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

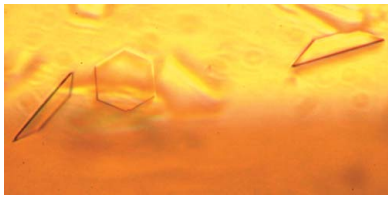
Preliminary structural investigations of the Eut-L shell protein of the ethanolamine ammonia-lyase metabolosome of *Escherichia coli*

Kiel Nikolakakis,[a] Akashi Ohtaki,[a] Keith Newton,[a] Arkadiusz Chworos[b] and Martin Sagermann[a,c]*

[a]Department of Chemistry and Biochemistry, University of California Santa Barbara, Santa Barbara, California 93106-9510, USA, [b]Department of Physics, University of California Santa Barbara, Santa Barbara, California 93106-9510, USA, and [c]Interdepartmental Program in BioMolecular Science and Engineering, University of California Santa Barbara, Santa Barbara, California 93106-9510, USA

Correspondence e-mail: sagermann@chem.ucsb.edu

Received 10 October 2008
Accepted 10 December 2008



The ethanolamine ammonia-lyase microcompartment is composed of five different shell proteins that have been proposed to assemble into symmetrically shaped polyhedral particles of varying sizes. Here, preliminary X-ray analysis of crystals of the bacterial microcompartment shell protein Eut-L from *Escherichia coli* is reported. Cloning, overexpression and purification resulted in highly pure protein that crystallized readily under many different conditions. In all cases the protein forms thin hexagonal plate-shaped crystals belonging to space group *P*3 that are of unusually high stability against different solvent conditions. The crystals diffracted to a resolution of 2.0 Å using synchrotron radiation but proved to be radiation-sensitive. Preparations of heavy-atom-derivatized crystals for use in determining the three-dimensional structure are under way.

1. Introduction

Bacterial microcompartments (BMCs) are supramolecular assemblies that harbor small-molecule metabolizing enzymes within the lumen of a polyprotein shell (Cannon *et al.*, 2001; Bobik, 2006). *Escherichia coli* and *Salmonella typhimurium*, for example, have evolved these specialized organelles for the degradation of small-molecular compounds such as ethanolamine and propanediol, respectively (Bradbeer, 1965; Faust *et al.*, 1990; Roof & Roth, 1988; Havemann & Bobik, 2003). The microcompartments enable these organisms to survive using these compounds as the only source of carbon, nitrogen and energy (Penrod & Roth, 2006).

The *Eut* operon of *E. coli* carries 17 genes that are intrinsic components of the ethanolamine ammonia-lyase BMC. Based on sequence homology, shell proteins (Kofoid *et al.*, 1999), catalytic lyase subunits, a transcription factor, two chaperones, a membrane transporter and catabolic enzymes as well as two proteins of unknown function have been identified.

One of the most fascinating aspects of these supramolecular assemblies is the formation of a membrane-like shell structure that separates the lumen of the compartment from the cytosol. This proteinaceous shell is thought to be composed of five different shell proteins: Eut-K, Eut-L, Eut-M, Eut-N and Eut-S. Recent studies suggest that the shell proteins themselves function as a semipermeable barrier that enables specific solutes to enter or leave the compartment (Penrod & Roth, 2006; Dou *et al.*, 2008).

Here, we present the cloning, overexpression, purification and preliminary crystallographic analyis of the Eut-L protein of the ethanolamine ammonia-lyase metabolosome of *E. coli.* Based on its sequence similarity to other BMC shell proteins, this protein has been predicted to be an intrinsic component of the microcompartment (Kofoid *et al.*, 1999).

2. Materials and methods

The Eut-L protein was cloned directly from the *E. coli* genome by selective PCR amplification. Three different versions of the protein were cloned as follows. (i) The primers Eut-L, 5′-CACCATGCCAGCTTTAGATTTGATTCGACCG-3′, and Eut-L_ANTI, 5′-TTACGCACGCTGGATTGGATTACG-3′, were used to clone the Eut-L-pET151 His-tagged version of the protein, (ii) Eut-L_NHIS was

cloned using the primers Eut-L_NHIS, 5′-CACCATGCATCATCACCATCACCATCCAGCTTTAGATTTGATTCGACCG-3′, and Eut-L_ANTI and (iii) Eut-L_CHIS was cloned using the primers Eut-L and Eut-L_CHIS, 5′-TTAATGGTGATGGTGATGATGCGCACGCTGGATTGGATTACG-3′. The cloning of the Eut-L sequence into the pET151 vector (Invitrogen) appends a cleavable 33-amino-acid tag sequence to its amino-terminus. The sequences for the short His_6-tagged versions of the protein, Eut-L_NHIS and Eut-L_CHIS, were cloned into pET101 vector (Invitrogen). Ligations and transformations were performed according to the manufacturer's specifications. A 50 ml overnight culture of the corresponding mutant was used to inoculate 2 × 1 l LB medium in baffled flasks. Protein expression was induced with 100 µ*M* isopropyl *β*-D-1-thiogalactopyranoside (ITPG) and 200 µ*M* (+)-arabinose (Fisher) for 3 h at 291 K (at 100 µ*M* ampicillin). Subsequently, cells were pelleted at 5000 rev min^{-1} using a Beckman JA10 rotor. 50 ml of a suspension of the harvested cells was then subjected to lysis *via* a French press at 6.9 MPa in the presence of 1 m*M* phenylmethanesulfonyl fluoride and a tablet of Boehringer protease-inhibitor cocktail. The lysate was spun for 20 min at 15 000 rev min^{-1} (JA20 rotor) and the protein was isolated *via* nickel-affinity chromatography according to the manufacturer's protocol (Qiagen). The resulting purity of the eluted protein and its oligomerization state were confirmed by SDS–PAGE, native PAGE and size-exclusion chromatography using a Sephadex 200 column (GE Healthcare). Dynamic light scattering was carried out in 50 µl quartz cuvettes on a Wyatt MiniDAWN spectrometer (Wyatt Technologies).

2.1. AFM imaging

Samples from crystallization drops were diluted 10–20-fold with deionized water; 10 µl samples of the resulting solutions were subsequently deposited onto a freshly cleaved V1-grade mica surface for 30 min under humidity-controlled conditions. Subsequently, the mica surface was washed with 4 × 500 µl HPLC pure water (Chromasolv from Sigma–Aldrich) and imaged under dried gaseous nitrogen or alternatively under water. AFM images were recorded in amplitude-modulated tapping mode using a MultiMode microscope equipped with an E scanner controlled by a Nanoscope IIIA (Veeco, Santa Barbara, California, USA). For experiments performed under water, silicon nitride probes (MSCT-AUHM) with resonance frequency ∼8 kHz and spring constant $k \simeq 0.03$ N m^{-1} (Veeco Probes, Santa Barbara, California, USA) were used. Images in air were acquired using silicon probes (NSC12 without Al) with resonance frequency ∼160–200 kHz and spring constant $k \simeq 4.5$–6.5 N m^{-1} (MicroMasch, Wilsonville, Oregon, USA). All images were processed and analyzed using NanoScope software (Veeco, Santa Barbara, California, USA) and leveled by a first-order plane fit in order to correct the sample tilt.

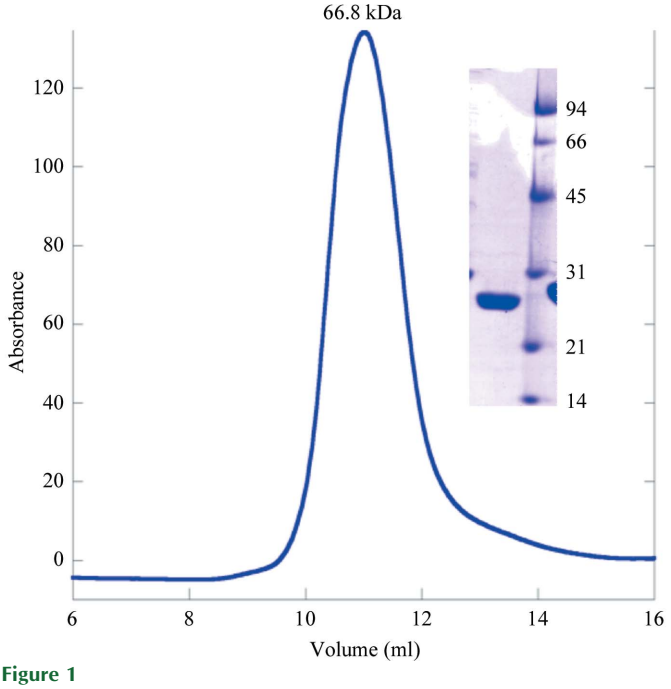


Figure 1
Superdex 200 size-exclusion chromatogram of Eut-L. The molecular weight was extrapolated from runs with commercially available molecular-weight standards. The chromatogram of Eut-L showed one distinct molecular-weight species of 68.81 kDa corresponding to a trimer. The inset picture shows an SDS–PAGE gel of the elution peak from the Sephadex 200 column (GE Healthcare) and shows the expected molecular weight of the monomer. The molecular weights of the standards are labelled in kDa.

2.2. Derivatization of the crystals

Selenomethione-derivatized proteins were obtained by growing cloned bacteria in selenomethionine-containing M9 minimal media (Molecular Dimensions Ltd). Protein overexpression and purification were performed in the same manner as described above. Incorporation of the artificial amino acid was confirmed by mass spectrometry using an ion-spray LC-QTOF MS/MS mass spectrometer as well as by X-ray fluorescence scans. Owing to the pronounced radiation-sensitivity of the crystals, a variety of heavy-metal compounds are currently being screened in order to enable efficient phasing. For crystallization trials, the protein sample was dialyzed against Standard Buffer (50 m*M* HEPES pH 7.0) and concentrated to a final concentration of about 1 mg ml^{-1}. Attempts to collect data using in-house copper X-ray radiation resulted in diffraction to 8 Å resolution at best. X-ray diffraction studies were therefore exclusively carried out using synchrotron radiation (Stanford Synchrotron Radiation Laboratory beamlines 9-1 and 11-1). Data processing was performed with the programs *MOSFLM* (Leslie, 1992; Collaborative Computational Project, Number 4, 1994) or *XDS* (Kabsch, 1993).

3. Results

Overexpression and purification of the Eut-L proteins resulted in ∼10 mg of protein per litre of bacterial culture. Upon dialysis, however, the protein readily precipitated, resulting in a final concentration of 1 mg ml^{-1} soluble protein (Bradford concentration assay). The version of Eut-L derived from the pET151 vector

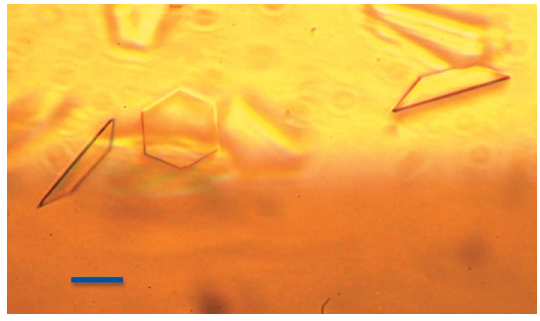


Figure 2
Crystals of Eut-L_CHIS obtained by hanging-drop vapor diffusion. The blue scale bar corresponds to 0.1 mm. The image was taken with a Leica SM-Lux microscope equipped with a Nikon D200 digital camera.

contained a 4.5 kDa tag that include a six-histidine-residue tag for metal-affinity chromatography and a TEV cleavage site for its removal. However, digestion of the protein under systematically varied conditions failed to remove the tag. Further structural studies were therefore carried out with recloned versions of the protein that contained only a short noncleavable His_6 tag at the N-terminus or at the C-terminus of the protein (referred to as Eut-L_NHIS or Eut-L_CHIS, respectively). Size-exclusion chromatography suggested that freshly prepared protein oligomerized readily into trimers in the presence or absence of 5 m*M* β-mercaptoethanol (Fig. 1). No monomeric or other higher order oligomerization states were observed in solution. The presence of a unique oligomer population was also confirmed by dynamic light-scattering experiments (data not shown).

Prior to crystallization trials, protein solutions were concentrated to about 1 mg ml^{-1} using Centriprep (Millipore). All crystallizations were carried out by hanging-drop vapor diffusion using commercially available screens from Hampton Research and Jena BioScience. Approximately 30% of all trial conditions yielded microcrystals over a wide range of pH values (crystals were observed within the pH range 4.6–8.5) and of different types of precipitants (salts, polyethylene glycols and alcohols). The best crystals of Eut-L_NHIS were obtained in 3.3 *M* ammonium acetate, 5% polyethylene glycol 400 (PEG 400) and 50 m*M* Tris buffer pH 7.5. These crystals grew as hexagonal plates to dimensions of approximately 0.2 × 0.2 × 0.02 mm. Preliminary X-ray diffraction analysis of these crystals revealed extremely anisotropic diffraction properties. When the X-ray beam was incident on the face of the crystal, good diffraction with sharp reflections was observed to 2.2 Å resolution. Perpendicular to this orientation, however, the crystals diffracted poorly to only 8 Å resolution with extremely streaky reflections, indicating substantial disorder along the *c* dimension of the crystal. However, crystals of Eut-L_CHIS proved to be of much better quality. In contrast to Eut-L_NHIS, Eut-L_CHIS crystals grew as single hexagonal plates with sharp edges (Fig. 2). Crystals grew in 2 *M* NaCl, 100 m*M* phosphate, 100 m*M* MES buffer pH 6.5 and 4% PEG 400. Diffraction patterns showed good diffraction along all dimensions to about 2.0 Å resolution and indexing of the data confirmed a trigonal space group. Data statistics for the best single-crystal data collection are given in Table 1.

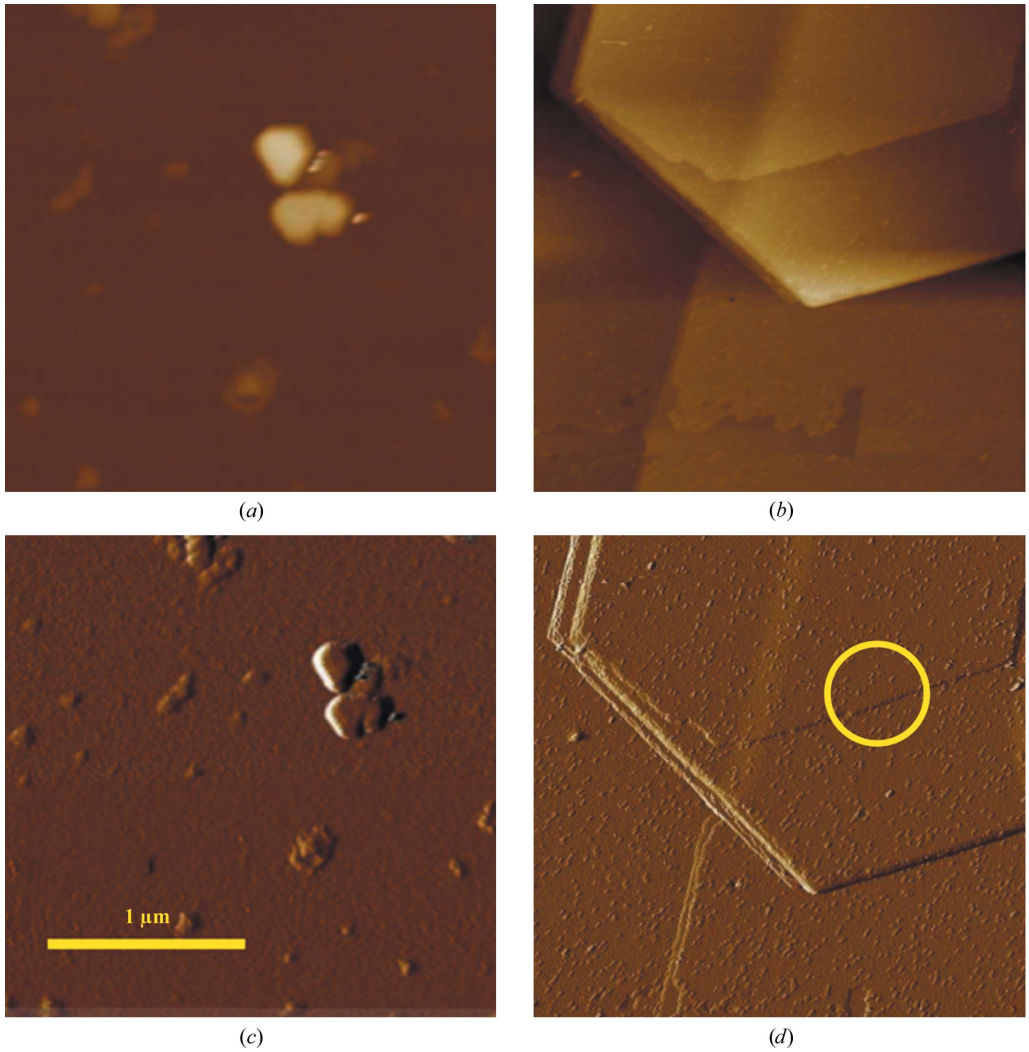


Figure 3
AFM images of Eut-L_CHIS crystals 2 h (*a*, *c*) or several days after setup (*b*, *d*). The images in (*a*) and (*c*) (imaged in height or amplitude mode, respectively) were scanned on dried crystals under nitrogen gas, whereas the images in (*b*) and (*d*) were obtained in water. In both cases, crystals were first diluted with deionized water before adsorption onto the mica surface. Individual lattice steps of the crystals (circle) corresponded to a thickness of 7.9 ± 0.5 nm and were likely to reflect the unit-cell parameters of the crystal along *c*. One other less frequent step size was identified corresponding to 4.0 ± 0.5 nm, which may be explained by the presence of two molecules in the asymmetric unit. The protein content of similarly treated crystals was confirmed by SDS–PAGE and mass spectrometry (data not shown).

One of the most surprising observations was the speed of crystallization and the remarkable stability of the crystals. Microcrystals appeared after 30–90 min. At high protein concentrations (>2 mg ml^{-1}), crystallization drops typically produced showers of microcrystals. Even through a decrease in the protein concentration produced fewer and larger crystals, the appearance of crystal showers was hardly controllable. Investigations of freshly prepared crystallization solutions by AFM also exhibited many small triangular crystals, confirming the rapid crystallization of the protein (Fig. 3*a*). Larger sizes were observed for older crystals (24 h) that closely reflect the shapes of the macroscopic crystals. The edges of the crystals exhibit steps similar in length to the *c* edge of the unit cell or half of this length (Fig. 3*b*). No crystals were observed in the protein stock solution.

Crystals also proved to be unusually stable with respect to changes in pH and ionic strength. Typically, protein crystals are highly sensitive to changes that perturb the solution equilibrium in which they were grown. Despite the fact that the crystals of Eut-L_CHIS were grown at high ionic strength, they appeared to remain stable when transferred into deionized water or buffer solutions of different pH or ionic strength without any visible deterioration in shape or morphology. In fact, crystals washed extensively with deionized water could still be imaged by AFM (Fig. 3), suggesting strong lattice interactions.

4. Discussion

The observation that Eut-L crystals grew rapidly into hexagonal plates under many different conditions may be indicative of the natural state and function of this protein within the structure of the shell of the ammonia-lyase compartment. Recent studies have shown that the carboxysomal shell proteins CcmK2 and CCmk4 also assemble into hexagonal configurations, which have been proposed to be an intrinsic feature of the carboxysome shell (Kerfeld *et al.*, 2005). This arrangement of the protein molecules resulted in tightly packed hexagonal tiles that form planar sheets within the crystal. The only openings within the sheet were pores in the centers of each hexamer. It was suggested that the conserved charged residues in the center of these pores allow the controlled exchange of solutes in and out of the interior of the compartment.

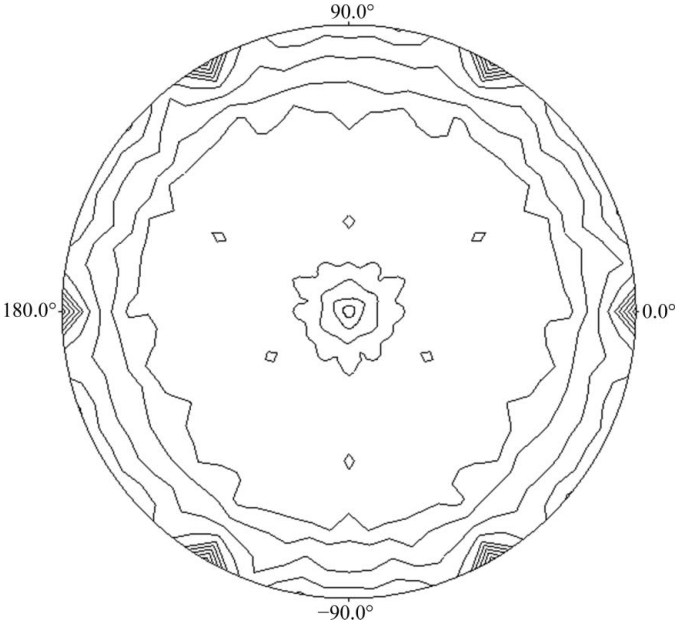


Figure 4
Plot of the 180° section of the self-rotation function calculated with *POLARRFN* (Collaborative Computational Project, Number 4, 1994). All data between 15 and 3.5 Å resolution were included in the calculation. The twofold axes are located in the *xy* plane and are currently treated as noncrystallographic symmetry (integration of the data in the corresponding higher symmetries resulted in unreasonably high symmetry *R* values).

Table 1
Data-collection statistics for a single Eut-L_CHIS crystal.

Values in parentheses are for the highest resolution shell.

X-ray source	SSRL BL9-1
Wavelength (Å)	0.979462
Space group	*P*3
Unit-cell parameters (Å, °)	$a = b = 67.38$, $c = 79.66$, $\alpha = \beta = 90.0$, $\gamma = 120.0$
No. of molecules per ASU	2
Resolution limits (Å)	19.95–1.95 (2.3–1.95)
Total No. of reflections	80659 (27135)
No. of unique reflections	25262 (8565)
$I/\sigma(I)$	10.0 (4.46)
R factor†	0.08 (0.278)
Completeness (%)	85.75 (74.3)
Overall mosaicity (°)	0.125
Wilson *B* factor (Å^2)	27.3

† As defined in *XDS*.

In contrast to previously determined carboxysome-shell proteins, no sixfold symmetry was observed in the Eut-L_CHIS crystals. Self-rotation functions and Patterson map calculations confirmed the trigonal symmetry of the collected data. Following integration of the data with *P*1 symmetry, self-rotation functions clearly revealed a threefold crystallographic axis perpendicular to the face of the crystal plate. Twofold axes perpendicular to the threefold axis were also observed (Fig. 4). No systematic absences of centric reflections were identified. As the twofold symmetries are not strong in all data sets collected, they were treated as noncrystallographic. The estimated Matthews coefficient (Matthews, 1968) for the crystal was $V_M = 2.21$ Å^3 Da^{-1}, indicating the possibility of two molecules in the asymmetric unit with a solvent content of 44.3%.

The amino-acid sequence of Eut-L did not exhibit any similarities to those of other proteins in the Protein Data Bank (PDB). *BLAST* sequence searches of its 219-amino-acid sequence revealed substantial sequence identity to proteins of other genera of prokaryotic organisms, such as *Salmonella*, *Shigella*, *Listeria* and *Clostridium*, which are also predicted to have ethanolamine ammonia-lyase-like microcompartments. The most pronounced homologies were identified to proteins of the propanediol-utilization metabolosome (emb|CAJ68791.1, with 54% sequence identity, and PduB, with 29% identity between residues 16 and 189). Secondary-structure predictions with the programs *PhD* (Rost *et al.*, 2004) and *SAM* (Karplus *et al.*, 1998) suggested a mixed α/β-protein.

As the Eut-L protein is crucially important to the function and the assembly of the microcompartment of ethanolamine ammonia-lyase, the determination of its structure is likely to provide important clues to the architecture and function of this supramolecular assembly.

We are extremely grateful to Richard C. Chapleau for help with size-exclusion chromatography and for help with mass spectrometry, Dr Andrea Tao for help with dynamic light scattering and Lisa Dunn and Drs Mike Soltis, Tzanko Duokov and Pete Dunten for the generously awarded beamtime as well for their expert help. AFM imaging was performed in the laboratory of Professor Helen Hansma at UCSB.

References

Bobik, T. A. (2006). *Appl. Microbiol. Biotechnol.* **70**, 517–525.
Bradbeer, C. (1965). *J. Biol. Chem.* **240**, 4669–4674.
Cannon, G. C., Bradburne, C. E., Aldrich, H. C., Baker, S. H., Heinhorst, S. & Shively, J. M. (2001). *Appl. Environ. Microbiol.* **67**, 5351–5361.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Dou, Z., Heinhorst, S., Williams, E. B., Murin, C. D., Shively, J. M. & Cannon, G. C. (2008). *J. Biol. Chem.* **283**, 10377–10384.
Faust, L. R., Connor, J. A., Roof, D. M., Hoch, J. A. & Babior, B. M. (1990). *J. Biol. Chem.* **265**, 12462–12466.
Havemann, G. D. & Bobik, T. A. (2003). *J. Bacteriol.* **185**, 5086–5095.
Kabsch, W. (1993). *J. Appl. Cryst.* **26**, 795–800.
Karplus, K., Barrett, C. & Hughey, R. (1998). *Bioinformatics*, **14**, 846–856.
Kerfeld, C. A., Sawaya, M. R., Tanaka, S., Nguyen, C. V., Phillips, M., Beeby, M. & Yeates, T. O. (2005). *Science*, **309**, 936–938.
Kofoid, E., Rappleye, C., Stojiljkovic, I. & Roth, J. (1999). *J. Bacteriol.* **181**, 5317–5329.
Leslie, A. G. W. (1992). *Jnt CCP4/ESF–EACBM Newsl. Protein Crystallogr.* **26**.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Penrod, J. T. & Roth, J. R. (2006). *J. Bacteriol.* **188**, 2865–2874.
Roof, D. M. & Roth, J. R. (1988). *J. Bacteriol.* **170**, 3855–3863.
Rost, B., Yachdav, G. & Liu, J. (2004). *Nucleic Acids Res.* **32**, W321–W326.