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

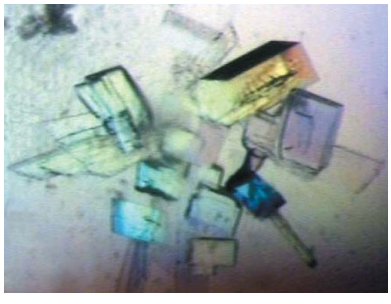
Crystallization and X-ray analysis of the *Schistosoma mansoni* guanidino kinase

Ayman M. Awama,[a] Patricia Paracuellos,[b] Sabine Laurent,[a] Colette Dissous,[c] Olivier Marcillat[a] and Patrice Gouet[d]*

[a]Institut de Chimie et Biochimie Moléculaire et Supramoléculaire, CNRS UMR 5246 et Université de Lyon, F-69622 Villeurbanne, France, [b]Phosphorylation des Protéines et Pathogénie Bactérienne, Institut de Biologie et Chimie des Protéines, CNRS et Université de Lyon, UMR 5086, IFR 128 'BioSciences Gerland-Lyon Sud', F-69367 Lyon CEDEX 07, France, [c]Institut Pasteur de Lille, INSERM U547–Université Lille 2, F-59019 Lille, France, and [d]Laboratoire de BioCristallographie, Institut de Biologie et Chimie des Protéines, CNRS et Université de Lyon, UMR 5086, IFR 128 'BioSciences Gerland-Lyon Sud', F-69367 Lyon CEDEX 07, France

Correspondence e-mail: p.gouet@ibcp.fr

Received 23 May 2008
Accepted 12 August 2008



The 716-amino-acid guanidino kinase from the parasitic flatworm *Schistosoma mansoni* results from the fusion of two guanidino kinase subunits. Crystals of this 80 kDa protein have been obtained in the monoclinic space group $P2_1$, with unit-cell parameters $a = 52.7$, $b = 122.1$, $c = 63.2$ Å, $\beta = 108.5^\circ$. Synchrotron data were collected to 2.8 Å resolution on ESRF beamline ID29. The structure was solved by the molecular-replacement method, using the 357-amino-acid structure of the arginine kinase from *Trypanosoma cruzi* as the search model.

1. Introduction

Guanidino kinases (GKs) or ATP:guanidino phosphotransferases constitute a family of enzymes that catalyse the reversible transfer of the ATP γ-phosphate to a guanidylic acceptor. GKs break the phosphodiester bond between O3β and Pγ of ATP and create a new bond between the released γ-phosphate and the primary amine of the guanidylic compound. Several phosphorylated guanidines, called phosphagens, have been described in a wide range of eukaryotic organisms. These compounds play a prominent role in cellular ATP-concentration buffering and differ in the substituents of the guanidine moiety.

The best known GKs are creatine kinase (CK) and arginine kinase (AK), but other GKs have also been described (Kassab *et al.*, 1965; Van Thoai *et al.*, 1966; Surholt, 1979; Suzuki & Furukohri, 1994; Suzuki *et al.*, 1997; Uda, Iwai *et al.*, 2005; Uda, Saishoji *et al.*, 2005; Uda, Tanaka *et al.*, 2005).

Sequence analysis of GKs (Muhlebach *et al.*, 1994) shows that these enzymes have an N-terminal 'specificity loop' (residues 60–80) with a length that is inversely proportional to the size of the substrate (Suzuki *et al.*, 1997). Thus, glycocyamine kinase, which catalyses the synthesis of the smallest phosphagen, exhibits the longest loop. CK has a single amino-acid deletion in the loop compared with glycocyamine kinase, in keeping with its slightly longer substrate. Arginine, taurocyamine, hypotaurocyamine and lombricine kinases, which allow the synthesis of large phosphagens, have a five-amino-acid deletion.

All three-dimensional structures of monomeric GKs known to date are composed of a small N-terminal domain of ~110 amino acids connected to a large C-terminal domain of ~270 amino acids (McLeish & Kenyon, 2005), as observed in the first crystallographic structure of a CK published by Fritz-Wolf *et al.* (1996). The ATP molecule binds in a cleft located between the two domains. GKs exist in a variety of oligomeric forms: monomeric or dimeric for AK (Cheung, 1973; Storey, 1977; Ratto & Christen, 1988), dimeric for cytosolic CK (Dawson *et al.*, 1965) and other less well characterized GKs (Kassab *et al.*, 1965; Brevet *et al.*, 1975) and octameric for mitochondrial CK (Marcillat *et al.*, 1987) and mitochondrial taurocyamine kinase (Uda, Saishoji *et al.*, 2005).

Atomic structures have been reported of dimeric CKs from rabbit muscle, chicken brain, human muscle and bovine retinal brain (Rao *et al.*, 1998; Eder *et al.*, 1999; Shen *et al.*, 2001; Tisi *et al.*, 2001), of octameric CKs from chicken sarcomeric mitochondria and human ubiquitous mitochondria (Fritz-Wolf *et al.*, 1996; Eder *et al.*, 2000) and of the monomeric AK from *Trypanosoma cruzi* (Fernandez *et al.*, 2007). In addition, crystallographic studies performed on *Limulus polyphemus* AK have provided a comprehensive view of the catalytic mechanism. These studies include the structural determination of the native form (Yousef *et al.*, 2003), the MgADP and NO_3^- complex (Pruett *et al.*, 2003), the MgADP and creatine complex (Azzi *et al.*, 2004) and the MgADP, NO_3^- and arginine transition-state analogue complex (TSAC; Zhou *et al.*, 1998; Yousef *et al.*, 2002). In the AK TSAC structure, the NO_3^- ion mimics the γ-phosphate as it is transferred between phosphoarginine and ADP. Together, these structures confirm previous small-angle X-ray scattering experiments showing a large conformational change (Dumas & Janin, 1983; Forstner *et al.*, 1996, 1998) from a substrate-free (open) form to a substrate-bound (closed) form. The closed form is obtained by hinge-bending domain displacement as observed in the *Torpedo californica* muscle-type CK complex with MgADP, NO_3^- and creatine (Lahiri *et al.*, 2002).

Schistosoma mansoni is a trematode flatworm that has been identified as the aetiological agent of schistosomiasis, a disease that affects about 200 million people worldwide. A protein expressed by this organism was classified as a guanidino kinase based on sequence similarity and preliminary enzymatic assays (Stein *et al.*, 1990). The *S. mansoni* guanidino kinase (SmGK) transcript is specifically associated with the cercaria life stage (Stein *et al.*, 1990; Curwen *et al.*, 2004) and is a potential drug target. SmGK cDNA has an open reading frame of 716 amino acids corresponding to a duplicated structure. The first (1–363) and second (364–716) subunits present a significant sequence identity of 53% (Fig. 1). Hence, this two-subunit protein mimics GK dimers such as muscle or brain creatine kinases and is likely to possess two active sites (Stein *et al.*, 1990). The recombinant enzyme shows catalytic activity with several guanidylic compounds, the highest activity being measured with taurocyamine. In SmGK, the loop 60–80 shows an uncharacterized six-amino-acid deletion in each equivalent subunit, suggesting that taurocyamine may not be the physiological substrate and that *S. mansoni* may use another phosphagen.

Here, we describe the purification and crystallization of recombinant SmGK and its structure determination by X-ray diffraction. This constitutes the first structural study of a GK other than AK and CK.

2. Protein expression and purification

2.1. Cloning of the SmGK cDNA

Total RNA was isolated from *S. mansoni* cercariae (Puerto Rican strain) by the method of Chirgwin *et al.* (1979) and purified using a cesium chloride gradient. First-strand cDNA synthesis was performed using the Thermoscript RT-PCR System (Invitrogen) and the oligo(dT)$_{20}$ primer. The full-length SmGK coding sequence (residues 1–716) was amplified using primers (SmCKf, 5′-GTGCTCAATCCATATGCAGGTTG-3′; SmCKr, 5′-GGAGCTCTAAAGTAATATTACAGTCCTTTCTCG-3′) complementary to the cDNA sequence (Genbank accession Nos. J05410 and EU042595; Stein *et al.*, 1990; Roger *et al.*, 2008) and confirmed by genomic data from the *S. mansoni* Genome Project (http://www.sanger.ac.uk/Projects/S_mansoni/). PCR was carried out with the PlatinumTaq Hi-Fi (Invitrogen) and PCR products were cloned into pCR 2.1 TOPO (Invitrogen). The SmGK coding sequence was further cloned into a pET21 plasmid (Novagen) without any tag and the pETSmGK construct was completely sequenced.

2.2. Expression of SmGK

The pETSmGK construct was used to transform the *Escherichia coli* BL21 (DE3) strain. A 25 ml overnight culture at 310 K was used to inoculate 1 l of 2% bactotryptone, 1% yeast extract, 0.5% NaCl, 0.2% glycerol, 50 m*M* K_2HPO_4 pH 7.2 supplemented with 50 µg ampicillin. The cells were grown at 303 K to an OD_{600} of 0.6 and

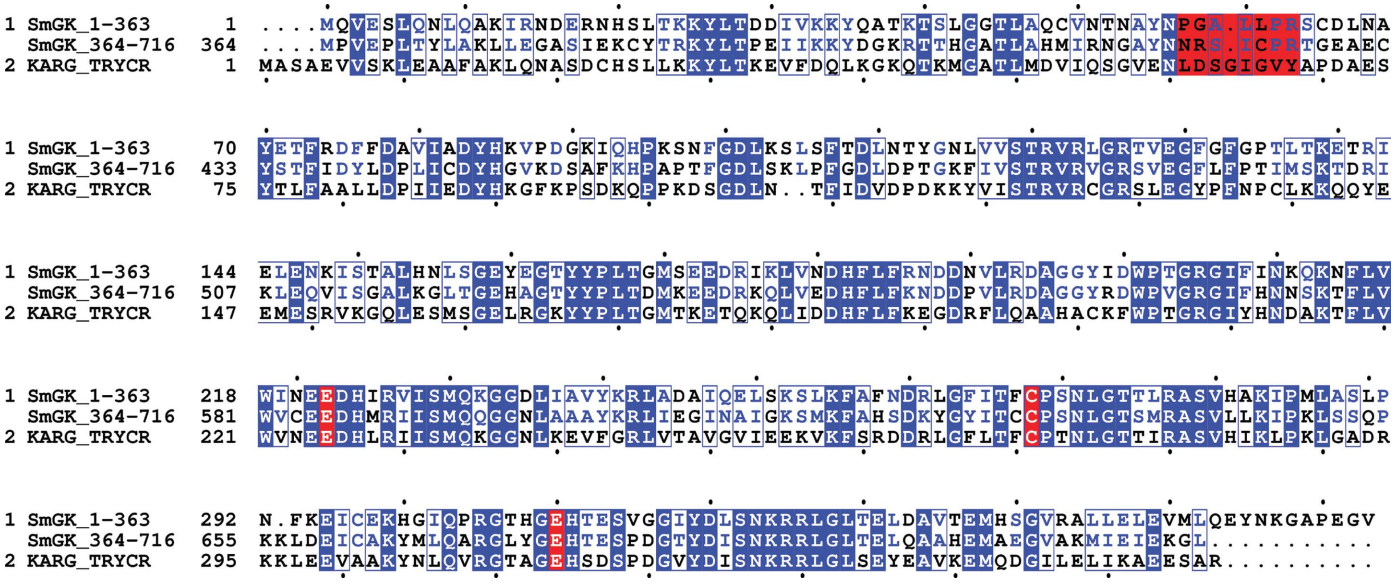


Figure 1
Multiple sequence alignment of the first subunit (1–363) of the guanidino kinase from *S. mansoni* (SmGK), the second subunit of SmGK (364–716) and *T. cruzi* AK. Identical residues are depicted in white on a blue background; similar residues are shown in blue type. Functionally important residues in guanidino kinases are highlighted with a red background (*i.e.* the N-terminal 'specificity loop', an invariant cysteine which plays a role in substrate binding and two glutamic acids which are putative catalytic residues). The figure was drawn with *ESPript* (Gouet *et al.*, 1999).

Table 1
Data-collection statistics for SmGK.

Values in parentheses are for the highest resolution shell.

Space group	$P2_1$
Unit-cell parameters (Å, °)	a = 52.7, b = 122.1, c = 63.2, β = 108.5
Maximum resolution (Å)	2.8 (2.9–2.8)
No. of observations	68251 (6251)
No. of unique reflections	18324 (1737)
Completeness (%)	98.9 (94.6)
$\langle I/\sigma(I)\rangle$	15.8 (5.5)
Redundancy	3.7 (3.6)
$R_{\rm merge}$† (%)	7.9 (27.0)
$V_{\rm M}$ (Å^3 Da^{-1})	2.4

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the ith measurement of reflection $I(hkl)$.

protein overexpression was induced in turn by the addition of 1 m*M* IPTG.

2.3. Purification of recombinant SmGK

5 h after induction, cells from a 1 l culture were pelleted, washed, resuspended in 50 ml 20 m*M* sodium phosphate buffer pH 6.0 and then lysed on ice by sonication. Cell debris was eliminated by centrifugation and the supernatant was loaded onto a Blue Sepharose 6 Fast Flow column (GE Healthcare) equilibrated with 20 m*M* sodium phosphate pH 6.0. The SmGK was eluted with 20 m*M* Tris pH 8.5 buffer. Enzymatically active fractions were pooled and loaded onto a Source 15 Q column (GE Healthcare) equilibrated in 20 m*M* Tris pH 8.5. Under these conditions, the enzyme elutes in the flow-through. This step substantially improves the purity by removing nucleotides that otherwise remain bound to the enzyme. SmGK was concentrated to 10 mg ml^{-1} by lyophilization and resolubilized in MilliQ water. The purity was verified by the presence of a single band on overloaded Coomassie Blue-stained sodium dodecyl sulfate–polyacrylamide gels. The protein concentration was determined spectrophotometrically (Lowry *et al.*, 1951; Bensadoun & Weinstein, 1976).

2.4. Taurocyamine kinase activity

The enzymatic activity at 303 K was measured by the pH-stat method using MgATP and taurocyamine as substrates. Typically, the reaction was started by the addition of 2.5 p*M* enzyme to 2 ml pH 8.2 standard assay medium containing 5 m*M* ATP, 5 m*M* magnesium acetate, 30 m*M* taurocyamine, 50 m*M* potassium acetate, 0.1 m*M* ethylenediaminetetraacetic acid and 5 m*M* DTT.

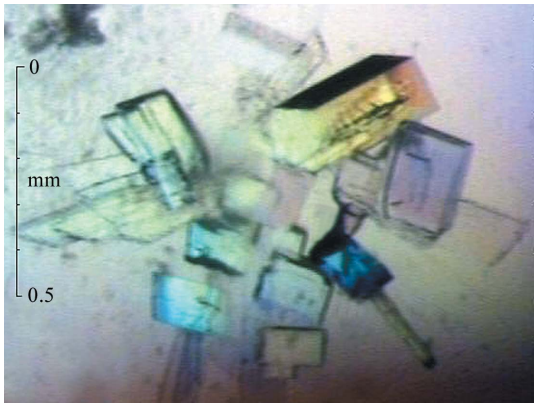


Figure 2
Crystals of *S. mansoni* guanidino kinase obtained at 290 K using the sitting-drop vapour-diffusion technique. Crystals with dimensions of 0.2 × 0.1 × 0.1 mm were obtained within 60 d.

3. Crystallization

Crystallization conditions were screened at 290 K by the sitting-drop vapour-diffusion method using commercial kits from Hampton Research, Molecular Dimensions and Nextal (Qiagen). Drops containing 2.5 µl protein solution at 10 mg ml^{-1} and 2.5 µl crystallization condition were equilibrated over 100 µl reservoir solution. One hit was found using a condition consisting of 25%(*w*/*v*) PEG 3000 and 100 m*M* MES pH 6.5. Crystals appeared within 20 d and reached dimensions suitable for diffraction experiments (0.2 × 0.1 × 0.1 mm) within 60 d (Fig. 2).

4. Preliminary X-ray analysis and structure determination

Prior to X-ray data collection, crystals were briefly soaked (30 s) in a cryoprotectant solution consisting of 5%(*v*/*v*) ethylene glycol added to the crystallization solution. They were tested at 100 K on a 345 mm MAR Research image-plate system using Cu $K\alpha$ radiation from an in-house Nonius FR591 rotating-anode X-ray generator. Crystals diffracted to 3.4 Å resolution in-house and belonged to space group $P2_1$, with unit-cell parameters a = 52.7, b = 122.1, c = 63.2 Å, β = 108.5°. The calculated $V_{\rm M}$ value is 2.4 Å^3 Da^{-1} for one molecule in the asymmetric unit, which corresponds to a solvent content of 49% (Matthews, 1968).

Synchrotron data were subsequently collected at 100 K to 2.8 Å resolution on ESRF beamline ID29, Grenoble, France. The crystal-to-detector distance was 367 mm, the oscillation step was 1° and the total oscillation range was 180°. The *XDS* program package (Kabsch, 1993) was used for data reduction and scaling (statistics are given in Table 1).

A *BLAST* search performed against the sequence database derived from the PDB showed that each ∼350-amino-acid SmGK subunit had a maximum sequence identity of ∼40% to the AK from the invertebrate *T. cruzi* (Fernandez *et al.*, 2007; Fig. 1). Thus, molecular-replacement calculations were carried out with the

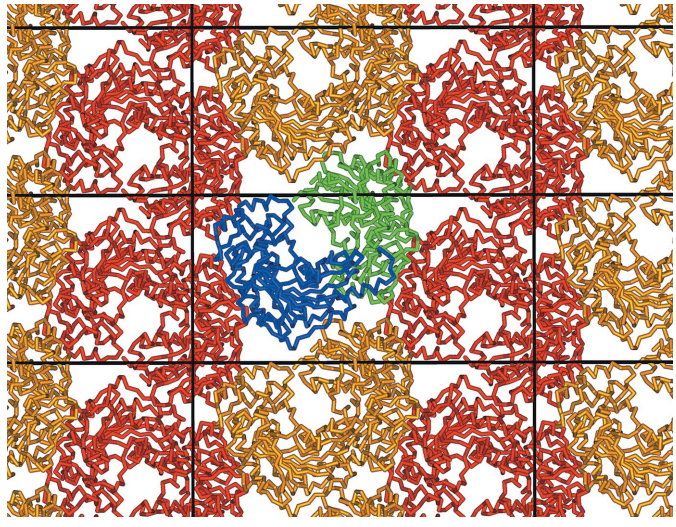


Figure 3
Crystal packing of *S. mansoni* guanidino kinase viewed down the *a* axis. Monomers are shown in orange and red. The monomer in the centre is displayed with its N-terminal subunit (1–363) in blue and its C-terminal subunit (364–716) in green. This figure was generated using *DS Visualizer* (Accelrys).

program *AMoRe* (Navaza, 2001) using the 1.9 Å substrate-free structure of *T. cruzi* AK (PDB code 2j1q) as the search model and a two-body search (the AK consists of 357 amino acids, while the asymmetric unit of the crystal contains one SmGK monomer of 716 amino acids). A clear solution was found with an *R* factor of 46% and a correlation coefficient of 0.35 to 3.4 Å resolution. 120 cycles of restrained positional refinement were performed with *CNS* (Brünger *et al.*, 1998) at 2.8 Å resolution and the *R* factor was reduced to 36% ($R_{\rm free}$ = 42%). Electron-density maps were calculated and displayed on a graphics workstation using the program *TURBO-FRODO* (Roussel & Cambillau, 1989). The $2F_{\rm o} - F_{\rm c}$ maps were of sufficient quality to observe an α-helix linking the two subunits of ∼350 amino acids in the asymmetric unit of the crystal and therefore unambiguously identified the monomeric structure of SmGK. This structure exhibits a U-shape and the C-terminal domain of the first subunit (1–363) interacts with the N-terminal domain of the second subunit (364–716) (Fig. 3). This arrangement of subunits differs from the dimeric arrangement observed in dimeric and octameric CKs, in which a pair of monomers are related by a twofold rotation axis. Crystallographic refinement and molecular rebuilding of SmGK are in progress.

The authors thank the ESRF staff at ID29 for technical support and assistance during X-ray data collection.

References

Azzi, A., Clark, S. A., Ellington, W. R. & Chapman, M. S. (2004). *Protein Sci.* **13**, 575–585.

Bensadoun, A. & Weinstein, D. (1976). *Anal. Biochem.* **70**, 241–250.

Brevet, A., Zeitoun, Y. & Pradel, L. A. (1975). *Biochim. Biophys. Acta*, **393**, 1–9.

Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-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.

Cheung, A. C. (1973). *Arch. Biochem. Biophys.* **154**, 28–39.

Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. (1979). *Biochemistry*, **18**, 5294–5299.

Curwen, R. S., Ashton, P. D., Johnston, D. A. & Wilson, R. A. (2004). *Mol. Biochem. Parasitol.* **138**, 57–66.

Dawson, D. M., Eppenberger, H. M. & Kaplan, N. O. (1965). *Biochem. Biophys. Res. Commun.* **21**, 346–353.

Dumas, C. & Janin, J. (1983). *FEBS Lett.* **153**, 128–130.

Eder, M., Fritz-Wolf, K., Kabsch, W., Wallimann, T. & Schlattner, U. (2000). *Proteins*, **39**, 216–225.

Eder, M., Schlattner, U., Becker, A., Wallimann, T., Kabsch, W. & Fritz-Wolf, K. (1999). *Protein Sci.* **8**, 2258–2269.

Fernandez, P., Haouz, A., Pereira, C. A., Aguilar, C. & Alzari, P. M. (2007). *Proteins*, **69**, 209–212.

Forstner, M., Kriechbaum, M., Laggner, P. & Wallimann, T. (1996). *J. Mol. Struct.* **383**, 217–222.

Forstner, M., Kriechbaum, M., Laggner, P. & Wallimann, T. (1998). *Biophys. J.* **75**, 1016–1023.

Fritz-Wolf, K., Schnyder, T., Wallimann, T. & Kabsch, W. (1996). *Nature (London)*, **381**, 341–345.

Gouet, P., Courcelle, E., Stuart, D. I. & Metoz, F. (1999). *Bioinformatics*, **15**, 305–308.

Kabsch, W. (1993). *J. Appl. Cryst.* **26**, 795–800.

Kassab, R., Pradel, L. A. & Van Thoai, N. (1965). *Biochim. Biophys. Acta*, **99**, 397–405.

Lahiri, S. D., Wang, P. F., Babbitt, P. C., McLeish, M. J., Kenyon, G. L. & Allen, K. N. (2002). *Biochemistry*, **41**, 13861–13867.

Lowry, O. H., Rosebrough, N. J., Farr, A. L. & Randall, R. J. (1951). *J. Biol. Chem.* **193**, 265–275.

McLeish, M. J. & Kenyon, G. L. (2005). *Crit. Rev. Biochem. Mol. Biol.* **40**, 1–20.

Marcillat, O., Goldschmidt, D., Eichenberger, D. & Vial, C. (1987). *Biochim. Biophys. Acta*, **890**, 233–241.

Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.

Muhlebach, S. M., Gross, M., Wirz, T., Wallimann, T., Perriard, J. C. & Wyss, M. (1994). *Mol. Cell. Biochem.* **133–134**, 245–262.

Navaza, J. (2001). *Acta Cryst.* D**57**, 1367–1372.

Pruett, P. S., Azzi, A., Clark, S. A., Yousef, M. S., Gattis, J. L., Somasundaram, T., Ellington, W. R. & Chapman, M. S. (2003). *J. Biol. Chem.* **278**, 26952–26957.

Rao, J. K., Bujacz, G. & Wlodawer, A. (1998). *FEBS Lett.* **439**, 133–137.

Ratto, A. & Christen, R. (1988). *Eur. J. Biochem.* **173**, 667–674.

Roger, E., Mitta, G., Moné, Y., Bouchut, A., Rognon, A., Grunau, C., Boissier, J., Théron, A. & Gourbal, B. E. F. (2008). *Mol. Biochem. Parasitol.* **157**, 205–216.

Roussel, A. & Cambillau, C. (1989). *TURBO-FRODO*. Silicon Graphics, Mountain View, California, USA.

Shen, Y., Tang, L., Zhou, H. & Lin, Z. (2001). *Acta Cryst.* D**57**, 1196–1200.

Stein, L. D., Harn, D. A. & David, J. R. (1990). *J. Biol. Chem.* **265**, 6582–6588.

Storey, K. B. (1977). *Arch. Biochem. Biophys.* **179**, 518–529.

Surholt, B. (1979). *Eur. J. Biochem.* **93**, 279–285.

Suzuki, T. & Furukohri, T. (1994). *J. Mol. Biol.* **237**, 353–357.

Suzuki, T., Kawasaki, Y., Furukohri, T. & Ellington, W. R. (1997). *Biochim. Biophys. Acta*, **1343**, 152–159.

Tisi, D., Bax, B. & Loew, A. (2001). *Acta Cryst.* D**57**, 187–193.

Uda, K., Iwai, A. & Suzuki, T. (2005). *FEBS Lett.* **579**, 6756.

Uda, K., Saishoji, N., Ichinari, S., Ellington, W. R. & Suzuki, T. (2005). *FEBS J.* **272**, 3521–3530.

Uda, K., Tanaka, K., Bailly, X., Zal, F. & Suzuki, T. (2005). *Int. J. Biol. Macromol.* **37**, 54–60.

Van Thoai, N., Di Jeso, F., Robin, Y. & Der Terrossian, E. (1966). *Biochim. Biophys. Acta*, **113**, 542–550.

Yousef, M. S., Clark, S. A., Pruett, P. K., Somasundaram, T., Ellington, W. R. & Chapman, M. S. (2003). *Protein Sci.* **12**, 103–111.

Yousef, M. S., Fabiola, F., Gattis, J. L., Somasundaram, T. & Chapman, M. S. (2002). *Acta Cryst.* D**58**, 2009–2017.

Zhou, G., Somasundaram, T., Blanc, E., Parthasarathy, G., Ellington, W. R. & Chapman, M. S. (1998). *Proc. Natl Acad. Sci. USA*, **95**, 8449–8454.