Projection Map of Aquaporin-9 at 7 Å Resolution

Hector Viadiu, Tamir Gonen and Thomas Walz

Department of Cell Biology, Harvard Medical School, Boston, MA 02115, USA

Aquaporin-9, an aquaglyceroporin present in diverse tissues, is unique among aquaporins because it is not only permeable to water, urea and glycerol, but also allows passage of larger uncharged solutes. Single particle analysis of negatively stained recombinant rat aquaporin-9 revealed a particle size characteristic of the tetrameric organization of all members of the aquaporin family. Reconstitution of aquaporin-9 into two-dimensional crystals enabled us to calculate a projection map at 7 Å resolution. The projection structure indicates a tetrameric structure, similar to GlpF, with each square-like monomer forming a pore. A comparison of the pore-lining residues between the crystal structure of GlpF and a homology model of aquaporin-9 locates substitutions in these residues predominantly to the hydrophobic edge of the tripathic pore of GlpF, providing first insights into the structural basis for the broader substrate specificity of aquaporin-9.

Introduction

Aquaporins are subdivided into two classes depending on whether they only conduct water (aquaporins) or also allow passage of small neutral solutes such as glycerol across the membrane (aquaglyceroporins). All members of the aquaporin family are impermeable to protons, thereby preserving transmembrane proton gradients. Aquaporin-9 (AQP9) was first identified in adipose tissue and liver, but has now also been found in many other tissues (e.g. lung, pancreas, testis, central nervous system, retina, thyroid gland, intestines and anterior pituitary). The brain expresses two AQP9 isoforms; the shorter, 25 kDa AQP9 isoform is an alternative splice form of the 32 kDa protein found in liver and is expressed in brain mitochondria. While this study hypothesized that mitochondrial AQP9 in brain cells may also play a role in Parkinson’s disease, a recent study provided evidence against functionally significant AQP expression in mitochondria. AQP9 is unique among aquaporins, because it has broader substrate specificity than any other known member of this family. Expression of AQP9 in Xenopus oocytes has shown it to be permeable not only to water, glycerol and urea but also to polyols, carbamides, purines, pyrimidines, nucleosides and monocarboxylates. AQP9 is also one of only two mammalian AQPs known to be permeable to arsenite and might thus contribute to the toxicity of arsenic ingestion. Since AQP9 is expressed in myeloid and lymphoid leukemia cell lines, AQP9 is also likely to contribute to the efficacy of arsenite as a chemotherapeutic agent for acute promyelocytic leukemia.

The atomic structures of seven aquaporins have been published to date. Six structures of pure water pores, those of red blood cell AQP1, Escherichia coli water pore AQPZ, lens-specific AQP0, rat AQP4, plant aquaporin SoPiP, and archaeabacterial AQPM, but only one of an aquaglyceroporin, the E. coli glycerol facilitator GlpF. All aquaporins assemble into tetramers, with each monomer forming an independent pore. They share the same fold consisting of two tandem repeats that sit in the membrane in opposite orientations. The first repeat contains helices 1 to 3 and reentrant loop B and the second repeat is formed...
by helices 4 to 6 and reentrant loop E. The two reentrant loops contain the aquaporin signature motif Asn-Pro-Ala (NPA) and form short pore-lining α-helices, named HB and HE. The first atomic model of an aquaporin, AQP1, already revealed the mechanism used by aquaporins to prevent proton conduction, which is based on the two NPA motifs and the opposing electrostatic fields of the two pore helices HB and HE. The proton exclusion mechanism was subsequently further refined using molecular dynamics simulations. Using the crystal structures of the *E. coli* water pore AQPZ and the *E. coli* glycerol facilitator GlpF, a molecular dynamics approach was also recently used to pinpoint the structural features that determine whether an aquaporin pore is permeable to glycerol or not. Once a structure for AQP9 is determined, the same approach should also be useful to elucidate why this aquaglyceroporin can have such a broad substrate specificity while still preventing proton conduction. Here, we present the first structural data for AQP9 and provide an initial explanation why AQP9 may accommodate conduction of larger solutes than any other aquaglyceroporins characterized to date.

**Results**

**Protein expression and purification**

Recombinant rat AQP9 (295 amino acid residues) was purified in octyl-β-D-glucopyranoside (OG). On a silver-stained SDS–PAGE gel, the majority of AQP9 runs as a monomer, but traces of dimers and tetramers are also present (Figure 1(a), lane 1). The identities of the dimer and tetramer bands were confirmed by immunoaffinity blots using anti-His-tag antibodies. The AQP9 monomer runs as a double band with apparent molecular masses of ~32 kDa and ~35 kDa. Incubation with N-glycosidase F (PNGase F) eliminated the ~35 kDa band (Figure 1(a), lane 2), showing that it contains the N-glycosylated form of AQP9. The likely glycosylation site is Asn142 in loop C, which is part of an N-glycosylation consensus sequence, Asn-Xaa-Ser/Thr. This differs from AQP1, in which the glycosylation site is located in loop A.

**Single particle electron microscopy analysis**

We used single particle electron microscopy to assess the quality of our AQP9 preparations and the suitability of the recombinant protein for 2D crystallization trials. Electron micrographs of purified AQP9 negatively stained with uranyl formate showed mono-disperse particles that were homogenous in size (Figure 1(b)). To avoid interference with the staining of the protein, the detergent was removed from AQP9 adsorbed to the carbon film by washing the grid with five drops of deionized water prior to staining with uranyl formate. A total of 6220 particles were selected from three images and classified into 24 groups. Representative class averages shown in Figure 1(c) reveal a square shape with a side length of about 70 Å, which correlates well with the side length of 68 Å for AQP1 tetramers seen in 2D crystals. As observed before with other AQP preparations, e.g. AQP1, only top views of AQP9 tetramers were seen in the class averages. The reason is that the protein presumably only adsorbs with its hydrophilic extramembranous moiety to the charged carbon film rather than with its hydrophobic detergent-covered transmembrane moiety. The single particle analysis indicated that AQP9 has the same tetrameric organization characteristic of all members of the aquaporin family and the preparations were suitable for 2D crystallization trials.

**Figure 1.** Protein purification and single particle electron microscopy analysis. (a) Silver stained SDS–10% (w/v) PAGE gel of purified recombinant rat AQP9 before (lane 1) and after deglycosylation with PNGase F (lane 2). (b) Image of negatively stained AQP9 showing that the OG-solubilized particles are mono-dispersed and homogeneous in size. The scale bar represents 10 nm. (c) Representative single particle averages. The side length of the individual panels represent 16.5 nm.
Two-dimensional crystallization

We reconstituted AQP9 into 2D crystals. While AQP9 could be reconstituted into lipid membranes with 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC), 1,2-dioleoyl-sn-glycero-3-phosphocholine (DOPC) and *E. coli* total lipid extract, the best ordered 2D crystals were formed with 1,2-dimyristoyl-sn-glycero-3-phosphocholine (DMPC). De-glycosylation of AQP9 did not have a noticeable effect on the order of the resulting 2D arrays, and 2D crystallization trials were therefore continued with glycosylated AQP9. To obtain well-ordered crystals, it proved essential to remove the His-tag. Magnesium and calcium ions promoted formation of 2D crystals. Calcium was used for further 2D crystallization trials, because it resulted in crystals that were better ordered than those grown in the presence of magnesium. The best 2D crystals were obtained by dialyzing the protein–lipid mixture against a buffer containing 20 mM Mes (pH 6), 150 mM NaCl, 50 mM CaCl2 and 1 mM dithiothreitol (Figure 2(a) and (b)). Initial crystallization trials were carried out at a final protein concentration in the reconstitution mixture of 1 mg/ml, later screens were performed at a protein concentration of 3 mg/ml, which increased the size of the formed arrays.

Projection map

AQP9 2D crystals were embedded in glucose and low-dose images were collected at liquid nitrogen temperature. The quality of the images was assessed by optical diffraction, and high-quality images (with no signs of specimen drift or charging) that showed diffraction spots beyond a resolution of 10 Å were selected for further image processing. After unbending and correction for the contrast transfer function (CTF), CTF plots typically showed good completeness of diffraction data to 7 Å resolution, and occasionally some strong diffraction spots to a resolution beyond 5 Å (Figure 2(c)). The AQP9 2D crystals have unit cell dimensions of $a=b=103\pm1$ Å and $\gamma=90^\circ$ ($n=9$) and display $p4_212$ symmetry (Table 1). The amplitude and phase data from the

Figure 2. Analysis of AQP9 2D crystals. (a) The image of a negatively stained AQP9 2D crystal clearly reveals the square crystal lattice. The scale bar represents 200 nm. (b) Fourier transform of the image shown in (a). The scale bar represents 5 nm$^{-1}$. (c) CTF plot of an image of a glucose-embedded 2D crystal. The circles represent CTF zero transitions and the boxed numbers depict the IQ values of the individual reflections as defined. The scale bar represents 1.7 nm$^{-1}$. (d) $p4_212$-symmetrized projection map of AQP9 at 7 Å resolution. A unit cell, lattice dimensions $a=b=103.0$ Å, $\gamma=90^\circ$, contains two AQP9 tetramers and is outlined in black. The scale bar represents 5 nm.
The best six images were merged to 7 Å resolution (Table 2) while applying p4212 symmetry. The final projection map of glucose-embedded AQP9 is shown in Figure 2(d). The projection map shows that a unit cell of the AQP9 2D crystals (outlined in black in Figure 2(d)) contains two tetramers that are incorporated into the membrane in opposite orientations.

Electron Diffraction

Most of the AQP9 2D crystals were vesicular, and larger vesicles tended to fold on themselves on the carbon film, reducing the useful crystalline area for structure analysis. Occasionally, AQP9 also formed single-layered crystalline sheets, which are particularly well suited to collect electron diffraction data. An electron diffraction pattern recorded from a glucose-embedded 2D crystalline sheet is shown in Figure 3. Sharp diffraction spots can be seen to a resolution of about 3.8 Å without background subtraction. The crystalline AQP9 sheets are currently not much larger than 500 nm. While this size suffices to collect diffraction patterns from untilted specimens, the crystals are currently too small to collect diffraction patterns from highly tilted specimens.

**Discussion**

The sequence identity among water-specific aquaporins and among glycerol-conducting aquaporins is higher than 35%, whereas the sequence identity between members of the two subgroups is generally less than 30%. This difference at the level of sequence identity is reflected in the projection structures of aquaporins; the projection map of the AQP9 tetramer (Figure 4(a)) resembles more that of the *E. coli* glycerol facilitator GlpF27 (Figure 4(b)) than to those of pure water pores, such as AQP125 or AQP028 (Figure 4(c)). Projections of the GlpF and AQP9 monomers have a squarish overall shape, whereas projections of monomers of pure water pores appear more wedge-shaped. The protein density in each monomer encloses a pore. GlpF has a larger pore compared to those in pure aquaporins,29 and our 7 Å projection map suggests that the pore in AQP9 is similar to the pore in GlpF. The round GlpF pore has a diameter of about 7 Å. Our AQP9 projection map shows an oval pore with approximate dimensions of 7 Å by 12 Å.

### Table 1. Internal phase residuals of all possible two-sided plane groups using the image shown in Figure 2(c)

<table>
<thead>
<tr>
<th>Two-sided plane group</th>
<th>Phase residual (deg.)</th>
<th>Number of comparisons</th>
<th>Target residual (deg.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>p1 2</td>
<td>20.1°</td>
<td>254</td>
<td>29.0</td>
</tr>
<tr>
<td>p12, b</td>
<td>35.1°</td>
<td>127</td>
<td>20.5</td>
</tr>
<tr>
<td>p12, a</td>
<td>80.1°</td>
<td>109</td>
<td>20.5</td>
</tr>
<tr>
<td>p12, b</td>
<td>18.2°</td>
<td>108</td>
<td>20.4</td>
</tr>
<tr>
<td>p12, a</td>
<td>28.1°</td>
<td>109</td>
<td>20.5</td>
</tr>
<tr>
<td>c12, b</td>
<td>76.7°</td>
<td>108</td>
<td>20.4</td>
</tr>
<tr>
<td>p22, b</td>
<td>62.5°</td>
<td>344</td>
<td>23.3</td>
</tr>
<tr>
<td>p22, a</td>
<td>61.0°</td>
<td>344</td>
<td>23.3</td>
</tr>
<tr>
<td>p22, b</td>
<td>63.7°</td>
<td>344</td>
<td>23.3</td>
</tr>
<tr>
<td>p22, a</td>
<td>28.5°</td>
<td>344</td>
<td>23.3</td>
</tr>
<tr>
<td>c22, a</td>
<td>62.5°</td>
<td>344</td>
<td>23.3</td>
</tr>
<tr>
<td>p42, b</td>
<td>26.9°</td>
<td>347</td>
<td>23.3</td>
</tr>
<tr>
<td>p42, a</td>
<td>55.0°</td>
<td>777</td>
<td>21.5</td>
</tr>
<tr>
<td>p42, b</td>
<td>24.9°</td>
<td>777</td>
<td>21.5</td>
</tr>
</tbody>
</table>

Internal phase residuals were determined using the program ALLSPACE34 from spots of IQ1 to IQ5 to 7 Å resolution. Only plane groups compatible with the AQP9 lattice are shown.

- **a** Phase residual versus other spots (90° random).
- **b** Target residual based on statistics taking Friedel weight into account.
- **c** Note that in space group p1 no phase comparison is possible, so the numbers given here are theoretical phase residuals based on the signal-to-noise ratio of the observed diffraction spots.
- **d** Within 30% of target residual.
- **e** Symmetry that yields a residual better than the target residual.
- **f** Within 20% of target residual.

### Table 2. Phase residuals in resolution shells

<table>
<thead>
<tr>
<th>Resolution shell (Å)</th>
<th>Number of phases</th>
<th>Mean value of Δαc (deg.)</th>
<th>Standard error (deg.)</th>
<th>Mean figure of merit</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>9.9</td>
<td>12.6</td>
<td>3.1</td>
<td>0.92</td>
</tr>
<tr>
<td>9.8</td>
<td>7.0</td>
<td>22.2</td>
<td>3.8</td>
<td>0.85</td>
</tr>
<tr>
<td>7.0</td>
<td>5.7</td>
<td>28.4</td>
<td>4.3</td>
<td>0.75</td>
</tr>
</tbody>
</table>

Δαc. Difference between the symmetry-imposed phase of 0° and 180° and the observed combined phase. 45° is random.
apparent enlargement of the AQP9 pore is due to a region of lower density that emanates from the round pore seen in GlpF (marked by an asterisk in Figure 4(a)), which may explain the permeability of AQP9 for bulkier solutes.

We used the homology server CPHmodels 2.0 to build a model of AQP9 based on the known GlpF structure (PDB id, 1LDA), which, among aquaporins of known 3D structure, has the highest level of identity with AQP9 (38%). GlpF has been described to have a tripathic pore with one hydrophobic corner and two sides that have either hydrogen-bond donor or acceptor groups. If we compare the AQP9 homology model with the GlpF crystal structure, we find that the weaker density of the AQP9 pore seen in our projection map extends towards the hydrophobic face of the pore.

An interesting pattern emerged when we looked at a multiple sequence alignment of rat, human and mouse AQP9 with GlpF, AQP1, AQP0 and AQPZ, four of the aquaporins with known atomic structures (Figure 5). The pore residues (marked by P in Figure 5) that are conserved in the aligned aquaporin sequences (Gly80, His82, Asn84, Asn216 and Arg219 in rat AQP9) all localized to the faces that present the hydrogen-bond donor and acceptor groups, predominantly in the center of the pore, specifically in loop B and pore helices HB and HE (Figure 6(a)). The residues that are only conserved in aquaglyceroporins but not in water-conducting aquaporins (Ile60, Val68, Leu172, Ile196, Ile200, Gly208, Gly212, Ala214 and Met215) all lie in the center of the pore on the hydrophobic side of the channel and are part of helices 2, 4 and 5 and loop E.
Finally, residues that form part of the pore and are conserved among all the AQP9 homologs but are different in GlpF (Phe64, Gly81, Met91, Val176, Phe180, Leu209 and Cys213) also fall into the hydrophobic face and are part of helices 2 and 4, loops B and E, and pore helix HB. Most of these substitutions are located close to the entrances of the pore (Figure 6(c)). By correlating the observed weaker density of the pore in our AQP9 projection map with the pattern of conserved residues in the aligned sequences, we can propose that the broader specificity found in AQP9 is related to changes in the hydrophobic face of the tripathic pore. These amino acid substitutions are located close to the openings of the pore (Figure 6(c)).

Figure 5. Multiple sequence alignment. Multiple sequence alignment of rat, human and mouse AQP9 with GlpF, AQP1, AQP0 and AQPZ (aquaporins with known 3D structures). Secondary structure assignments are shown according to the crystal structure of GlpF. Residue numbering follows the rat AQP9 sequence. Residues in green are conserved in all aquaporins, residues in blue are only conserved in glycerol-conducting aquaporins, and residues in red are only conserved in AQP9 homologs. Residues marked with a P form part of the pore in the GlpF structure, according to the program HOLE. A green P indicates a pore residue in GlpF that is conserved in all aquaporins, a blue P a pore residue in GlpF that is conserved in all glyceroporins, a red P a pore residue in GlpF that is different from all the AQP9 sequences shown, and an uncolored P a pore residue in GlpF that is identical to one or two AQP9 sequences.
of the pore and may thus allow larger solutes easier access to the pore. On the other hand, the central constricted region of the pore is conserved in all aquaporins, indicating that AQP9 uses the same proton exclusion mechanism as all the other aquaporins. For a larger solute to permeate the AQP9 constriction site may thus require a solute-induced change in this pore region similar but probably larger than the one seen in molecular dynamics simulations of glycerol permeating the GlpF pore.\(^{23}\)

Our data provide a tentative explanation for how the AQP9 pore can have such broad solute specificity without compromising proton exclusion. Our electron diffraction data (Figure 3) showing reflections in unprocessed diffraction patterns to at least 3.8 Å resolution is a promising first step towards determining the structure of the first eukaryotic aquaglyceroporin at atomic resolution.

### Materials and Methods

#### Protein expression and purification

A construct of an N-terminal 6x(His) tag separated from the rat AQP9 gene by a TEV cleavage site was expressed in Sf9 insect cells using the Bac-to-Bac baculovirus expression system (Invitrogen). Cells were harvested 48 h after virus infection. Membranes were prepared as described,\(^{30}\) and AQP9 was purified as described\(^{31}\) with some modifications. Briefly, membranes were washed with 2 M urea and alkaline buffer (pH 11) prior to solubilization with 5% (w/v) OG in 250 mM NaCl and 20 mM Tris (pH 8.5) for 15 min at 25 °C. After centrifugation at 100,000g for 45 min at 4 °C, the OG concentration of the supernatant was lowered to 2% (w/v) by dilution. Solubilized 6x(His)-tagged AQP9 was adsorbed overnight to Ni-NTA agarose at 4 °C. The resin was washed with 20 mM histidine and the protein eluted with 150 mM imidazole in 250 mM NaCl and 20 mM Tris (pH 8.5) for 2–4 h at 4 °C. The His-tag was removed by digestion with TEV-protease and gel filtration chromatography prior to crystallization. One liter of Sf9 culture yielded approximately 0.1 mg of pure AQP9. His-tagged AQP9 was detected in a Western blot using anti-His-tag antibodies conjugated with peroxidase and developed with an enhanced chemiluminescence (ECL) kit (Amersham).

#### 2D crystallization

AQP9 at concentrations ranging from 1 to 3 mg/ml was mixed with DMPC, DOPC, POPC, and *E. coli* total lipid extract (Avanti) at lipid-to-protein ratios (LPRs) ranging from 0.4 to 1 (w/w). The mixtures were placed in 50 μl
dialysis buttons and dialyzed against 20 mM Mes (pH 6), 150 mM NaCl, 50 mM CaCl$_2$ and 1 mM dithiothreitol at 25 °C for seven days with daily buffer exchanges. The best crystals were obtained with an LPR of 0.4 to 0.6.

**Electron microscopy**

OG-solubilized AQP9 was prepared by the conventional negative staining protocol as described. Images were taken with a FEI Tecnai F20 electron microscope equipped with a LaB$_6$ filament and operated at 200 kV. Images were taken at a magnification of 50,000× on Kodak SO-163 film using low-dose procedures. Images were developed for 12 min with full-strength Kodak D-19 developer at 20 °C.

AQP9 2D crystals were adsorbed to a thin carbon film on a molybdenum grid and quickly frozen in the presence of 20% (w/v) glucose and transferred to a Gatan cryotransfer holder. Images were taken with an FEI Tecnai F20 electron microscope equipped with a field emission gun and operated at 200 kV. Images were taken at a magnification of 50,000× on Kodak SO-163 film using low-dose procedures and developed as described above. The image quality and the order of the 2D crystals was assessed with a JEOL JFO-3000 laser diffractometer, and only drift-free images showing diffraction spots beyond 10 Å resolution were used for further analysis.

**Image processing**

For the single particle electron microscopy analysis, micrographs were digitized with a Zeiss SCAI scanner using a step size of 7 μm and processed with the MRC software package, as described. The six best images were merged and used to calculate a projection map to enhance the high-symmetry. A negative temperature factor of 25 °C was used for seven days with daily buffer exchanges. The best crystals were obtained with an LPR of 0.4 to 0.6.

**Sequence alignment and homology modeling**

Multiple sequence alignment of rat, human and mouse AQP9 with the sequences of aquaporins of known 3D structure was calculated with T-Coffee™ with the option to consider the structural alignment for calculations. To calculate a homology model of rat AQP9, we used the CPHmodels 2.0 Server.

**Acknowledgements**

We acknowledge contributions by Dr Matthias Hediger and Dr Ingeborg Schmidt-Krey during the early stages of this project. The authors thank Dr Stephen C. Harrison for continuous support and advice. This work was supported by NIH grant GM62580 (to T.W.). H.V. was a fellow of the Damon-Runyon Cancer Research Fund.

**References**


**Edited by W. Baumeister**

(Received 25 October 2006; received in revised form 13 December 2006; accepted 15 December 2006) Available online 20 December 2006