Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Affinity tags can reduce merohedral twinning of membrane protein crystals

Anna Backmark,[a] Maria Nyblom,[a] Susanna Törnroth-Horsefield,[b] Urszula Kosinska-Eriksson,[b] Kristina Nordén,[c] Maria Fellert,[c] Per Kjellbom,[c] Urban Johanson,[c] Kristina Hedfalk,[b] Karin Lindkvist-Petersson,[d] Richard Neutze[b] and Rob Horsefield[b]*

[a]Department of Chemical and Biological Engineering, Chalmers University of Technology, PO Box 462, SE-405 30 Gothenburg, Sweden, [b]Department of Chemistry, University of Gothenburg, PO Box 462, SE-405 30 Gothenburg, Sweden, [c]Center for Molecular Protein Science, Department of Biochemistry, Lund University, PO Box 124, SE-221 00 Lund, Sweden, and [d]Department of Cell and Molecular Biology, University of Gothenburg, Box 462, SE-405 30 Gothenburg, Sweden

Correspondence e-mail: rob@chem.gu.se

Received 14 July 2008
Accepted 15 September 2008

This work presents a comparison of the crystal packing of three eukaryotic membrane proteins: human aquaporin 1, human aquaporin 5 and a spinach plasma membrane aquaporin. All were purified from expression constructs both with and without affinity tags. With the exception of tagged aquaporin 1, all constructs yielded crystals. Two significant effects of the affinity tags were observed: crystals containing a tag typically diffracted to lower resolution than those from constructs encoding the protein sequence alone and constructs without a tag frequently produced crystals that suffered from merohedral twinning. Twinning is a challenging crystallographic problem that can seriously hinder solution of the structure. Thus, for integral membrane proteins, the addition of an affinity tag may help to disrupt the approximate symmetry of the protein and thereby reduce or avoid merohedral twinning.

1. Introduction

Membrane protein crystallization remains a challenging task (Iwata, 2003), with approximately 150 unique membrane protein structures reported to date. Numerous factors influence success in recovering well diffracting membrane protein crystals, including the overproduction yield and protein purity. Recombinant overproduction of eukaryotic membrane proteins often results in low yields and thus affinity tags are frequently used to facilitate purification. One common affinity tag is the addition of six or more consecutive histidine residues (a His tag) to the protein sequence (Derewenda, 2004). His tags allow purification by metal-affinity chromatography (normally Ni^{2+}). Common variations of the basic tag include the addition of a linker or a cleavage site so that the His tag can be removed prior to crystallization. A recent survey of entries in the Protein Data Bank (PDB; Berman *et al.*, 2000) reported that the presence of a His tag only has a minor influence on the protein structure (Carson *et al.*, 2007). Since membrane protein structures constitute approximately 1% of the entries in the PDB, it is possible that additional challenges specific to membrane proteins may have been overlooked in this survey.

Membrane proteins crystallize as type I or type II crystals (Ostermeier & Michel, 1997). Type I crystals are essentially stacked two-dimensional crystals with repeating protein–lipid layers held together by hydrophobic contacts between molecules within the layers and soluble interactions between the layers. Type II crystals, which are more common, result from detergent-solubilized protein and the interactions form between soluble domains of the protein that extend out of the micelle. In either case, membrane proteins frequently crystallize as stacked assemblies with repeating hydrophobic and hydrophilic layers. Layered crystals, and particularly those with weak or nonspecific interactions between the layers, are susceptible to crystal-growth defects such as merohedral twinning (Parsons, 2003). Thus, for crystallization of integral membrane proteins, especially those that lack any large soluble domain, there is a significant risk of merohedral twinning. Published examples of integral membrane proteins that suffer from twinning include bacteriorhodopsin (Luecke *et al.*, 1998; Faham *et al.*, 2004), cytochrome b_6f complex (Kurisu *et al.*, 2003), the ammonium transporter Amt-1 (Andrade *et al.*, 2005),

Table 1
Summary of crystallographic data for selected aquaporins.

Aquaporin	Source	PDB code	His tag	Resolution (Å)	Space group	Twin fraction α† (%)	Average $\langle I^2\rangle/\langle I\rangle^2$‡	References
AQP0§	Bovine	1ymg		2.2	$I422$	—		Harries *et al.* (2004)
AQP1	Bovine	1j4n		2.2	$P42_12$	—		Sui *et al.* (2001)
AQP1	Human			3.1	$I4$	50	1.56	Unpublished work
AQP5	Human		C-terminal	3.1	$P312$	17¶	2.27	Unpublished work
AQP5§	Human	3d9s		2.0	$P2_12_12_1$††	46	1.65	Horsefield *et al.* (2008)
PIP2;1, pH 8.0§	Spinach	1z98		2.1	$I4$	—		Törnroth-Horsefield *et al.* (2006)
PIP2;1, pH 5.6§	Spinach	2b5f	C-terminal	3.9	$P2_12_12$	—		Törnroth-Horsefield *et al.* (2006)
PIP2;1 S115E	Spinach		N-terminal	2.2	$I4$	—		Unpublished work
PIP2;1, pH 6.0	Spinach			2.3	$P4$	45	1.36	Unpublished work
AqpZ	*E. coli*	1rc2	Cleaved	2.5	$P4$	—		Savage *et al.* (2003)
AqpZ§	*E. coli*	2abm	Cleaved	3.2	$P4_122$	—		Jiang *et al.* (2006)
AqpM§	*M. marburgensis*	2f2b	Cleaved	1.68	$I4$	—		Lee *et al.* (2005)
GlpF	*E. coli*	1fx8	Cleaved	2.2	$I422$	—		Fu *et al.* (2000)

† Estimated using the Merohedral Crystal Twinning Server or the program *SHELXL*; see §2.4 for details. ‡ Average values of $\langle I^2\rangle/\langle I\rangle^2$ were obtained from a 'Perfect Twinning Test' using the Merohedral Crystal Twinning Server (see §2.4 for details), where 1.5 corresponds to perfectly twinned and 2.0 to untwinned. § Structure factors available for download from the PDB. ¶ 'Partial twinning test'; 2 along *c*, *a*, *b*; $\langle H\rangle = 0.312670$, $\langle H^2\rangle = 0.147459$; twin fraction = 0.177387 ± 0.009943. †† Unit-cell parameters were $a = 90.48$, $b = 90.64$, $c = 184.39$ Å. The diffraction data were from a crystal with near-perfect pseudo-merohedral twinning with twin operator $(h, k, l) \rightarrow (k, h, -l)$ (010 100 00−1), swapping the *a* and *b* axes, and an estimated twinning fraction of 0.46.

archaerhodopsin (Yoshimura & Kouyama, 2008) and human aquaporin 5 (Horsefield *et al.*, 2008). Twinning is the consequence of two or more domains within a single crystal having different orientations. If the reciprocal lattices (*i.e.* diffraction patterns) of the domains overlay perfectly, the twinning is merohedral. Each measured intensity is then the sum of the intensities from the individual domains and this is usually apparent when analyzing the data (Dauter, 2003), for example the cumulative intensity distributions (Rees, 1980). The twin fraction (α) describes the relative sizes of the domains in the crystal. If α is small, *e.g.* <20% (partial twinning), twinning can be easily accounted for when solving and refining the structure (detwinning). However, twinning is associated with complications such as pseudo-symmetry which make it difficult to determine the real space group of the crystal (Zwart *et al.*, 2008). Of greater concern, for severely twinned data, *e.g.* $\alpha = 50\%$ (perfect twinning), molecular replacement can sometimes be successful, but in other cases it may be impossible to overcome the twinning and the structure cannot be solved. As a result, the number of diffracting membrane protein crystals suffering from twinning is probably under-reported in the literature. While in fortuitous examples crystals without twinning may be found (Belrhali *et al.*, 1999), a rational strategy to overcome twinning is attractive.

In this work, we discuss the influence of a His tag on the crystallization of three eukaryotic membrane proteins: human aquaporin 1 (HsAQP1), human aquaporin 5 (HsAQP5) and the spinach plasma membrane aquaporin (SoPIP2;1). Aquaporins (AQPs) are integral membrane-channel proteins that facilitate the flux of water across biological membranes (King *et al.*, 2004). A number of X-ray structures of AQPs are already available (Table 1). Most AQPs have a compact and rather symmetric overall structure with short N- and C-termini and no large soluble domains. As a result, AQPs readily crystallize in stacked-layer arrangements with the protein molecules typically oriented perfectly orthogonal to the plane of the membrane (Fig. 1 and Fig. S1[1]). Here, we present a comparison of AQP structures solved with and without an affinity tag and the effects of the tag on crystal packing and twinning are discussed. Even though only aquaporins, membrane proteins of very similar structure, have been considered here, we conclude that the addition of a tag potentially provides a useful strategy to counter merohedral twinning of membrane protein crystals.

2. Methods

2.1. Cloning, overproduction and membrane preparation

The generation of the constructs for nontagged and C-terminally His-tagged HsAQP1 (Nyblom *et al.*, 2007) and SoPIP2;1 (Törnroth-Horsefield *et al.*, 2006) as well as for nontagged HsAQP5 (Horsefield *et al.*, 2008) has previously been described. Constructs of HsAQP5 with a C-terminal His tag and SoPIP2;1 mutant S115E with an N-terminal His tag were generated using a similar approach. Constructs without tags comprised the basic protein sequence without further additions, with the exception of HsAQP5, where an extra serine was incorporated after the initiation methionine to provide an optimal translational start site. Constructs with a C-terminal His tag included a Myc epitope and six consecutive histidines. The N-terminal tag of SoPIP2;1 contained six histidines and a linker (Myc and His-tag sequences are shown in bold): HsAQP5–AAASFL**EQKLISEEDL**NSAVD**HHHHHH**; SoPIP2;1–L**EQKLISEEDL**NSAVD**HHHHHH**; MS**HHHHHH**SSGLVPRGSH–SoPIP2;1(S115E).

All constructs were cloned into the expression vector pPICZB (Invitrogen) and overproduced in the methylotrophic yeast *Pichia pastoris*. The methods for overproduction, cell breakage and membrane preparation have been described previously (Törnroth-Horsefield *et al.*, 2006; Nyblom *et al.*, 2007; Horsefield *et al.*, 2008). In brief, protein overproduction in *P. pastoris*, wild-type strain X-33 (Invitrogen), was induced by methanol and all constructs gave an exceptional yield of membrane protein. Crude membranes were recovered by ultracentrifugation and washed with urea buffer and NaOH to remove peripheral proteins. *N*-Octyl-β-D-glucopyranoside (OG; Anatrace) was used for solubilization of HsAQP1 and SoPIP2;1, while *n*-nonyl-β-D-glucopyranoside (NG; Anatrace) was used for solubilization of HsAQP5.

2.2. Purification and crystallization

The procedures for nontagged and C-terminally tagged HsAQP1 (Nyblom *et al.*, 2007) and SoPIP2;1 (Törnroth-Horsefield *et al.*, 2006) as well as for nontagged HsAQP5 (Horsefield *et al.*, 2008) have been described previously. HsAQP5 with a His tag was purified according to Horsefield *et al.* (2008), but Ni^{2+}-affinity chromatography replaced

[1] Supplementary material has been deposited in the IUCr electronic archive (Reference: YT5010). Services for accessing this material are described at the back of the journal.

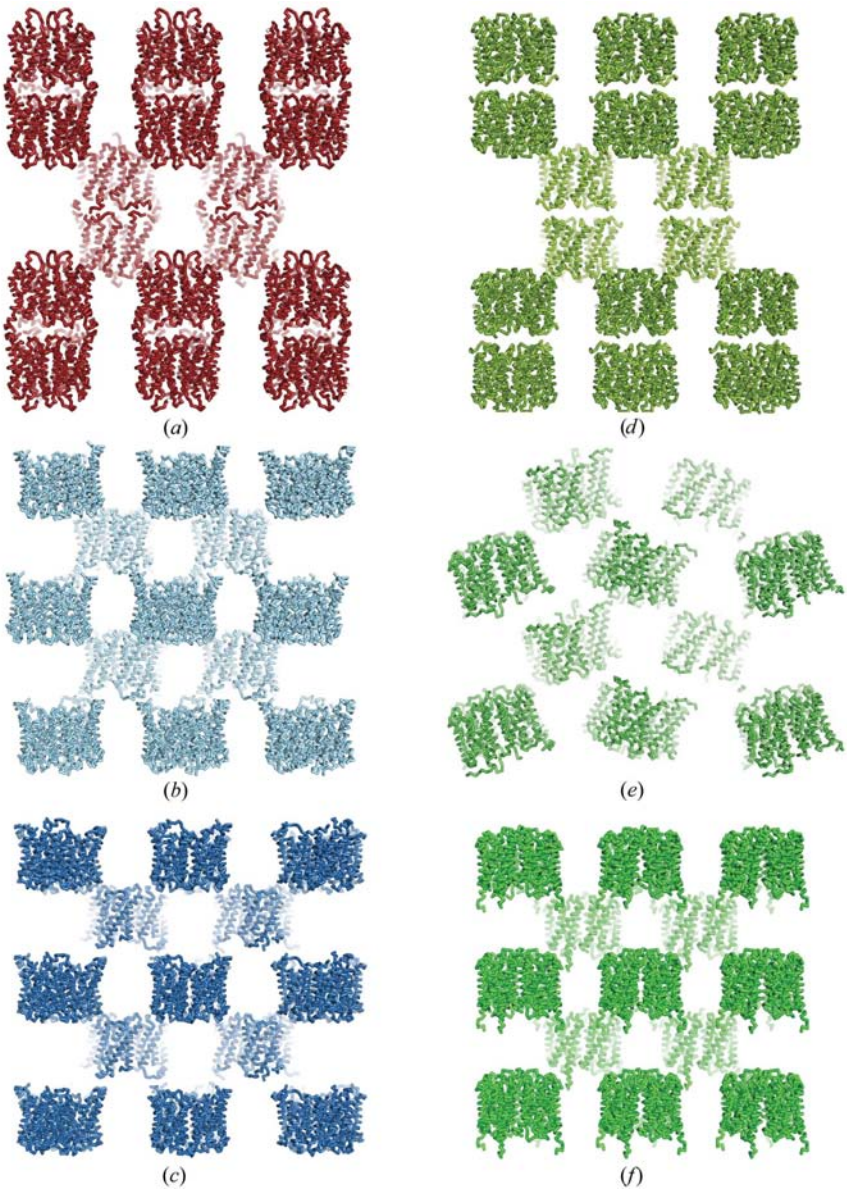

Figure 1
The influence of affinity tags on the crystallization of eukaryotic aquaporins. Crystal packing of (*a*) nontagged HsAQP1, (*b*) nontagged HsAQP5, (*c*) C-terminally tagged HsAQP5, (*d*) nontagged SoPIP2;1, (*e*) C-terminally tagged SoPIP2;1 and (*f*) N-terminally tagged S115E mutant SoPIP2;1.

the cation-exchange step. Crystals of tagged HsAQP5 were obtained under identical conditions to those used for HsAQP5 without an affinity tag (Horsefield *et al.*, 2008). Crystals of SoPIP2;1 S115E with an N-terminal His tag were obtained using the same purification and crystallization procedure as used for nontagged SoPIP2;1 (Törnroth-Horsefield *et al.*, 2006). Twinned crystals of the nontagged SoPIP2;1 construct were obtained with the same protocol using MES buffer pH 6.0 (Table 1).

2.3. Data processing, molecular replacement and refinement

Details of these methods have been provided elsewhere (Törnroth-Horsefield *et al.*, 2006; Horsefield *et al.*, 2008). In brief, diffraction data were processed and scaled with *MOSFLM* (Leslie, 1992) and *SCALA* (Kabsch, 1988). Molecular-replacement calculations were made using the programs *MOLREP* (Vagin & Teplyakov, 2000) and *AMoRe* (Navaza, 1994) from the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994). Model building in *O* (Jones *et al.*, 1991) and *Coot* (Emsley & Cowtan, 2004) was combined with refinement in *REFMAC* (Murshudov *et al.*, 1997), *CNS* (Brünger *et al.*, 1998) and *SHELXL* (Sheldrick, 2008).

2.4. Twin-fraction determination

Detection of twinning and estimation of the twin fraction (α) took place using the Merohedral Crystal Twinning Server (http://nihserver.mbi.ucla.edu/Twinning; Yeates, 1997) and Merohedral Twin Detector: Padilla–Yeates Algorithm (http://nihserver.mbi.ucla.edu/pystats; Padilla & Yeates, 2003). In the case of nontagged HsAQP5, the twin fraction was estimated and refined using *SHELXL* (Sheldrick, 2008).

3. Results and discussion

The three-dimensional crystal packing for the constructs of HsAQP1, HsAQP5 and SoPIP2;1 are shown in Fig. 1. HsAQP1 could only be crystallized in the absence of the His tag (Fig. 1*a*). A short summary of the crystallographic data is provided in Table 1, which also contains similar information for several X-ray structures of AQPs available in the PDB. Only three structures, HsAQP5 (Horsefield *et al.*, 2008), the open conformation of SoPIP2;1 (Törnroth-Horsefield *et al.*, 2006) and the S115E mutant of SoPIP2;1, were crystallized in the presence of a His tag. Although AQPZ (Savage *et al.*, 2003; Jiang *et al.*, 2006), AQPM (Lee *et al.*, 2005) and GlpF (Fu *et al.*, 2000) were recombinantly overproduced and purified using a His tag, the tag was removed by treatment with proteases prior to crystallization. HsAQP5 and SoPIP2;1 are the only examples where structures are available in both the presence and the absence of a tag. In the case of the SoPIP2;1 mutant S115E, the addition of an N-terminal His tag had little influence on the resolution of the structure (Table 1), but did alter the crystal packing (Fig. 1*f*). However, a C-terminal extension of wild-type SoPIP2;1 produced a significant alteration to the crystal packing (Fig. 1*e*) and the resolution was reduced from 2.1 to 3.9 Å (Table 1). Likewise, for HsAQP5 a C-terminal His tag changed the crystal packing (Fig. 1*c*) and caused a significant reduction in diffraction resolution, which declined from 2.0 to 3.1 Å (Table 1).

Another observation highlighted in Table 1 is that constructs of HsAQP1, HsAQP5 and SoPIP2;1 without a His tag all yielded crystals that suffered from merohedral twinning. In the case of HsAQP5 partial twinning (17%) was also observed for the tagged crystal but did not hinder solution and refinement of the structure. However, in the absence of a tag twinning was severe (>45%) for all three AQPs. For crystals of HsAQP1 and crystals of SoPIP2;1 grown at pH 6.0, both lacking the tag, twinning prevented the structures from being

solved. Although nontagged HsAQP5 was also severely twinned (46%), the structure could be solved and refined using *SHELXL* (Sheldrick, 2008), implementing twinning-compatible composite OMIT maps during model building (Horsefield *et al.*, 2008). Some of the other AQP structures in Table 1 had structure factors that were available for download. Interestingly, of those that did (see Table 1) and that were crystallized without a tag, some showed slight twinning according to the Merohedral Crystal Twinning Server (Yeates, 1997) and Merohedral Twin Detector: Padilla–Yeates Algorithm (Padilla & Yeates, 2003; data not shown; see §2.4). However, no twinning issues were reported and twinning does not appear to have influenced the solution and refinement of these structures (Fig. S1).

From these observations, two correlations are apparent: the presence of a tag limits the final diffraction resolution of the crystals and the presence of a tag can prevent or reduce the severity of merohedral twinning. The observed reduction in resolution can be explained by increased flexibility and heterogeneity of the protein sample containing the C-terminal tag, steric effects of the tag influencing crystal packing and (in some cases) a greater solvent content of the crystal.

We hypothesize that the positive influence of an affinity tag on twinning derives from disruption of the approximately cubic overall shape of the AQP tetramer (55 × 65 × 65 Å in the case of HsAQP5). As the residues of the tag must be incorporated within the crystal, they are likely to change the dimensions of the asymmetric unit. This can reduce the risk of packing errors caused by insertion of the AQP in randomly alternating orientations within the same packing symmetry. Comparison of the crystal packing of HsAQP5 with and without the C-terminal extension serves to illustrate this point. In crystals of HsAQP5 without a His tag (Fig. 1*b*) the tetramers pack with the soluble surfaces facing each other in close contact. In contrast, there is a gap between the equivalent (cytoplasmic) tetramer surfaces in crystals of His-tagged HsAQP5 (Fig. 1*c*). This gap accommodates the bulky disordered tag and affects the crystal packing such that a different space group is obtained. The average spacing between each protein layer in crystals of nontagged HsAQP5 is 46.2 Å, whereas it is 50.5 Å in the presence of the affinity tag. Interestingly, for SoPIP2;1 with a C-terminal His tag the increased spacing between layers within the crystal allowed a 16 Å movement of one of the loops in the protein (loop *D*), revealing the open conformation of this gated AQP (Törnroth-Horsefield *et al.*, 2006).

A recent survey of more than 1000 entries in the PDB with annotated His tags showed little influence of His tags on the final protein structure. Slight shifts towards higher *B* factors and greater solvent content for tagged structures were observed, but neither the resolution nor the *R* factors seemed to be affected (Carson *et al.*, 2007). Although His tags do not seem to have a significant effect on the structure of soluble proteins (Carson *et al.*, 2007), this does not appear to be the case for eukaryotic AQPs. In contrast, sequence modifications of regions extending out from the detergent micelles are shown here to alter the crystal packing, affect the final diffraction resolution and to reduce or prevent twinning of three eukaryotic membrane proteins. It is possible that a minimal fusion construct containing only six histidine residues, rather than the extended tags described here, may not be sufficient to show such effects. Additionally, this work only considers aquaporins, membrane proteins of very similar structure. However, even with these limitations in mind, the observations presented here may be extrapolated to other membrane proteins such that when encountering membrane protein crystals suffering from merohedral twinning, a strategy of parallel experiments using a variety of fusion-tag constructs may be an attractive approach.

This work was supported by the Swedish Science Research Council (VR) and the Swedish Foundation for Strategic Research (SSF), The European Commission Integrated Project E-MEP, the Carl Tryggers Stiftelse, the Stiftelsen Olle Engkvist Byggmästare, the University of Gothenburg Quantative Biology Platform, the National Research School in Functional Genomics and Bioinformatics, Formas and the Research School in Pharmaceutical Sciences (FLÄK).

References

Andrade, S. L., Dickmanns, A., Ficner, R. & Einsle, O. (2005). *Proc. Natl Acad. Sci. USA*, **102**, 14994–14999.
Belrhali, H., Nollert, P., Royant, A., Menzel, C., Rosenbusch, J. P., Landau, E. M. & Pebay-Peyroula, E. (1999). *Structure*, **7**, 909–917.
Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. & Bourne, P. E. (2000). *Nucleic Acids Res.* **28**, 235–242.
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.
Carson, M., Johnson, D. H., McDonald, H., Brouillette, C. & DeLucas, L. J. (2007). *Acta Cryst.* D**63**, 295–301.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Dauter, Z. (2003). *Acta Cryst.* D**59**, 2004–2016.
Derewenda, Z. S. (2004). *Methods*, **34**, 354–363.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Faham, S., Yang, D., Bare, E., Yohannan, S., Whitelegge, J. P. & Bowie, J. U. (2004). *J. Mol. Biol.* **335**, 297–305.
Fu, D., Libson, A., Miercke, L. J., Weitzman, C., Nollert, P., Krucinski, J. & Stroud, R. M. (2000). *Science*, **290**, 481–486.
Harries, W. E. C., Akhavan, D., Miercke, L. J. W., Khademi, S. & Stroud, R. M. (2004). *Proc. Natl Acad. Sci. USA*, **101**, 14045–14050.
Horsefield, R., Nordén, K., Fellert, M., Backmark, A., Törnroth-Horsefield, S., Terwisscha van Scheltinga, A. C., Kvassman, J., Kjellbom, P., Johanson, U. & Neutze, R. (2008). *Proc. Natl Acad. Sci. USA*, **105**, 13327–13332.
Iwata, S. (2003). *Methods and Results in Crystallization of Membrane Proteins.* La Jolla: International University Line.
Jiang, J., Daniels, B. V. & Fu, D. (2006). *J. Biol. Chem.* **281**, 454–460.
Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.
Kabsch, W. (1988). *J. Appl. Cryst.* **21**, 916–924.
King, L. S., Kozono, D. & Agre, P. (2004). *Nature Rev. Mol. Cell Biol.* **5**, 687–698.
Kurisu, G., Zhang, H., Smith, J. L. & Cramer, W. A. (2003). *Science*, **302**, 1009–1014.
Lee, J. K., Kozono, D., Remis, J., Kitagawa, Y., Agre, P. & Stroud, R. M. (2005). *Proc. Natl Acad. Sci. USA*, **102**, 18932–18937.
Leslie, A. G. W. (1992). *Jnt CCP4/ESF–EACBM Newsl. Protein Crystallogr.* **26**.
Luecke, H., Richter, H. T. & Lanyi, J. K. (1998). *Science*, **280**, 1934–1937.
Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.
Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.
Nyblom, M., Öberg, F., Lindkvist-Petersson, K., Hallgren, K., Findlay, H., Wikstrom, J., Karlsson, A., Hansson, Ö., Booth, P. J., Bill, R. M., Neutze, R. & Hedfalk, K. (2007). *Protein Expr. Purif.* **56**, 110–120.
Ostermeier, C. & Michel, H. (1997). *Curr. Opin. Struct. Biol.* **7**, 697–701.
Padilla, J. E. & Yeates, T. O. (2003). *Acta Cryst.* D**59**, 1124–1130.
Parsons, S. (2003). *Acta Cryst.* D**59**, 1995–2003.
Rees, D. C. (1980). *Acta Cryst.* A**36**, 578–581.
Savage, D. F., Egea, P. F., Robles-Colmenares, Y., O'Connell, J. D. III & Stroud, R. M. (2003). *PLoS Biol.* **1**, E72.
Sheldrick, G. M. (2008). *Acta Cryst.* A**64**, 112–122.
Sui, H., Han, B. G., Lee, J. K., Walian, P. & Jap, B. K. (2001). *Nature (London)*, **414**, 872–878.
Törnroth-Horsefield, S., Wang, Y., Hedfalk, K., Johanson, U., Karlsson, M., Tajkhorshid, E., Neutze, R. & Kjellbom, P. (2006). *Nature (London)*, **439**, 688–694.
Vagin, A. & Teplyakov, A. (2000). *Acta Cryst.* D**56**, 1622–1624.
Yeates, T. O. (1997). *Methods Enzymol.* **276**, 344–358.
Yoshimura, K. & Kouyama, T. (2008). *J. Mol. Biol.* **375**, 1267–1281.
Zwart, P. H., Grosse-Kunstleve, R. W., Lebedev, A. A., Murshudov, G. N. & Adams, P. D. (2008). *Acta Cryst.* D**64**, 99–107.