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

Structure of hypothetical Mo-cofactor biosynthesis protein B (ST2315) from *Sulfolobus tokodaii*

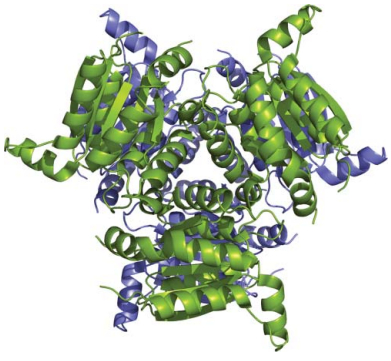
Svetlana V. Antonyuk,[a] Richard W. Strange,[a]* Mark J. Ellis,[b] Yoshitaka Bessho,[c,d] Seiki Kuramitsu,[d,e] Akeo Shinkai,[d] Shigeyuki Yokoyama[c,d,f] and S. Samar Hasnain[a]*

[a]Molecular Biophysics Group, School of Biological Sciences, University of Liverpool, Crown Street, Liverpool L69 7ZB, England, [b]STFC Daresbury Laboratory, Warrington, Cheshire WA4 4AD, England, [c]Systems and Structural Biology Center, Yokohama Institute, RIKEN, 1-7-22 Suehiro, Tsurumi, Yokohama 230-0045, Japan, [d]RIKEN SPring-8 Center, Harima Institute, 1-1-1 Kouto, Sayo, Hyogo 679-5148, Japan, [e]Department of Biological Sciences, Graduate School of Science, Osaka University, 1-1 Machikaneyama, Toyonaka, Osaka 560-0043, Japan, and [f]Department of Biophysics and Biochemistry, Graduate School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

Correspondence e-mail: r.strange@liverpool.ac.uk, s.s.hasnain@liverpool.ac.uk

Received 25 June 2009
Accepted 22 October 2009

PDB Reference: ST2315, 3iwt, r3iwtsf.



The structure of a probable Mo-cofactor biosynthesis protein B from *Sulfolobus tokodaii*, belonging to space group $P6_422$ with unit-cell parameters $a = b = 136.68$, $c = 210.52$ Å, was solved by molecular replacement to a resolution of 1.9 Å and refined to an R factor and R_{free} of 16.8% and 18.5%, respectively. The asymmetric unit contains a trimer, while the biologically significant oligomer is predicted to be a hexamer by size-exclusion chromatography. The subunit structure and fold of ST2315 are similar to those of other enzymes that are known to be involved in the molybdopterin- and molybdenum cofactor-biosynthesis pathways.

1. Introduction

As part of the RIKEN Structural Genomics Initiative, proteins from *Sulfolobus tokodaii* were targeted for high-throughput structure determination (Sugahara *et al.*, 2008). Using the Pfam database (Bateman *et al.*, 2004), the protein ST2315 is predicted to function in the biosynthesis pathway of the molybdopterin cofactor. Here, we report the crystal structure of ST2315 at 1.90 Å resolution.

Molybdopterin is an essential cofactor that is found at the active site in all molybdenum-containing enzymes except nitrogenase, where it has a role in electron transfer and in the redox function of the Mo atom (Schwarz & Mendel, 2006). The biosynthesis of molybdopterin is a five-step process that starts with guanosine triphosphate. This is converted by two enzymatic reactions, which are performed in *Escherichia coli* by MoaA and MoaC, to cyclic pyranopterin monophosphate (cPMP). This unstable intermediate (originally called 'precursor Z') is then converted to molybdopterin by the insertion of dithiolene sulfurs into the pterin side chain by the action of molybdopterin synthase. In *E. coli*, molybdopterin synthase consists of a heterodimer of MoaD and MoaE enzymes. Of these enzymes, MoaC has recently been structurally characterized in *S. tokodaii* (Yoshida *et al.*, 2008). The probable role of ST2315 in *S. tokodaii* is predicted to be the activation of sulfur-free molybdopterin synthase by the transfer of sulfur to its MoaD subunit (Schwarz & Mendel, 2006), similar to the function of MoaB in *E. coli*. It has been suggested that MoaB may also bind intermediate pterin compounds during molybdopterin biosynthesis (Bader *et al.*, 2004).

2. Materials and methods

2.1. Cloning, expression and purification

The gene encoding ST2315 (gi:15922647) was amplified *via* PCR using *S. tokodaii* strain 7 genomic DNA and was cloned into the pET-21a expression vector (Merck Novagen, Darmstadt, Germany). The expression vector was introduced into *E. coli* Rosetta (DE3) strain (Merck Novagen, Darmstadt, Germany) and the recombinant strain was cultured in 2.5 l LB medium containing 30 µg ml^{-1} chloramphenicol and 50 µg ml^{-1} ampicillin. The harvested cells (9 g) were lysed by sonication in 16 ml 20 m*M* Tris–HCl buffer pH 8.0 containing 500 m*M* NaCl, 5 m*M* β-mercaptoethanol and 1 m*M* phenylmethyl-

Table 1
Data-collection and refinement parameters.

Values in parentheses are for the highest resolution shell (1.95–1.90 Å).

Space group	$P6_422$
Unit-cell parameters (Å)	$a = b = 136.68$, $c = 210.52$
Resolution (Å)	30.0–1.90
Unique reflections	91510
Completeness (%)	98.5 (97.0)
Redundancy	9.0 (4.0)
$R_{\rm merge}$† (%)	6.5 (61.0)
$I/\sigma(I)$	22.0 (2.0)
R factor‡ (%)	16.8
$R_{\rm free}$‡ (%)	18.7
B factors (Å^2)	
Wilson plot	31.0
Protein	48.4
Water	50.9
Mg^{2+} ions	42.8
Glycerol	64.8
PEG	61.5
R.m.s. deviations	
Bond distances (Å)	0.015
Bond angles (°)	1.48
ESU§ (Å)	0.065

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the observed intensity and $\langle I(hkl)\rangle$ is the average intensity of multiple symmetry-related observations of that reflection. ‡ $R = \sum_{hkl} ||F_{\rm obs}| - |F_{\rm calc}|| / \sum_{hkl} |F_{\rm obs}|$. $R_{\rm free}$ is the same but calculated for a test set not used in structural refinement. § Estimated standard uncertainty based on maximum likelihood as implemented in *REFMAC*.

sulfonyl fluoride on ice. The cell lysate was heat-treated at 363 K for 12 min and centrifuged at 15 000*g* for 30 min at 277 K. The supernatant was desalted by fractionation on a HiPrep 26/10 column (GE Healthcare Biosciences). The sample was applied onto a Resource S column (GE Healthcare Biosciences) equilibrated with 20 m*M* MES buffer pH 6.0 and eluted with a linear (0–0.5 *M*) gradient of NaCl. The fractions that eluted in 0.35 *M* NaCl were further purified using a hydroxyapatite CHT10-I column (Bio-Rad Laboratories) with a linear gradient of 0.01–0.5 *M* potassium phosphate buffer pH 7.0. The target sample, which eluted in the 0.44 *M* potassium phosphate fraction, was collected and applied onto a HiLoad 16/60 Superdex 200pg column (GE Healthcare Biosciences) equilibrated with 20 m*M* Tri–HCl buffer pH 8.0 containing 200 m*M* NaCl. The protein sample was analyzed by SDS–PAGE and confirmed by N-terminal amino-acid sequencing. After concentration to 23.1 mg ml^{-1} by ultrafiltration, the protein yield was 21.9 mg from 9 g of cells.

2.2. Crystallization

Crystallization was performed by the microbatch-under-oil method at 291 K. 0.5 µl of crystallization reagent consisting of 0.1 *M* Tris–HCl buffer pH 7.0 containing 10%(*w*/*v*) PEG 8000 and 0.2 *M* $MgCl_2$ (Wizard II condition No. 43, Emerald BioSystems) was mixed with 0.5 µl 23.1 mg ml^{-1} protein solution. The mixture was then covered with 15 µl silicone and paraffin oil. Crystals suitable for X-ray data collection appeared within 3 d and reached final dimensions of 0.8 × 0.2 × 0.2 mm. The crystals (Fig. 1) were immersed in their reservoir solution with 25% glycerol added as cryoprotectant, flash-cooled in a nitrogen-gas stream and stored in liquid nitrogen prior to transport to the synchrotron.

2.3. Data collection and processing

Experiments were performed at the Daresbury Synchrotron Radiation Source (SRS) using the combined crystallography/X-ray absorption beamline 10.1, employing a Si(111) sagittally focused monochromator tuned to a wavelength of 1.074 Å. Diffraction data were recorded at 100 K from a single crystal. The crystal diffracted to a resolution of 1.9 Å. Images were recorded using a MAR Mosaic 225 CCD detector and processed (indexed, integrated and scaled) using *HKL*-2000 (Otwinowski & Minor, 1997). The crystal was found to belong to space group $P6_422$, with unit-cell parameters $a = b = 136.68$, $c = 210.52$ Å, $\gamma = 120°$. Data-collection and processing parameters are shown in Table 1.

2.4. Structure solution and refinement

Structure solution was accomplished by molecular replacement with *MOLREP* (Vagin & Teplyakov, 1997), using as a search model the molybdenum cofactor-biosynthesis protein B from *Bacillus cereus* with 39% identity and 67% similarity to the target sequence and consisting of a trimer of entry 1y5e (C. Chang, M. Zhou, J. Abdullah & A. Joachimiak, unpublished work) from the Protein Data Bank (Abola *et al.*, 1987). The optimum solution, with a score of 0.38 and an *R* factor of 55% (the next highest score and *R* factor were 0.23 and 61.9%, respectively), was found for one trimer in the asymmetric unit, equivalent to a high solvent content of 74%. Attempts to fit two trimers in the asymmetric unit in accordance with the Matthews coefficient produced a poor molecular-replacement model (score 0.28, *R* factor 59.3%). Rigid-body refinement to 3 Å resolution using *REFMAC* (Murshudov *et al.*, 1997) improved the *R* factor to 51%. Subsequently, several cycles of restrained refinement reduced the *R* factor and $R_{\rm free}$ to 38.5% and 41.7%, respectively. At this stage, significant rebuilding using the amino-acid sequence for *S. tokadaii* was carried out using the automated model-building routines of *ARP*/*wARP* (Perrakis *et al.*, 1999) combined with isotropic *B*-factor refinement using *REFMAC* and graphical intervention using *Coot* (Emsley & Cowtan, 2004). At the final resolution of 1.9 Å, the *R* factor and $R_{\rm free}$ were 16.8% and 18.5%, respectively. Refinement statistics are shown in Table 1. The stereochemistry of the final model was checked using *MolProbity* (Davis *et al.*, 2007). The structure was deposited in the PDB under accession code 3iwt. The protein interfaces, surfaces and assemblies service *PISA* (Krissinel & Henrick, 2007) at the European Bioinformatics Institute (http://www.ebi.ac.uk/msd-srv/prot_int/pistart.html) was used to calculate the molecular properties described below.

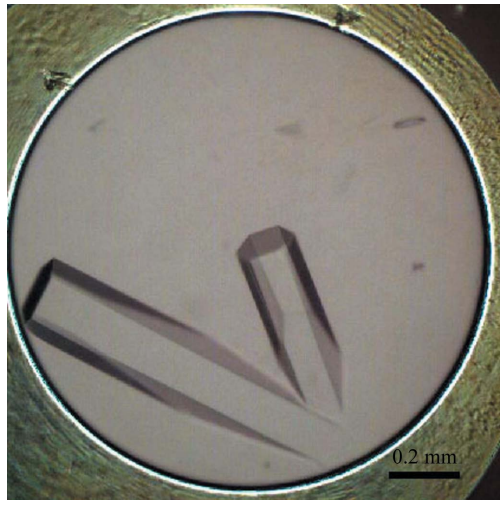


Figure 1
Crystals of *S. tokodaii* ST2315.

3. Results and discussion

The final model of ST2315 contains a trimer in the asymmetric unit with a total of 4039 protein atoms and 575 water molecules. Eight PEG molecules, two glycerol molecules and two hydrated Mg^{2+} atoms were also included in the model and were located on the surface of the protein molecule. Ramachandran plots calculated using *PROCHECK* (Laskowski *et al.*, 1993) showed that 96.1% of residues are in allowed regions, with 3.9% in additionally allowed regions of the structure. Each subunit of the trimer contains 166 residues, with the first 12 residues at the N-terminal ends missing because the electron density was too weak in these sections to be built correctly. The subunit forms a globular domain consisting of nine helices, including two 3_{10}-helices, and six β-strands (Fig. 2).

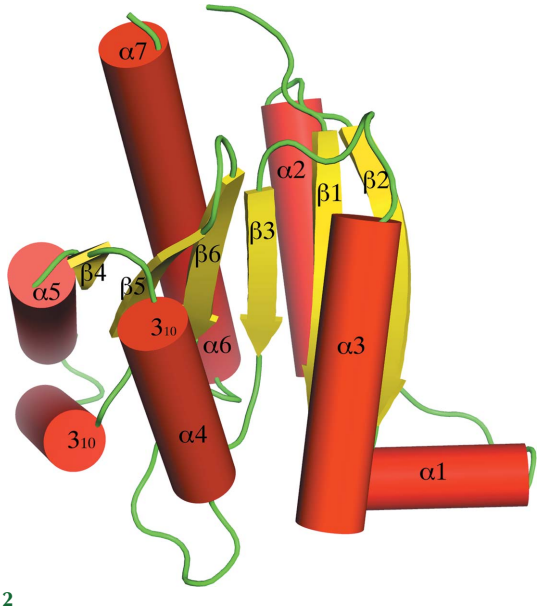


Figure 2
The secondary structure and overall fold of the ST2315 subunit. The core domain consists of six β-strands (yellow), two 3_{10}-helices and seven α-helices (red cylinders). The intersubunit interface is located on the left-hand side of the figure, in the region of β-strand 3, helix α5 and the 3_{10}-helix.

Table 2
Summary of potential hydrogen bonds and salt bridges.

(*a*) Between adjacent subunits of the ST2315 trimer.

		Distance (Å)
Glu112 OE1	Arg104 NH1	2.96
Glu112 OE2	Arg104 NH2	2.82
Asp116 OD2	Arg119 NE	2.84
Asp116 OD1	Arg119 NH2	2.88
Arg109 NH1	Ser94 O	2.95
Arg109 NH2	Tyr93 O	2.82
Glu112 N	Tyr93 OH	2.96
Tyr148 OH	Tyr133 OH	2.88
Glu110 O	Tyr93 OH	3.48
Leu164 O	Tyr133 OH	2.80

(*b*) Between adjacent ST2315 trimers.

		Distance (Å)
Lys66 NZ	Asp108 OD1	2.54
Leu106 O	Lys66 NZ	2.54
Thr74 OG1	Thr74 OG1	2.70

The trimer buries 17% of the solvent-accessible surface area of each subunit. Intermolecular contacts between adjacent subunits include two salt bridges, involving residues Glu112 and Asp116 of one subunit and Arg104 and Arg119 of the second subunit, and six potential hydrogen bonds between residues Arg109, Glu110, Glu112, Leu164 and Tyr148 of one subunit and Tyr93, Ser94 and Tyr133 of the second subunit (Table 2). The final model of ST2315 has an overall C^{α}-atom r.m.s.d. of 0.85 Å for 398 atoms of the trimer and 0.7 Å for 125 atoms of one subunit compared with the molecular-replacement search model from *B. cereus* (PDB code 1y5e). The only significant structural difference between these subunits lies in the region of helix α1 (residues 25–38) on the external arms of the ST2315 trimer (Fig. 3), which is a short loop in the *B. cereus* protein. Inspection of the crystallographic packing shows that the closest contact between the molecules involves the α1 helices of adjacent trimers.

Size-exclusion chromatography suggests that ST2315 exists as a hexamer in solution and as for *E. coli* MoaB (Sanishvili *et al.*, 2004; Bader *et al.*, 2004) this oligomeric form is likely to be the biologically significant unit. A hexamer (or dimer of trimers) of ST2315 (Fig. 3) is

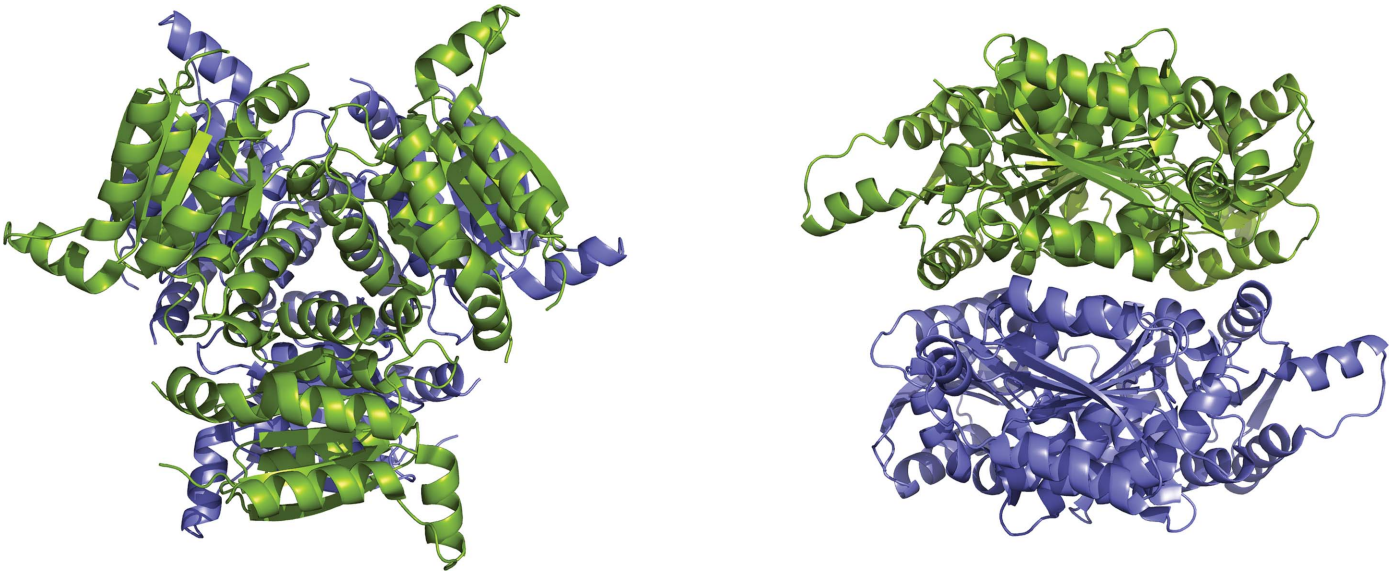


Figure 3
Two orientations of the possible ST2315 hexamer are shown parallel and perpendicular to the rotation axis of the crystallographic trimer. The hexamer is generated from the crystallographic trimer by the orthogonal transformation $-x$, y, $-z$. The molecular-graphics figures were all obtained using *PyMOL* (DeLano, 2008).

generated by the crystallographic symmetry. Formation of the hexamer buries approximately 10% of the solvent-accessible area of each subunit. The main interacting regions involve the α3 helices (residues Lys66–Ser78) from each subunit of one trimer and the symmetrically equivalent α3 helices of the adjacent trimer. Additional contacts are made between the α3 helices and the α4 and 3_{10}-helices (residues Thr102–Asp108) of the interacting trimers. These contacts include one salt bridge (Lys66–Asp108) and two potential hydrogen bonds (Thr74–Thr74 and Lys66–Leu106) (Table 2). The same secondary-structure elements are involved in forming the intertrimer interfaces of the MoaBs from *E. coli* and *B. cereus*. The interacting residues are generally not conserved, although *B. cereus* retains the Lys–Asp salt bridge between the interacting trimers.

Alignments made using the *DALI* server (Holm & Sander, 1994) show that the *E. coli* MoaB subunit (29% identity) has a *Z* score of 25.7 and an r.m.s.d. of 1.2 Å with respect to ST2315, while the molecular-replacement search model subunit from *B. cereus* has a *Z* score of 27.1. There is also a strong structural similarity between ST2315 and enzymes involved in molybdopterin synthesis, including *E. coli* MogA (*Z* score 21.1, r.m.s.d. 1.7 Å; Liu *et al.*, 2000), the G domain of the human neuroreceptor anchoring protein gephyrin (*Z* score 23.5, r.m.s.d. 1.7 Å), which has a dual role in glycine receptor clustering and in molybdoenzyme activity (Feng *et al.*, 1998; Sola *et al.*, 2001), and the CNX1 plant protein (*Z* score 23.5, r.m.s.d. 1.7 Å; Schwarz *et al.*, 2001). Unlike previously identified MoaB-type proteins, these other enzymes are trimeric in the solution state. In ST2315 the trimer may also be a stable oligomer, as implied by the number of hydrogen bonds and salt bridges and the amount of buried surface area per subunit.

We thank Mr Yoshihiro Agari and Dr Akio Ebihara for their help in sample preparation. This work was supported in part by the RIKEN Structural Genomics/Proteomics Initiative (RSGI), the National Project on Protein Structural and Functional Analyses, Ministry of Education, Culture, Sports, Science and Technology of Japan. This work was supported by the Synchrotron Radiation Department at the Science and Technology Facilities Council, Daresbury Laboratory UK, and X-ray data were collected on beamline 10.1 at the Synchrotron Radiation Source, which was supported by Biotechnology and Biological Sciences Research Council Grant BB/E001971 (to SSH and RWS)

References

Abola, A., Bernstein, F. C., Bryant, S. H., Koetzle, T. F. & Weng, J. (1987). *Crystallographic Databases – Information Content, Software Systems, Scientific Applications*, edited by F. H. Allen, G. Bergerhoff & R. Sievers, pp. 107–132. Bonn/Cambridge/Chester: Data Commission of the International Union of Crystallography.

Bader, G., Gomez-Ortiz, M., Haussmann, C., Bacher, A., Huber, R. & Fischer, M. (2004). *Acta Cryst.* D**60**, 1068–1075.

Bateman, A., Coin, L., Durbin, R., Finn, R. D., Hollich, V., Griffiths-Jones, S., Khanna, A., Marshall, M., Moxon, S., Sonnhammer, E. L., Studholme, D. J., Yeats, C. & Eddy, S. R. (2004). *Nucleic Acids Res.* **32**, D138–D141.

Davis, I. W., Leaver-Fay, A., Chen, V. B., Block, J. N., Kapral, G. J., Wang, X., Murray, L. W., Arendall, W. B. III, Snoeyink, J., Richardson, J. S. & Richardson, D. C. (2007). *Nucleic Acids Res.* **35**, W375–W383.

DeLano, W. L. (2008). *PyMOL Molecular Viewer.* DeLano Scientific, Palo Alto, California, USA. http://www.pymol.org.

Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.

Feng, G., Tintrup, H., Kirsch, J., Nichol, M. C., Kuhse, J., Betz, H. & Sanes, J. R. (1998). *Science*, **282**, 1321–1324.

Holm, L. & Sander, C. (1994). *Proteins*, **19**, 165–173.

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

Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.

Liu, M. T. W., Wuebbens, M. M., Rajagopalan, K. V. & Schindelin, H. (2000). *J. Biol. Chem.* **275**, 1814–1822.

Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

Perrakis, A., Morris, R. & Lamzin, V. S. (1999). *Nature Struct. Biol.* **6**, 458–463.

Sanishvili, R., Beasley, S., Skarina, T., Glesne, D., Joachimiak, A., Edwards, A. & Savchenko, A. (2004). *J. Biol. Chem.* **279**, 42139–42146.

Schwarz, G. & Mendel, R. R. (2006). *Annu. Rev. Plant Biol.* **57**, 623–647.

Schwarz, G., Schrader, N., Mendel, R. R., Hecht, H.-J. & Schindelin, H. (2001). *J. Mol. Biol.* **312**, 405–418.

Sola, M., Kneussel, M., Heck, I. S., Betz, H. & Weissenhorn, W. (2001). *J. Biol. Chem.* **276**, 25294–25301.

Sugahara, M. *et al.* (2008). *J. Struct. Funct. Genomics*, **9**, 21–28.

Vagin, A. & Teplyakov, A. (1997). *J. Appl. Cryst.* **30**, 1022–1025.

Yoshida, H., Yamada, M., Kuramitsu, S. & Kamitori, S. (2008). *Acta Cryst.* F**64**, 589–592.