conditions were optimized using laboratory-made solutions *via* the hanging-drop vapour-diffusion method in a 24-well Linbro plate. Finally, wtSjTGR crystals obtained from a condition consisting of 0.2 *M* potassium sodium tartrate tetrahydrate, 0.1 *M* HEPES pH 7.0, 20% PEG 3350 were used for data collection. coSjTGR microcrystals were observed using 0.2 *M* KSCN, 20% PEG 3350 pH 7.0. However, no coSjTGR crystals suitable for X-ray diffraction were obtained even after extensive optimization trials.

2.4. Data collection and processing

Cryoprotection of the crystals was achieved by rapid soaking in the reservoir solution supplemented with 25%(*v*/*v*) glycerol. Diffraction data were collected in a stream of cold nitrogen at 100 K on beamline BL17U1 at the Shanghai Synchrotron Radiation Facility (SSRF), People's Republic of China. X-ray diffraction data were collected using an ADSC Q315r CCD detector at a wavelength of 0.9707 Å. The crystal was oscillated by 1.0° per frame over a range of 360° to obtain sufficient multiplicity. The data sets were scaled and merged using the *HKL*-2000 program package (Otwinowski & Minor, 1997).

2.5. Structure determination

The structure of SjTGR was solved by the molecular-replacement method using *PHENIX* (Adams *et al.*, 2010). *Coot* (Emsley & Cowtan, 2004) was used for model building and manual adjustments.

3. Results and discussion

An alignment of the wild-type and the codon-optimized SjTGR genes (wtSjTGR and coSjTGR) indicating the codon changes is shown in Fig. 1. The expression of SjTGR from *E. coli* BL21 strain with recombinant pET-28a-wtSjTGR and with pET-28a-coSjTGR vectors successfully yielded soluble SjTGRs from the cell lysate. Interestingly, the recombinant proteins produced from different vectors had different colours: wtSjTGR was yellow and coSjTGR was colourless

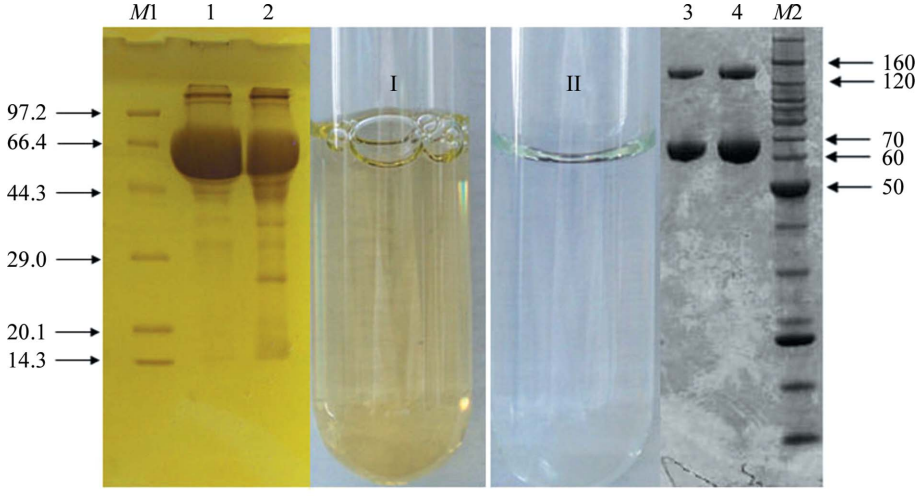

Figure 2
SDS–PAGE analysis of wtSjTGR and coSjTGR. (*a*) Purified wtSjTGR shown on a 12% polyacrylamide SDS gel. Lane *M*1, molecular-weight marker (labelled in kDa); lanes 1 and 2, eluted fraction from DEAE column. (*b*) Purified coSjTGR shown on a 12% polyacrylamide SDS gel. Lane *M*2, molecular-weight marker (labelled in kDa); lanes 3 and 4, eluted fraction from DEAE column. I, solution sample of wtSjTGR. II, solution sample of coSjTGR.

(Fig. 2). Preliminary analysis of the electron density suggests that FAD molecules present in wtSjTGR generate the yellow colour.

Crystals of wtSjTGR (Fig. 3) suitable for X-ray diffraction data collection (0.3 × 0.3 × 0.2 mm) were obtained by mixing 2 µl 20 mg ml^{-1} protein solution with 2 µl reservoir solution consisting of 0.2 *M* potassium sodium tartrate tetrahydrate, 20%(*w*/*v*) PEG 3350 in bis-tris buffer pH 6.8. Diffraction data were collected to 2.4 Å resolution (Fig. 3). The data-collection and processing statistics are given in Table 1. Auto-indexing was performed with *HKL*-2000 and the results indicated that the crystals belonged to space group $P2_12_12_1$, with unit-cell parameters a = 84.185, b = 86.47, c = 183.164 Å. One dimer exists in the asymmetric unit, with a corresponding V_M of 2.6 Å^3 Da^{-1} and solvent content of 52.8% (Matthews, 1968). Only small crystals were observed for coSjTGR using the same methods and screening kit.

In this study, high-level expression of both wtSjTGR and coSjTGR genes was achieved in an *E. coli* system. wtSjTGR contained bound FAD, whereas coSjTGR had no bound cofactor. Single crystals suitable for X-ray diffraction data collection were only obtained for wtSjTGR. However, only microcrystals of coSjTGR (Fig. 2) could be grown even after condition optimization (Fig. 2). These results may suggest that codon usage regulates not only the protein expression level but also protein folding. Protein chaperones and sufficient time are required for protein folding. Codon optimization leads to higher translation rates and reduces the time available for folding. Codon optimization is probably the main reason that coSjTGR presents non-optimal codon usage for FAD capture, and affects the protein folding. Previous studies have indicated that codon usage regulates protein folding (Komar *et al.*, 1999; Siller *et al.*, 2010; Zhou *et al.*, 2009), and can change the protein structure and function in some cases (Zhou *et al.*, 2013; Kimchi-Sarfaty *et al.*, 2007).

Table 1
X-ray diffraction data and processing statistics.

Values in parentheses are for the outermost resolution shell.

X-ray source	BL17U1, SSRF
Wavelength (Å)	0.9707
Temperature (K)	100
Detector	ADSC Quantum 315r
Space group	$P2_12_12_1$
Resolution (Å)	43.24–2.35 (2.432–2.35)
Unit-cell parameters (Å, °)	a = 84.185, b = 86.470, c = 183.164, $\alpha = \beta = \gamma$ = 90
Total reflections	56540
Unique reflections	5506
Completeness (%)	99.9 (99.23)
R_{merge}† (%)	14.5 (94.8)
$\langle I/\sigma(I)\rangle$	20.44 (4.0)
Multiplicity	7.1 (7.3)

† $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of the ith observation of unique reflection hkl and $\langle I(hkl)\rangle$ is the average over symmetry-related observations of unique reflection hkl.

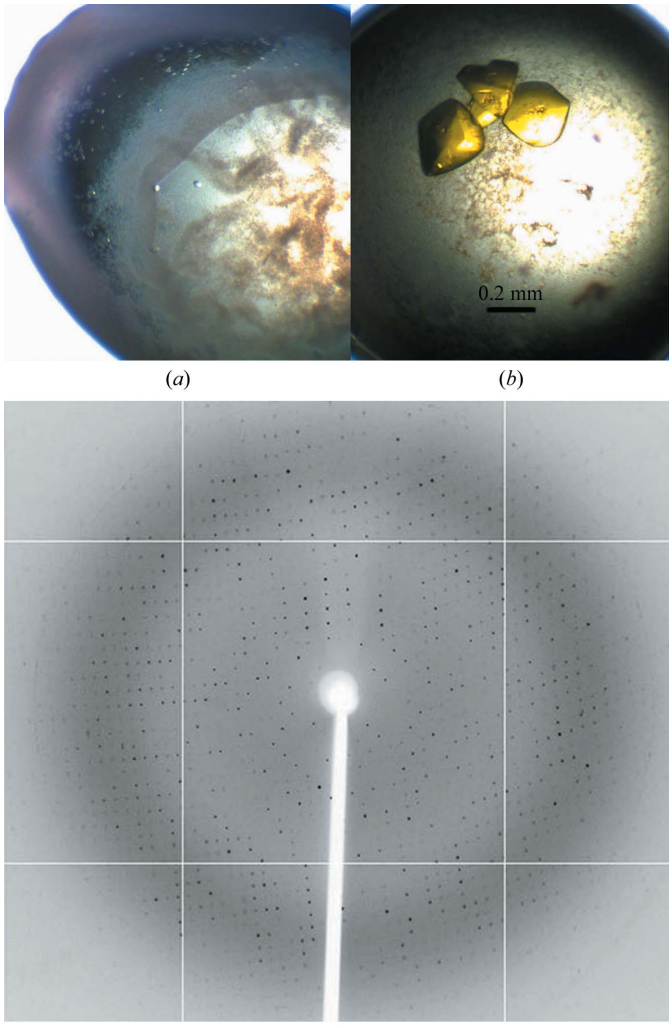

Figure 3
Crystals of coSjTGR and wtSjTGR. (*a*) Microcrystals of coSjTGR grown with PEG 3350 as precipitant. (*b*) Typical wtSjTGR crystals with dimensions of 0.3 × 0.3 × 0.2 mm that diffracted to a resolution of 2.4 Å. (*c*) Diffraction pattern of wtSjTGR; data were detectable to 2.4 Å resolution.

We thank Dr Junhu Chen from the National Institute of Parasitic Diseases, Chinese Center for Disease Control and Prevention for kindly providing the plasmid from the *S. japonicum* cDNA library. We are grateful to the National Science Foundation of China (31260206, 30860073), for a General Financial Grant from the China Postdoctoral Science Foundation (2012M520352) and to the Fund for the Jiangxi 'Jinggang Star' Young Scientist Training Program (2011BCB23026).

References

Adams, P. D. *et al.* (2010). *Acta Cryst.* D**66**, 213–221.
Alger, H. M. & Williams, D. L. (2002). *Mol. Biochem. Parasitol.* **121**, 129–139.
Bradford, M. M. (1976). *Anal. Biochem.* **72**, 248–254.
Brinkmann, U., Mattes, R. E. & Buckel, P. (1989). *Gene*, **85**, 109–114.
Drummond, D. A. & Wilke, C. O. (2008). *Cell*, **134**, 341–352.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Gustafsson, C., Govindarajan, S. & Minshull, J. (2004). *Trends Biotechnol.* **22**, 346–353.
Han, Y., Zhang, M., Hong, Y., Zhu, Z., Li, D., Li, X., Fu, Z. & Lin, J. (2012). *Parasitol Int.* **61**, 475–480.
Ikemura, T. (1985). *Mol. Biol. Evol.* **2**, 13–34.
Johansson, L., Gafvelin, G. & Arnér, E. S. (2005). *Biochim. Biophys. Acta*, **1726**, 1–13.
Kimchi-Sarfaty, C., Oh, J. M., Kim, I.-W., Sauna, Z. E., Calcagno, A. M., Ambudkar, S. V. & Gottesman, M. M. (2007). *Science*, **315**, 525–528.
Kleber-Janke, T. & Becker, W. M. (2000). *Protein Expr. Purif.* **19**, 419–424.
Komar, A. A., Lesnik, T. & Reiss, C. (1999). *FEBS Lett.* **462**, 387–391.
Kuntz, A. N., Davioud-Charvet, E., Sayed, A. A., Califf, L. L., Dessolin, J., Arnér, E. S. & Williams, D. L. (2007). *PLoS Med.* **4**, e206.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Otero, L., Bonilla, M., Protasio, A. V., Fernández, C., Gladyshev, V. N. & Salinas, G. (2010). *BMC Genomics*, **11**, 237.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Plotkin, J. B. & Kudla, G. (2011). *Nature Rev. Genet.* **12**, 32–42.
Puigbò, P., Guzmán, E., Romeu, A. & Garcia-Vallvé, S. (2007). *Nucleic Acids Res.* **35**, W126–W131.

Puigbò, P., Romeu, A. & Garcia-Vallvé, S. (2008). *Nucleic Acids Res.* **36**, D524–D527.
Siller, E., DeZwaan, D. C., Anderson, J. F., Freeman, B. C. & Barral, J. M. (2010). *J. Mol. Biol.* **396**, 1310–1318.
Song, L., Li, J., Xie, S., Qian, C., Wang, J., Zhang, W., Yin, X., Hua, Z. & Yu, C. (2012). *PLoS One*, **7**, e31456.
Springer, B. A. & Sligar, S. G. (1987). *Proc. Natl Acad. Sci. USA*, **84**, 8961–8965.
Zhou, M., Guo, J., Cha, J., Chae, M., Chen, S., Barral, J. M., Sachs, M. S. & Liu, Y. (2013). *Nature (London)*, **495**, 111–115.
Zhou, T., Weems, M. & Wilke, C. O. (2009). *Mol. Biol. Evol.* **26**, 1571–1580.