Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization and preliminary crystallographic analysis of the *Escherichia coli* water channel AqpZ

Brenda V. Daniels, Jian-Sheng Jiang and Dax Fu*

Biology Department, Brookhaven National Laboratory, Upton, NY 11973-5000, USA

Correspondence e-mail: dax@bnl.gov

Received 26 August 2003
Accepted 24 December 2003

AqpZ is a 24 kDa integral membrane protein that facilitates water movement across the plasma membrane of *Escherichia coli*. In this study, the first crystallization and preliminary X-ray analysis of AqpZ are described. AqpZ was overexpressed and purified with a yield of 13 mg of purified AqpZ per litre of cell culture. The purified AqpZ was shown to be a monodisperse species consisting of tetrameric protein–detergent complexes. A crystallization condition for producing diffraction-quality crystals was identified. Initial X-ray analysis indicated that the diffraction limit of AqpZ extended to 3.6 Å. Crystals were found to belong to space groups $P4_122$ or $P4_322$, with unit-cell parameters $a = b = 119.04$, $c = 380.23$ Å.

1. Introduction

Fluid balance is a ubiquitous biological process in all living organisms. Rapid transmembrane fluxes of water and small non-electrolytes are mediated by a superfamily of fluid-transporting membrane channels (Preston & Agre, 1991; Preston *et al.*, 1992). Based on their permeability properties, these channels are subdivided into aquaporins, a subfamily of strict water-selective channels, and aquaglyceroporins, a homologous subfamily of less selective channels with permeability for small neutral solutes such as glycerol. More than 150 aquaporins and aquaglyceroporins have been identified in organisms ranging from bacteria to mammals, including ten human isoforms (AQP0–9). The *Escherichia coli* aquaporin AqpZ and aquaglyceroporin GlpF share 25% pairwise sequence identity with all the human isoforms identified thus far (Unger, 2000; Beitz & Schultz, 1999). Previous structural analysis has led to structure determinations of AQP1 (Sui *et al.*, 2001) and GlpF (Fu *et al.*, 2000). AQP1 and GlpF show strikingly similar structures. Both crystallized as tetramers, with each monomeric subunit as an asymmetric unit. The monomers are folded in an α-barrel structure, with a central channel surrounded by six full-length transmembrane helices and two half-spanning helices that are joined head-to-head in the middle of the membrane to give the appearance of a kinked transmembrane span. In both AQP1 and GlpF an internal quasi-twofold symmetry is evident corresponding to the two tandem repeats in the linear sequence that may arise from an intragenic gene-duplication event (Pao *et al.*, 1991). Despite sequence homology and structural similarity, it is not clear why GlpF selects glycerol over water while the homologous AqpZ selects water over glycerol. Here, we report the crystallization and preliminary crystallographic analysis of AqpZ. Determining the structure of AqpZ will allow a direct structural comparison with GlpF in order to identify structural determinants for water–glycerol selectivity.

2. Methods

2.1. Protein expression and purification

The entire open reading frame sequence of AqpZ was obtained by PCR using the genomic DNA of *E. coli* BL21 strain as a template and a pair of AqpZ-specific primers (forward primer 5′-CGGCAGCCATATGTTCAGAAAATTA-GCAGC-3′, reverse primer 5′-GGATCCTC-GAGTTAATCACGCTTTTCCAG-3′). A host strain BL21(DE3)pLysS was used for over-expression of AqpZ, which was induced with 0.2 m*M* isopropyl-β-D-thiogalactoside (IPTG). Cells were grown at 310 K in Luria–Bertani (LB) broth containing 100 µg ml^{-1} ampicillin. 4 h after induction, cells were harvested and membrane vesicles were prepared by three passages of the cells through an ice-chilled microfluidizer (Microfluidics) at 6.9 MPa; the resulting membrane vesicles were collected by centrifugation at 140 000*g* for 45 min. The membrane pellet was washed with Tris buffer [20 m*M* Tris, 300 m*M* NaCl, 2 m*M* 2-mercaptoethanol (β-ME) pH 8.0] and membrane proteins were then extracted with a detergent buffer [100 m*M* NaCl, 20 m*M* HEPES pH 7.5, 7% β-D-glucopyranoside (β-OG, Anatrace), 20%(*w*/*v*) glycerol, 2.0 m*M* β-ME]. Insoluble cellular debris was pelleted from the supernatant by an additional centrifugation step at 140 000*g* for 30 min. The detergent extract was passed through a DEAE column (Amersham Pharmacia) and the flowthrough was loaded onto an Ni^{+}–NTA column (Qiagen). The column was washed free of non-specific protein

 DOI: 10.1107/S090744490302972X

binding with a wash solution [20 m*M* HEPES pH 7.5, 300 m*M* NaCl, 20%(*w*/*v*) glycerol, 0.75% *β*-OG, 0.5 m*M* *β*-ME, 50 m*M* imidazole]. AqpZ was eluted using wash solution with additional imidazole added to 500 m*M*. The high salt content of the purified AqpZ was removed using a desalting column (Econo-Pac 10 DG, BioRad) pre-equilibrated with desalting buffer (20 m*M* HEPES pH 7.5, 100 m*M* NaCl, 20% glycerol, 0.75% *β*-OG, 0.5 m*M* *β*-ME). The N-terminal His tag was removed by thrombin digestion. AqpZ was further purified by sizing HPLC using a TSK 4000 gel-filtration column (Tosoh Corporation) with the following solution as the mobile phase: 20 m*M* HEPES pH 7.0, 100 m*M* NaCl, 20% glycerol, 0.75% *β*-OG, 0.5 m*M* *β*-ME. HPLC-purified AqpZ was concentrated to 10 mg ml^{-1} using an Amicon stirred cell with a 10 kDa molecular-weight cutoff (MWCO) membrane. Prior to crystallization setups, concentrated AqpZ was dialyzed in a DispoDialyzer (50 kDa MWCO, Spectrum Laboratories) against 2.0 m*M* HEPES pH 7.0, 20% glycerol, 0.75% *β*-OG and 4 m*M* DTT.

2.2. Mass-spectrometric analysis

Sinapinic acid matrix was prepared as a saturated solution in a 6:3:1 mixture of acetonitrile, water and 10% trifluoroacetic acid. Aliquots of protein (10 mg ml^{-1}) were spotted onto the sample plate using the sandwich method (Beavis & Chait, 1996). MALDI–TOF mass spectroscopy data were collected at a Voyager Biospectrotometry Workstation operating in the linear mode. The spectra were externally calibrated using a calibration mixture containing cytochrome *c* and bovine serum albumin.

2.3. Crystallization

Crystallization trials were set up by the standard hanging-drop vapor-diffusion method (McPherson, 1990). 2 µl protein sample was mixed with 2 µl reservoir solution, deposited on a siliconized cover slip and equilibrated against 1 ml reservoir solution sealed in Linbro plates. Initial crystallization conditions were obtained using a modified Hampton Research Cryoscreen kit. Specifically, 7.5% *β*-OG stock solution was added to 0.75% to each of the Hampton screen conditions to make a final reservoir solution containing 0.75% *β*-OG and 90% of the original screen condition. Crystal nucleation and growth took place at room temperature. After the appearance of plate-like crystals in the initial trials, optimization conditions were obtained using Hampton Additive Screens with the addition of 100 m*M* $MgSO_4$.

2.4. Data collection and processing

X-ray data were collected at a wavelength of 1.1 Å on beamline X25 at the National Synchrotron Light Source, Brookhaven National Laboratory using an ADSC Q315 detector. Glycerol was added to the crystal drop to 25% prior to freezing the crystals in liquid nitrogen. Pre-frozen crystals were immediately mounted directly in a nitrogen-gas stream at 100 K. The crystal-to-detector distance was set to 450 mm and 347 images were collected in 1° oscillations, using 30 or 60 s exposures. The data were indexed, integrated and merged with *HKL*2000 (*DENZO* and *SCALEPACK*; Otwinowski & Minor, 1997).

3. Results and discussion

3.1. Purification and characterization

AqpZ yields after nickel-affinity purification reached 10–13 mg per litre of cell culture. The protein yield depended on the cell density at the time of IPTG induction. The optimal cell density for IPTG induction was OD_{600} = 0.6–0.9. The purification of AqpZ was achieved by single-step metal-affinity chromatography. The purity was verified by SDS–PAGE (Fig. 1*a*), showing a major protein band (lane 6) eluted from the Ni^+–NTA column by 500 m*M* imidazole. Lane 7 is the same purified AqpZ sample after thrombin digestion. The downward shift of the protein band is consistent with tag removal by thrombin digestion. The digestion of 50 mg AqpZ using 50 U thrombin appeared to be complete after overnight incubation at room temperature. The apparent molecular weight of the purified AqpZ in SDS–PAGE is around 75 kDa. This is consistent with the previous report that AqpZ migrates as an oligomer rather than a monomer in SDS–PAGE (Borgnia *et al.*, 1999).

To further confirm the molecular identity of the purified AqpZ, we determined the molecular weight by mass spectroscopy. The mass spectrum in Fig. 1(*b*) was obtained with AqpZ after partial thrombin digestion. The major mass peak at 23976.41 Da is consistent

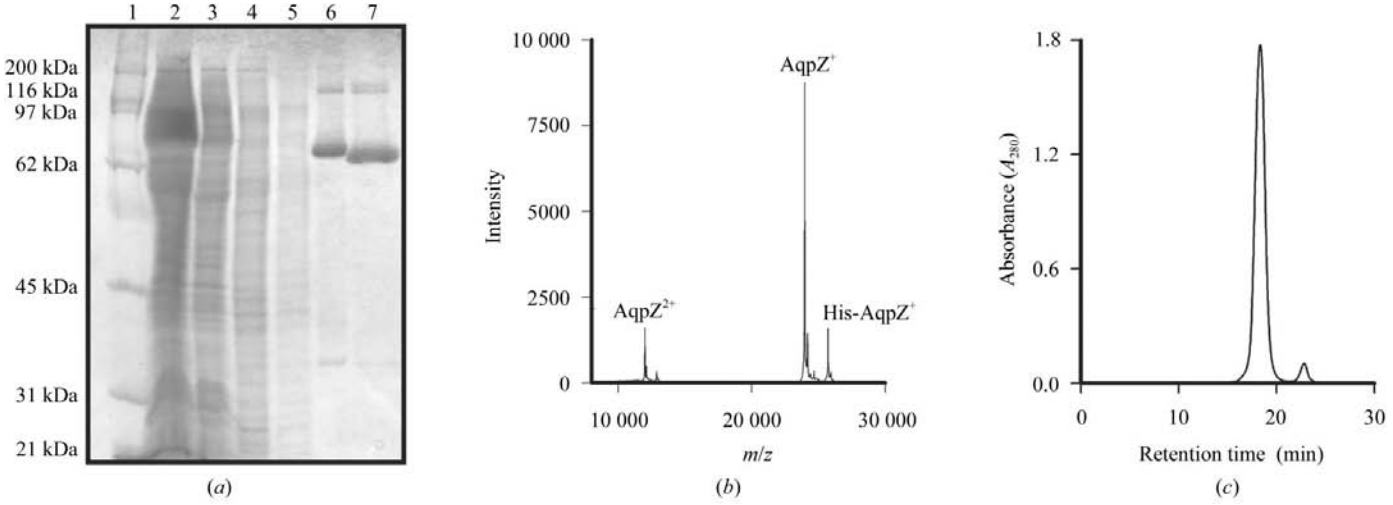
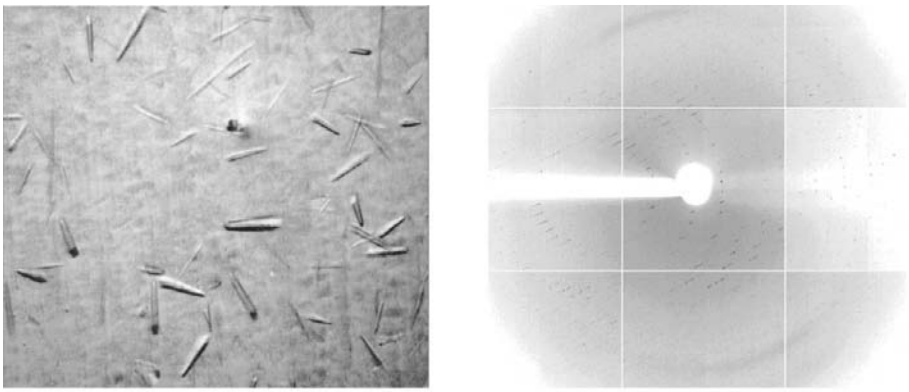


Figure 1
(*a*) SDS–PAGE analysis of AqpZ purification. Proteins were stained with Coomassie Blue. Lane 1, molecular-weight markers; lane 2, *β*-OG membrane extract; lane 3, Ni–NTA^+ flowthrough; lanes 4 and 5, first and second column washes; lane 6, column elution; lane 7, purified His-AqpZ after thrombin cleavage. (*b*) MALDI–MS of partially digested His-AqpZ. The main mass peak corresponds to AqpZ, while the minor peak corresponds to His-AqpZ. The +1 and +2 charge states of AqpZ are indicated. (*c*) Sizing HPLC chromatogram of AqpZ. 50 µl of 10 mg ml^{-1} AqpZ was injected into an TSK 4000 column run at a flow rate of 0.6 ml min^{-1}.



(*a*) (*b*)

Figure 2
X-ray diffraction-quality AqpZ crystals. (*a*) AqpZ crystals in a hanging drop, with average dimensions 20 × 20 × 120 mm. (*b*) Typical X-ray diffraction image of AqpZ crystals. The diffraction image was recorded with 1° oscillation on beamline X25 at the National Synchrotron Light Source.

with the tag-removed AqpZ, which has an expected molecular weight of 23 979.1 Da. The minor peak at 25 724.87 Da matches well with His-AqpZ, which has an expected molecular weight of 25 729.95 Da.

Sizing HPLC analysis of AqpZ showed a major monodisperse species with a retention time of about 18 min (Fig. 1*c*). The monodispersity of AqpZ was confirmed by dynamic light-scattering analysis (data not shown). The retention time of AqpZ was found to be near-identical to that of GlpF, suggesting that AqpZ is a tetramer in the detergent solution.

3.2. Crystallization

AqpZ crystallized under many different conditions, with crystal formation taking between a few weeks and nine months. Initial screening suggested PEG 3350 to be the main precipitant. We thus focused on PEG 3350 conditions using Hampton Crystal Screen kits with modifications to accommodate the need for detergents. Microcrystals first appeared as needles, as well as in droplets of protein phase separations that grew into thin plates. The best condition appeared to be 16% PEG 3350, 0.75% β-OG, 20% glycerol, 0.064 *M* sodium acetate pH 4.6, 0.128 *M* $(NH_4)_2SO_4$ and 100 m*M* $MgSO_4$. We further refined the crystallization conditions using Hampton Additive Screens and obtained diffraction-quality crystals that grew in the above crystallization condition supplemented with 10 m*M* $CoCl_2$. Typical crystals grew to dimensions of 20 × 20 × 120 µm (Fig. 2)

3.3. Data collection and processing

A complete native data set was collected from a single crystal on beamline X25 at the National Synchrotron Light Source, Brookhaven National Laboratory. The crystal diffracted to 3.2 Å with 100% completeness. Analysis of the average intensity as a function of resolution shell using *TRUNCATE* (Collaborative Computational Project, Number 4, 1994; French & Wilson, 1978) indicated that AqpZ diffraction was strongly anisotropic, with rapid intensity fall-off along the c^* axis. We therefore report the data to 3.6 Å with 93.3% completeness (Table 1). The data were indexed as primitive tetragonal, with unit-cell parameters $a = b = 119.04$, $c = 380.23$ Å. Systematic absences eliminated space groups such as $P4$, $P42_12$ or $P4_32_12$. Diffraction data were merged using space group $P4_122$ or $P4_322$, giving an $R_{\rm merge}$ of 14.3% (34% for the outer shell) and $I/\sigma(I) = 9.8$ (3.9 for the outer shell). The total number of observed reflections was 436 370, with 30 767 unique reflections. A Matthews coefficient (Matthews, 1968) of 6.46 Å^3 Da^{-1} was calculated assuming one AqpZ tetramer per asymmetric unit, which corresponds to 81% solvent content by volume. The Matthews coefficient and solvent content were 3.23 Å^3 Da^{-1} and 62%, respectively, assuming the presence of two tetramers in the asymmetric unit. Self-rotation function calculations (Tong & Rossmann, 1990) confirmed the fourfold crystallographic axis and revealed additional twofold non-crystallographic symmetry, in agreement with the assumption that there are two AqpZ tetramers per asymmetric unit. The estimated mosaicity of the crystal was in the range 0.3–0.35°. This is considerably lower than that of the diffraction obtained from the homologous *E. coli* GlpF crystal, which had a 1.0–1.5° mosaicity with a diffraction limit of 2.2 Å resolution. Improving the diffraction limit of AqpZ may be possible by further refining the crystallization conditions in order to increase crystal sizes.

Table 1
Summary of crystallographic data collection and processing.

Values in parentheses refer to the highest resolution shell (3.73–3.6 Å).

Wavelength (Å)	1.1
Resolution limits (Å)	3.6
Total reflections	436370
Unique reflections	30767
Completeness (%)	93.3 (80.0)
$R_{\rm merge}$† (%)	14.3 (33.8)
$I/\sigma(I)$	9.8 (3.9)
Multiplicity	14.2 (5.3)
Space group	$P4_122/P4_322$
Unit-cell parameters (Å)	$a = b = 119.04$, $c = 380.23$

† $R_{\rm merge} = \sum|I - \langle I\rangle|/\sum I$.

We are grateful to Dr M. Becker (National Synchrotron Light Source, Brookhaven National Laboratory) for assistance in X-ray diffraction experiments. This work was supported by a NIH grant (GM65137 to DF).

References

Beavis, R. & Chait, B. (1996). *Methods Enzymol.* **270**, 519–551.
Beitz, E. & Schultz, J. E. (1999). *Curr. Med. Chem.* **6**, 457–467.
Borgnia, M. J., Kozono, D., Calamita, G., Maloney, P. C. & Agre, P. (1999). *J. Mol. Biol.* **291**, 1169–1179.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
French, G. S. & Wilson, K. S. (1978). *Acta Cryst.* A**34**, 517–525.
Fu, D., Libson, A., Miercke, L. J. W., Weitzman, C., Nollert, P., Krucinski, J. & Stroud, R. M. (2000). *Science*, **290**, 481–486.
McPherson, A. (1990). *Eur. J. Biochem.* **189**, 1–23.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Pao, G. M., Wu, L. F., Johnson, K. D., Hofte, H., Chrispeels, M. J., Sweet, G., Sandal, N. N. & Saier, M. H. Jr (1991). *Mol. Microbiol.* **1**, 33–37.
Preston, G. M. & Agre, P. (1991). *Proc. Natl Acad. Sci. USA*, **88**, 11110–11114.
Preston, G. M., Carroll, T. P., Guggino, W. B. & Agre, P. (1992). *Science*, **256**, 385–387.
Sui, H., Bong-Gyoon, H., Lee, J. K., Walian, P. & Jap, B. K. (2001). *Nature (London)*, **414**, 872–878.
Tong, L. & Rossmann, M. G. (1990). *Acta Cryst.* A**46**, 783–792.
Unger, V. M. (2000). *Nature Struct. Biol.* **7**, 1082–1084.