Supporting Information

Ho et al. 10.1073/pnas.0902725106

SI Materials and Methods

Protein Expression and Purification. The full-length human aquaporin 4 gene (NCBI accession no. NM_001650) was purchased through Origene. The expression construct was designed with an N-terminal 8xHis followed by a flag tag (DYKDDDDK) and a human rhinovirus 3C protease cleavage site (LEVLFQ↓GP) and cloned into the EcoRI and NotI sites of pPICZ expression vector (Invitrogen). The expression vector was then electroporated into *Pichia pastoris* X-33 cells (Invitrogen) using the Bio-Rad Gene Pulser Xcell System following a standard yeast electroporation protocol. Transformation was then selected on YPD plates with 50 μg/mL Zeocin (Invitrogen). Four colonies were restreaked and tested for expression. For production, the yeast was cultured in BMY media (Invitrogen) in Fernbach flasks at 30 °C for 24 h, then the temperature was lowered to 26 °C, and methanol was added directly to the cultures to a final concentration of 2.5%. The cultures were grown for another 48 h before harvest. Cultures harvested by centrifugation at 4 °C at 6,000 × *g* for 10 min. Pellets were washed once with TBS buffer with 1 mM β-ME, and 1 mM PMSF and pelleted again. Cells were then resuspended with the same buffer and lysed by bead beating with glass beads. Broken and unlysed cells were removed by centrifugation at 4 °C at 6,000 × *g* for 10 min while the membranes remained in the supernatant. The membranes were then pelleted at 160,000 × *g* at 4 °C for 1 h. Pellets were resuspended in MR Buffer (25 mM Tris·HCl, pH 7.4 at room temperature, 250 mM NaCl, 10% glycerol, 1 mM β-ME) and stored at −80 °C until further processing.

To begin purification, resuspended membrane was solubilized by adding 400 mM n-octyl-β-D-glucopyranoside (OG) (Anatrace) to a final concentration of 200 mM and stirred at 4 °C for 1 h. Unsolubilized material was pelleted at 160,000 × *g* at 4 °C for 30 min; 5 M imidazole, pH 7.4, was added to the supernatant to a final concentration of 50 mM. The supernatant was then batch bound with Ni-NTA resins (Qiagen) for 2 h, loaded onto a Bio-Rad Econo Column and washed with MR Buffer with 40 mM OG and 50 mM imidazole, and then eluted with 300 mM imidazole. Imidazole was removed using Econo-Pac DG10 desalting column (Bio-Rad) equilibrated with MR Buffer with 40 mM OG. The N-terminal tag was cleaved by His-tagged MBP fusion of human rhinovirus 3C protease (His-6-MBP-3C) at 4 °C overnight. Uncleaved hAQP4 (<5%) and 6xHis-MBP-3C were removed the next day with TALON resin (Clontech). Then, hAQP4 was concentrated in a 50,000 molecular weight cut-off Amicon spin concentrator (Millipore) and further purified by size exclusion chromatography on a Superdex 200 10/300 GL column (GE Healthcare) in 25 mM Citrate, pH 6.0, 50 mM NaCl, 5% glycerol, 40 mM OG, and 2 mM DTT (Crystal Buffer). Purified hAQP4 was concentrated again to 30 mg/mL and stored as aliquots at −80 °C. The final yield of hAQP4 was ≈15 mg of purified and concentrated protein per liter of cells.

Trypsinolysis of hAQP4. Concentrated hAQP4 (post size exclusion chromatography) was diluted with MR Buffer to 1 mg/mL, and immobilized TPCK-trypsin (Cat 20230; Thermo Scientific) was added. The mixture was incubated by gentle rocking at 25 °C overnight. Immobilized TPCK-trypsin was removed by filtration, and trypsinized hAQP4 was further purified in the same way as the full length on size exclusion chromatography. MALDI-MS and N-terminal sequencing were performed on trypsinized hAQP4, and the stable trypsinized product was identified as N-ENIMV…PDVEFK-C, which removes the N-terminal 19 residues and the C-terminal 64 residues. Trypsinized hAQP4 was concentrated to 30 mg/mL and stored at −80 °C until crystallization.

Crystallization. Full-length hAQP4 was crystallized in 25% PEG 2000-MME, 50 mM Citrate, pH 6.0, 5% Glycerol, 40 mM OG, and 2 mM DTT by hanging drop vapor diffusion at 25 °C. Bipyramidal crystals grew to a maximum size of ≈100 μm within 1 week and diffracted to a best resolution of 8 Å.

Trypsinized hAQP4 was crystallized in 25% PEG2000-MME, 50 mM Mes, pH 6.5, 5% Glycerol by sitting drop vapor diffusion at 25 °C. Both bipyramidal crystals and rod crystals were found within 1 week. The Silver Bullet Screen condition A1 (Hampton Research) was found helpful in growing the bipyramidal crystals. The rod crystals grew to a maximum size of 1.5 mm and diffracted to a best resolution of 6 Å. The bipyramidal crystals grew to a maximum size of 50 μm and diffracted to a best resolution of 1.8 Å. The best freezing condition for the crystals was a 3-s soak in a solution of 50% paraffin (vol/vol) and 50% paratone-N (vol/vol) (Hampton Research) before flash-freezing in liquid nitrogen.

Proteoliposome Assay. Lipid stocks were made from *Escherichia coli* polar lipid extracts (Avanti Polar Lipids) at 50 mg/mL in water plus 4 mM β-mercaptoethanol and stored at −80 °C. Before the essay, stocks were thawed and 250 μL lipid and 100 μL water were mixed in a 16 × 125 mm glass culture tubes (VWR) and overlaid with argon to reduce oxidation. Liposomes were formed by sonicating the culture tubes in a bath sonicator (Laboratory Supplies). The reconstitution buffer mixture was formed by mixing (in order): 100 mM Mops, pH 7.5, 43 mM β-OG, full-length or trypsinized hAQP4 (final concentration 0.8 mg/mL), and sonicated lipids (final concentration 8 mg/mL) in a total volume of 1.5 mL. It was then dialyzed against the assay buffer (20 mM Hepes, pH 7.4) in a 25,000 molecular weight cut off Spectra/Por Float-a-lyzer (Spectrum Laboratories) for 48 h. Liposomes were then harvested by centrifugation at 75,000 × *g* at 4 °C for 1 h. For water conduction, liposomes were resuspended in 1 mL of assay buffer. For glycerol conduction, liposomes were resuspended in 1 mL 20 mM Hepes, pH 7.3 plus ≈550 mM glycerol (equal in osmolarity to sucrose buffer: 20 mM Hepes, pH 7.3 with 570 mM sucrose). On mixing 100 μL of liposome resuspension with 100 μL of sucrose buffer in a stopped-flow apparatus, vesicle shrinkage was monitored by light scattering at 440 nm over time at 12 °C. For assaying water conduction in the presence of acetazolamide and rizatriptan, liposomes were incubated for 1 h in assay buffer with 1 or 10 mM of compound, and assayed against the sucrose buffer with compound. The resulting data points were fitted to a single exponential curve to extract the rate constant.

Molecular Dynamic (MD) Simulations of hAQP4 and bAQP1. MD simulations were carried out with the Gromacs 3.3.1 simulation software (1). The simulations were performed on a 256 processor computer cluster running Microsoft Windows 2003 Compute Cluster Server operating system. We used the hAQP4 X-ray crystal structure as a starting conformation. The tetramer was centrally placed into a palmitoyloleoylphosphatidylethanolamine (POPE) lipid bilayer area containing 364 lipids, and was solvated on both sides with 25507 SPC water molecules (2). The OPLSAA force field (3, 4) was used to model the protein. Lipid parameters were an extension of the parameters

of Berger et al. (5) that were optimized for the OPLSAA force field (L. Monticelli and P. Tieleman, personal communication); 45 chloride ions and 45 sodium ions were added to the periodic simulation box to create a salt concentration of 50 mM. The total system size was 108,939 atoms.

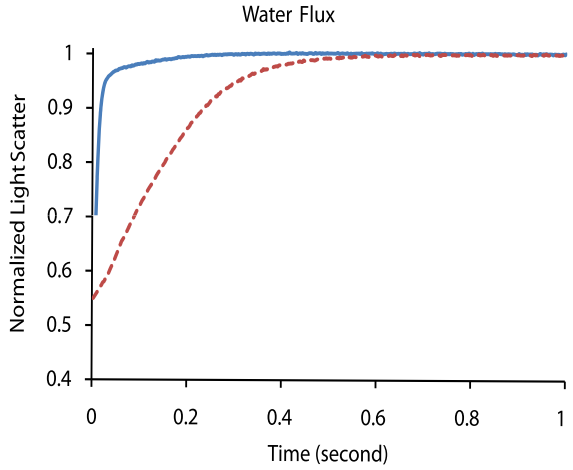
After energy minimization, 2 500-ps simulations were carried out with the protein atoms frozen in their crystallographic positions. The first simulation contained the crystallographic waters, whereas the second simulation did not. A third 1-ns simulation was then carried out with the starting positions taken from the end of the first simulation in which the heavy atom (N, C, and O) positions of the protein were restrained to their crystallographic position with a harmonic spring-like restraint, having a force constant of 1,000 $kJ{\cdot}mol^{-1}{\cdot}nm^{-2}$. A fourth 1-ns simulation was then performed, starting from the end of the third simulation, in which the restraints were removed and the protein allowed to move freely.

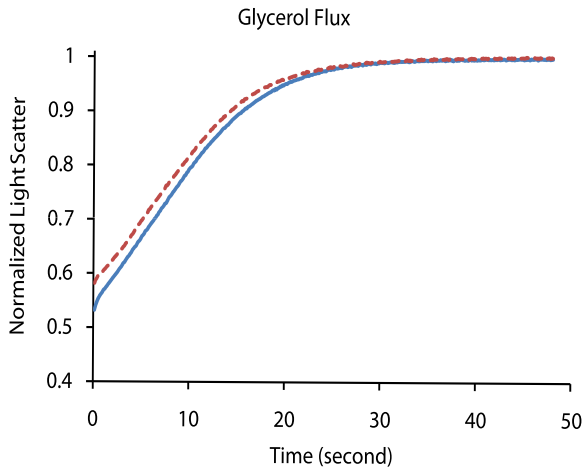
The number of water molecules along each of the four channels were collected every 1 ps and used to calculate the number densities in Fig. 5*C*. In the second simulation it took the water <200 ps to fully populate the channel region. Number densities were calculated over 100-ps intervals and block averaged to test for convergence. In all cases, convergence of the water density in the channel region was observed.

Throughout these simulations, the temperature was kept constant by coupling the system to a temperature bath of 298 K (6). Likewise, the pressure was kept constant, in the third and fourth simulations, by coupling the system to a pressure bath at 1.01 bar (7). The xy (membrane plane) and z (membrane normal) directions, respectively, were independently coupled with a coupling constant τ of 1 ps. Electrostatic interactions were calculated using PME (8) with an electrostatic cut off of 1.2 nm. The Lennard–Jones potential, describing the van der Waals interaction, was cut off at 0.9 nm. The Settle algorithm (9) was used to constrain the bond lengths and angles of the water molecules, and Lincs was used to constrain all other bond lengths (10), allowing a time-step of 2 fs.

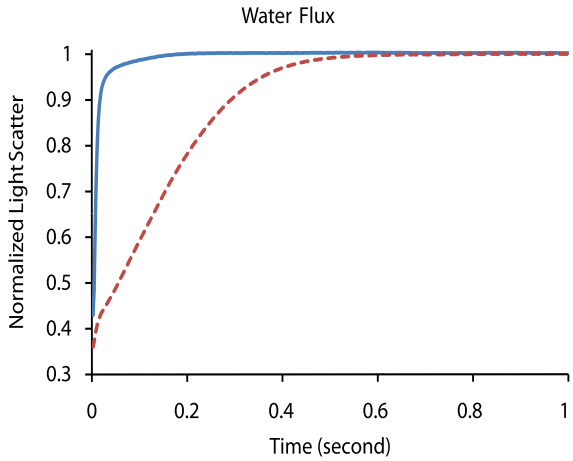
1. Lindahl E, Hess B, Van Der Spoel D (2001) GROMACS 3.0: A package for molecular simulation and trajectory analysis. *J Mol Model* 7:306–317.
2. Pullman B, ed (1981) *Intermolecular Forces: Proceedings of the Fourteenth Jerusalem Symposium on Quantum Chemistry and Biochemistry Held in Jerusalem, Israel, April 13–16, 1981* (Reidel, Dordrecht, The Netherlands).
3. Xu Z, Luo HH, Tieleman DP (2007) Modifying the OPLS-AA force field to improve hydration free energies for several amino acid side chains using new atomic charges and an off-plane charge model for aromatic residues. *J Comput Chem* 28:689–697.
4. Jorgensen WL, Maxwell DS, TiradoRives J (1996) Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J Am Chem Soc* 118:11225–11236.
5. Berger O, Edholm O, Jahnig F (1997) Molecular dynamics simulations of a fluid bilayer of dipalmitoylphosphatidylcholine at full hydration, constant pressure, and constant temperature. *Biophys J* 72:2002–2013.
6. Nose S (1984) A Molecular-Dynamics Method for Simulations in the Canonical Ensemble. *Mol Phys* 52:255–268.
7. Parrinello M, Rahman A (1981) Polymorphic Transitions in Single-Crystals: A New Molecular-Dynamics Method. *J Appl Phys* 52:7182–7190.
8. Darden T, York D, Pedersen L (1993) Particle Mesh Ewald: An N. Log(N) Method for Ewald Sums in Large Systems. *J Chem Phys* 98:10089–10092.
9. Miyamoto S, Kollman PA (1992) Settle: An Analytical Version of the Shake and Rattle Algorithm for Rigid Water Models. *J Comput Chem* 13:952–962.
10. Hess B, Bekker H, Berendsen HJC, Fraaije JGEM (1997) LINCS: A linear constraint solver for molecular simulations. *J Comput Chem* 18:1463–1472.
11. Hiroaki Y, et al. (2006) Implications of the aquaporin-4 structure on array formation and cell adhesion. *J Mol Biol* 355, 628–639.

(I) Full Length Protein





(II) Trypsinized Protein



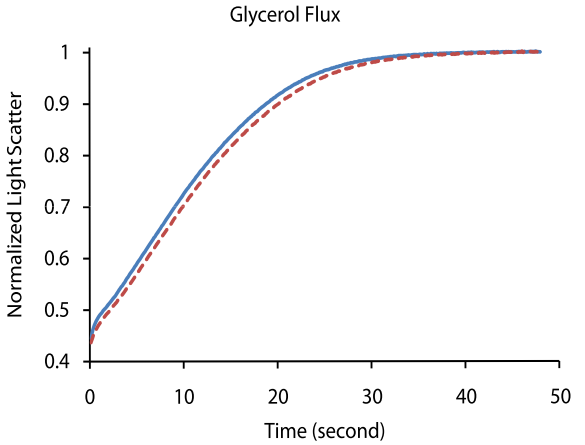
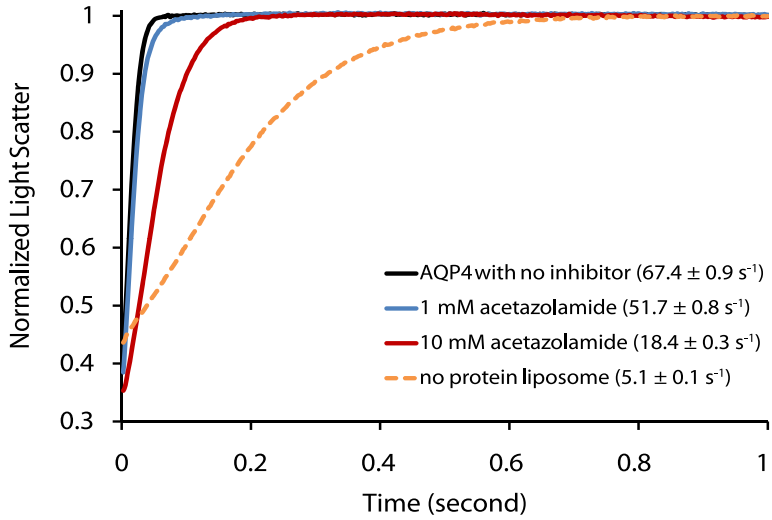


Fig. S1. (*a*) Water and glycerol conduction assays for the (*i*) full-length and (*ii*) trypsinized protein. Proteoliposome with AQP4 is shown as blue line. Liposome without protein is shown as red dashes. Plots were generated from average values from 6 replicate measurements. Rates were determined from fitting a single exponential curve to the plot. (*b*) Water conduction assays in the absence and presence of 1 or 10 mM of (*i*) acetazolamide and (*ii*) rizatriptan. Results using the full-length human AQP4 are displayed, but the trypsinized protein yielded similar results. Proteoliposomes treated with 1 or 10 mM of TEA conducted as fast as without. Empty liposomes were assayed with and without inhibitors and no change of water conduction was observed. Plots were generated from average values from 6 replicate measurements. Rates were determined from fitting a single exponential curve to the plot.

(I)



(II)

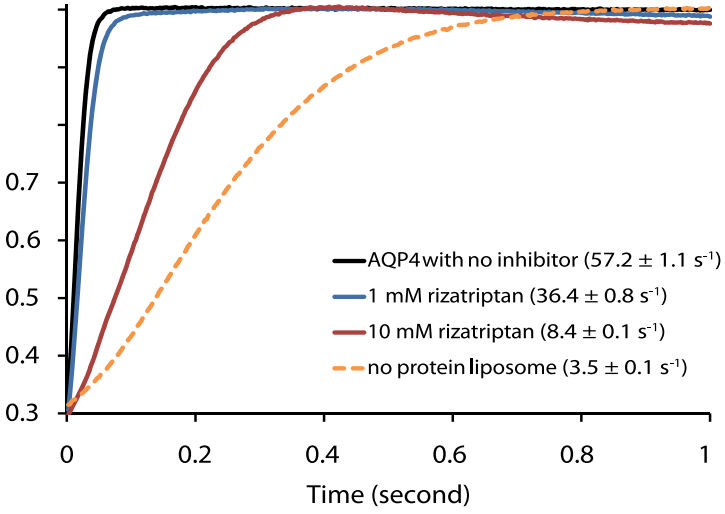


Fig. S1 continued.

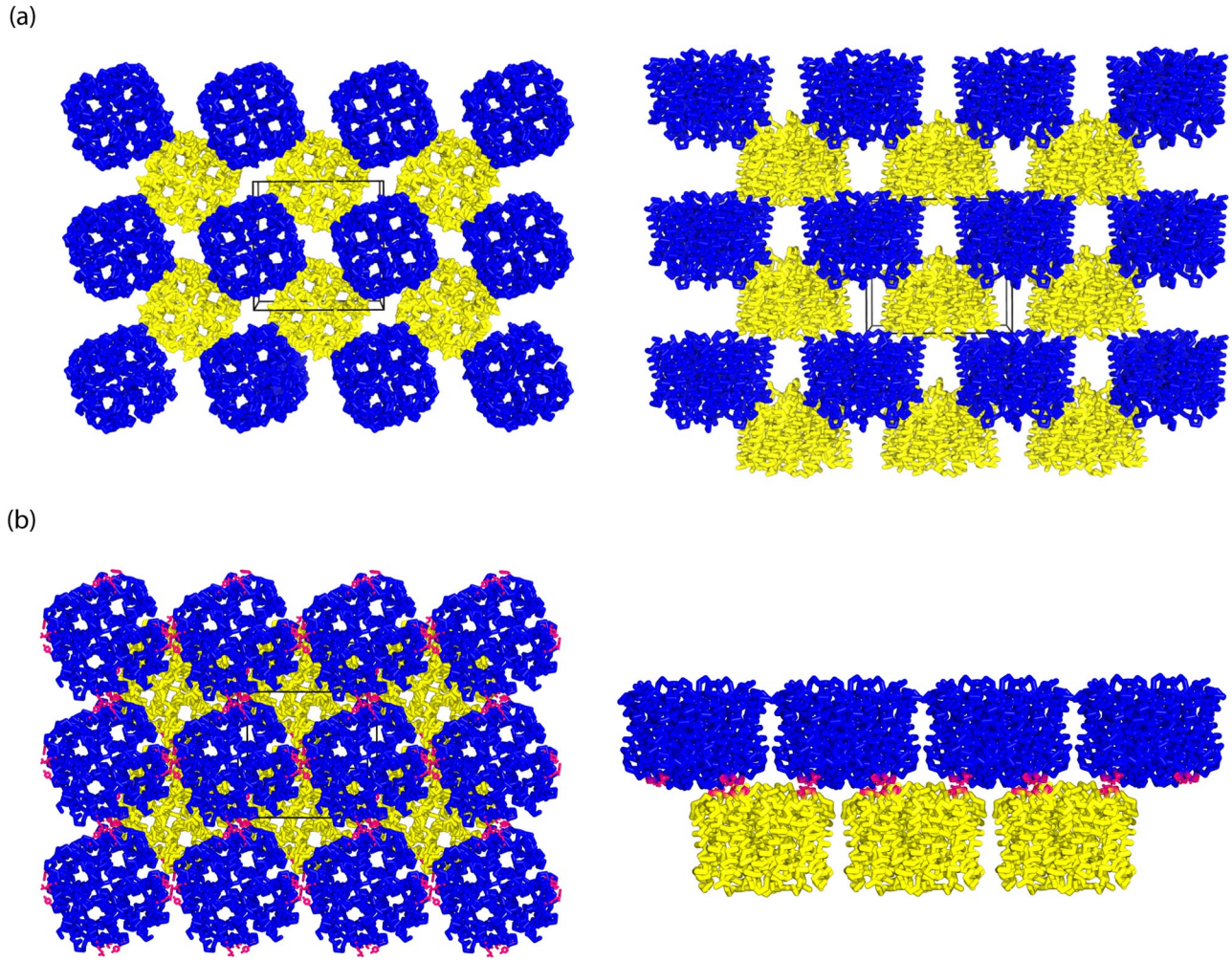


Fig. S2. Crystal lattice packing of (*a*) hAQP4 (space group = $P42_12$, a = 82.1Å, c = 76.4 Å) and (*b*) rAQP4 (space group = $P42_12$, a = 69 Å). (*Left*) Viewing the packing down the 4-fold axis. (*Right*) Viewing the packing from the side. Tetramers that are not in the same horizontal plane are in different colors. Unit cell is shown in black. (*b*) R108, G157, W231, I239, and Y250 are highlighted in pink, showing the interaction of the tetramers in the same plane (*Left*). (*b*) P139 and V142 from the rAQP4 3_{10} helix are highlighted in pink, showing that they are involved in the formation of the rAQP4 2D 2-layered crystal (*Right*) (12).

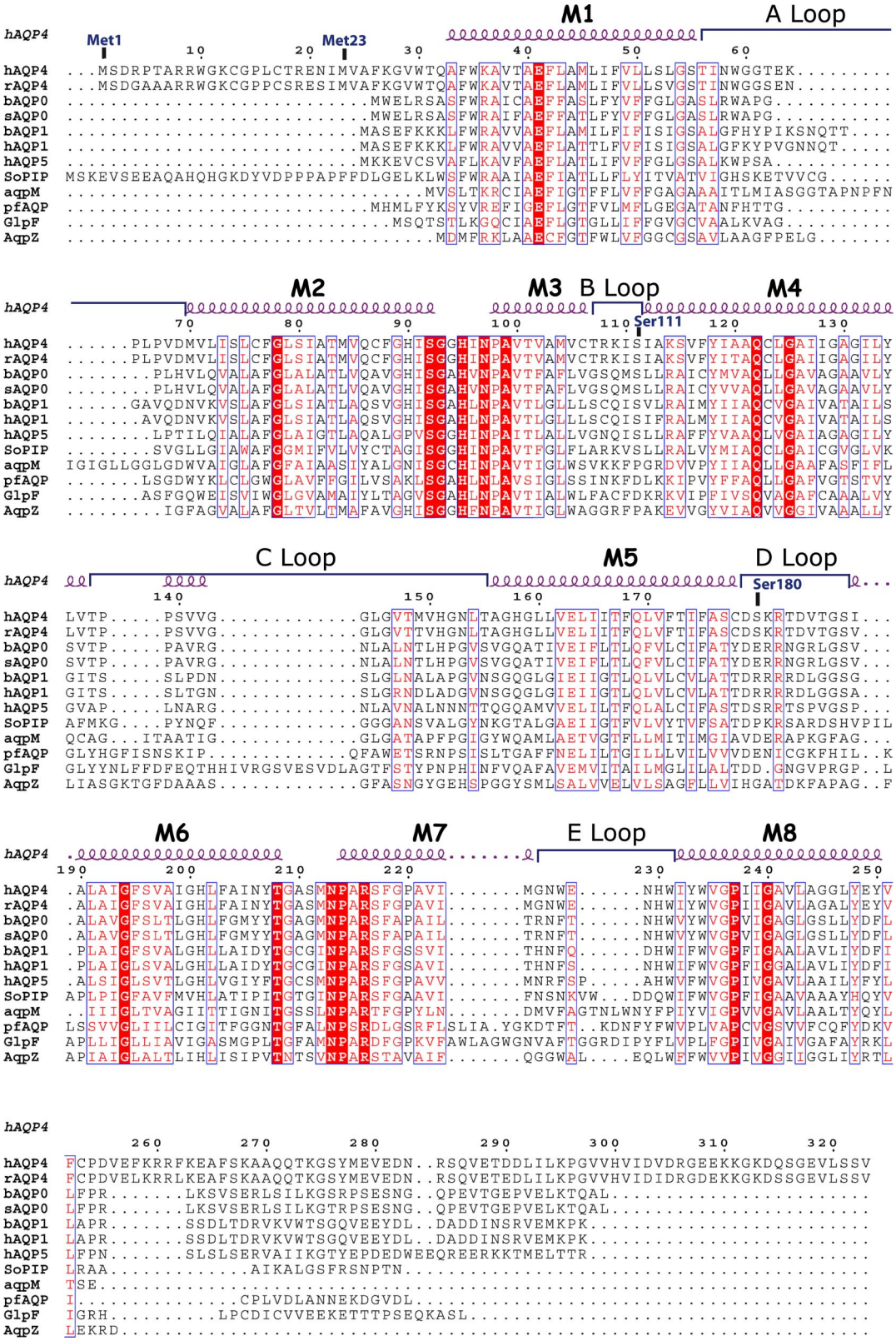


Fig. S3. Protein sequence alignment of all of the aquaporin structures solved to date: human and rat AQP4, bovine and sheep AQP0, bovine and human AQP1, human AQP5, spinach AQP SoPIP2;1, archeal AQP AqpM, *Plasmodium falciparum* PfAQP, *E. coli* GlpF, and AqpZ. Transmembrane helices and loop regions are defined. Initiating methionine for M1 and M23 isoforms of AQP4 are labeled. Ser-111 and Ser-180 are potential phosphorylation sites that affect gating for AQP4. The C-terminal SSV of AQP4 is the ligand that recruits α-syntrophin.

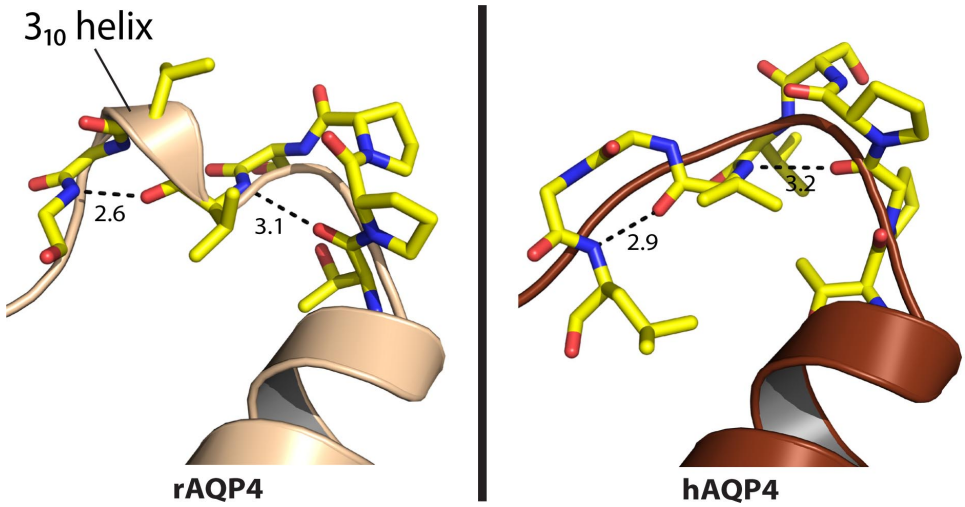


Fig. S4. Comparison of the C loop between rat AQP4 (light brown) and human AQP4 (dark brown). Notice the 3_{10} helix is missing in human AQP4.

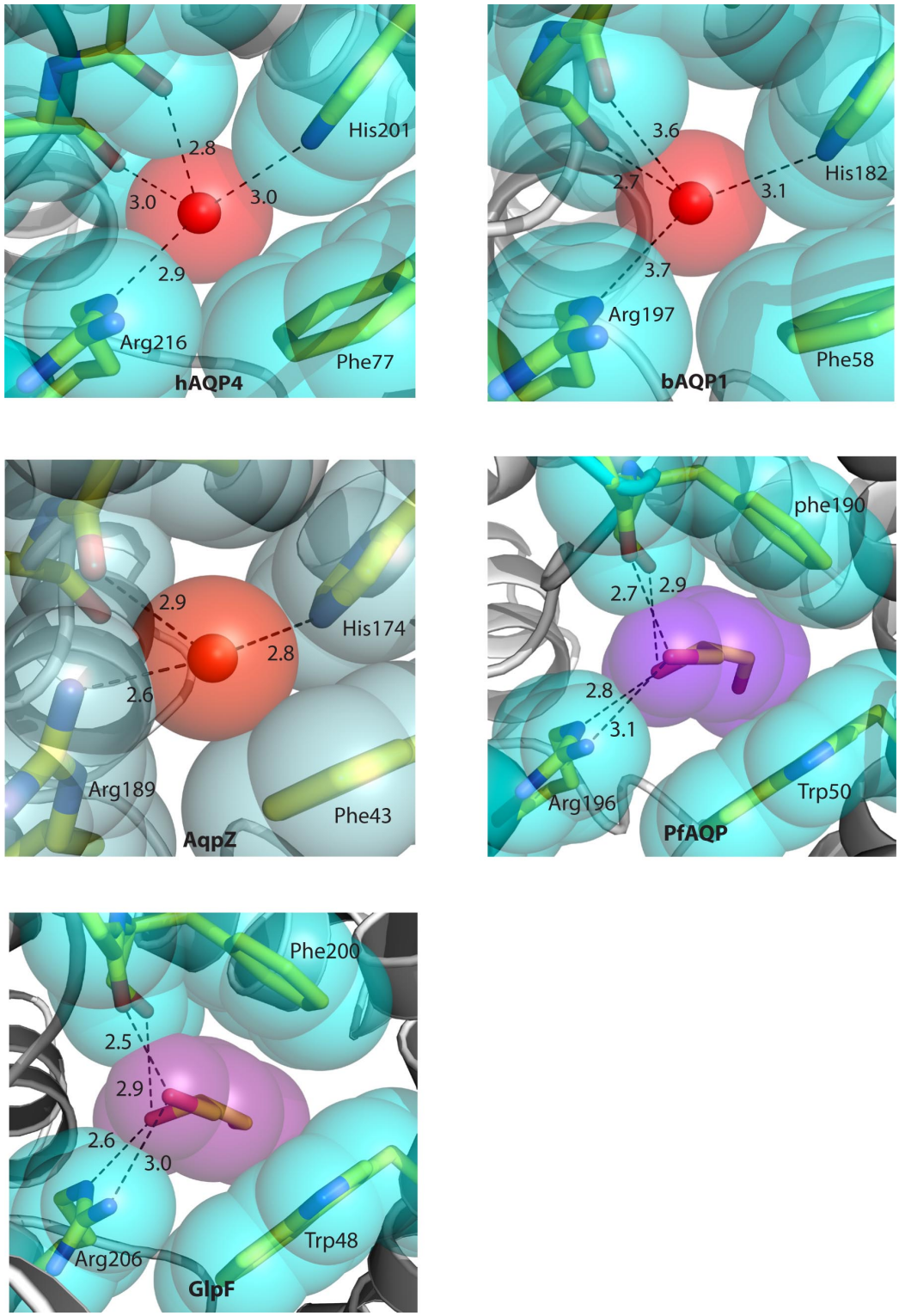


Fig. S5. Sphere representations of the van der Waals contact of the selectivity filter (SF) residues of hAQP4, bAQP1, AqpZ, PfAQP, and GlpF, viewing from the extracellular side down the conducting pore.

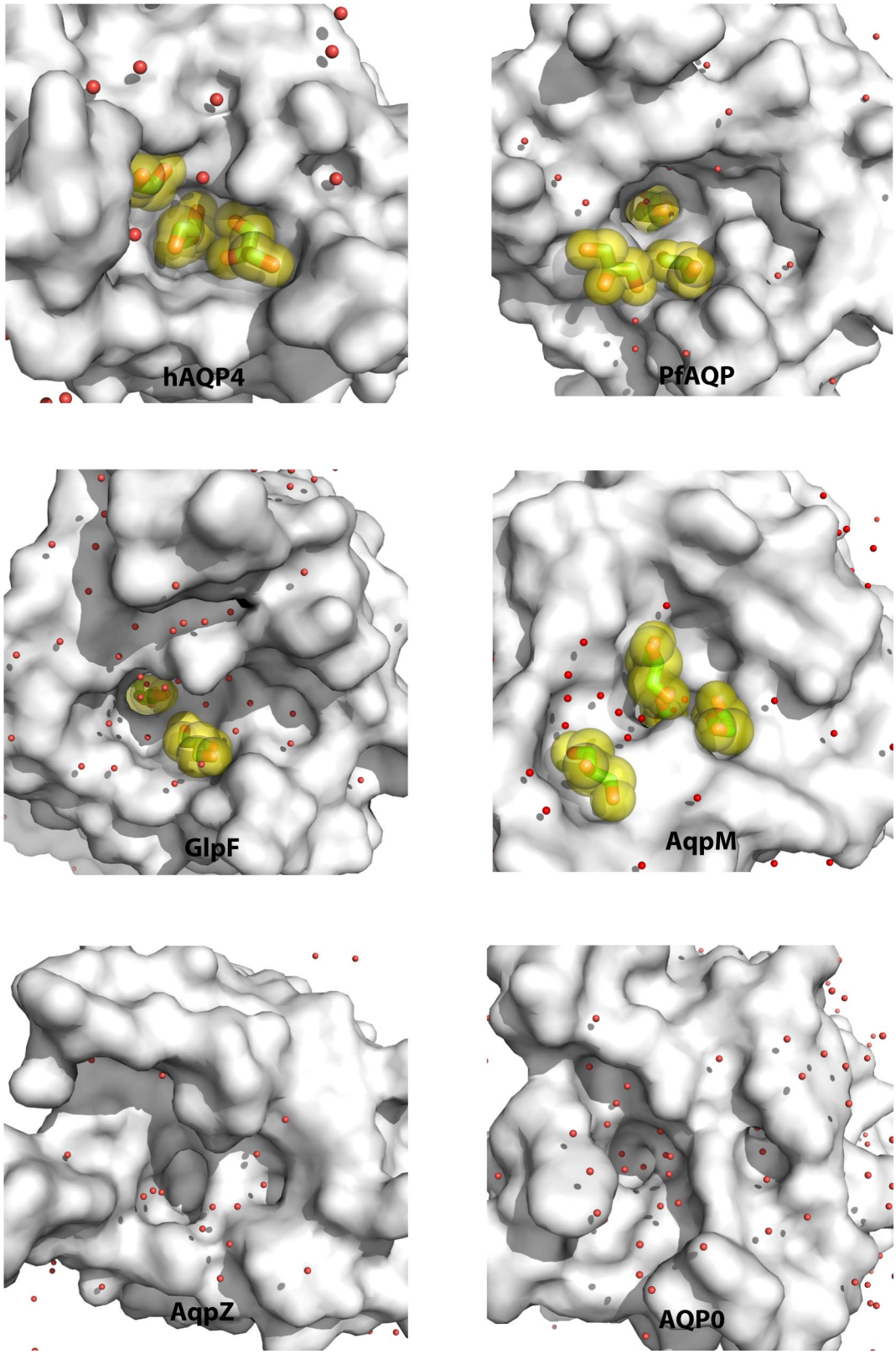


Fig. S6. Comparison of the surfaces of the extracellular vestibule of human AQP4, PfAQP, *E. coli* GlpF, *E. coli* AqpZ, Archaeal AqpM, and bovine AQP0. All are in the same orientation. Glycerol molecules are shown as green sticks with yellow surfaces. Water molecules are shown as red spheres with gray shadows.

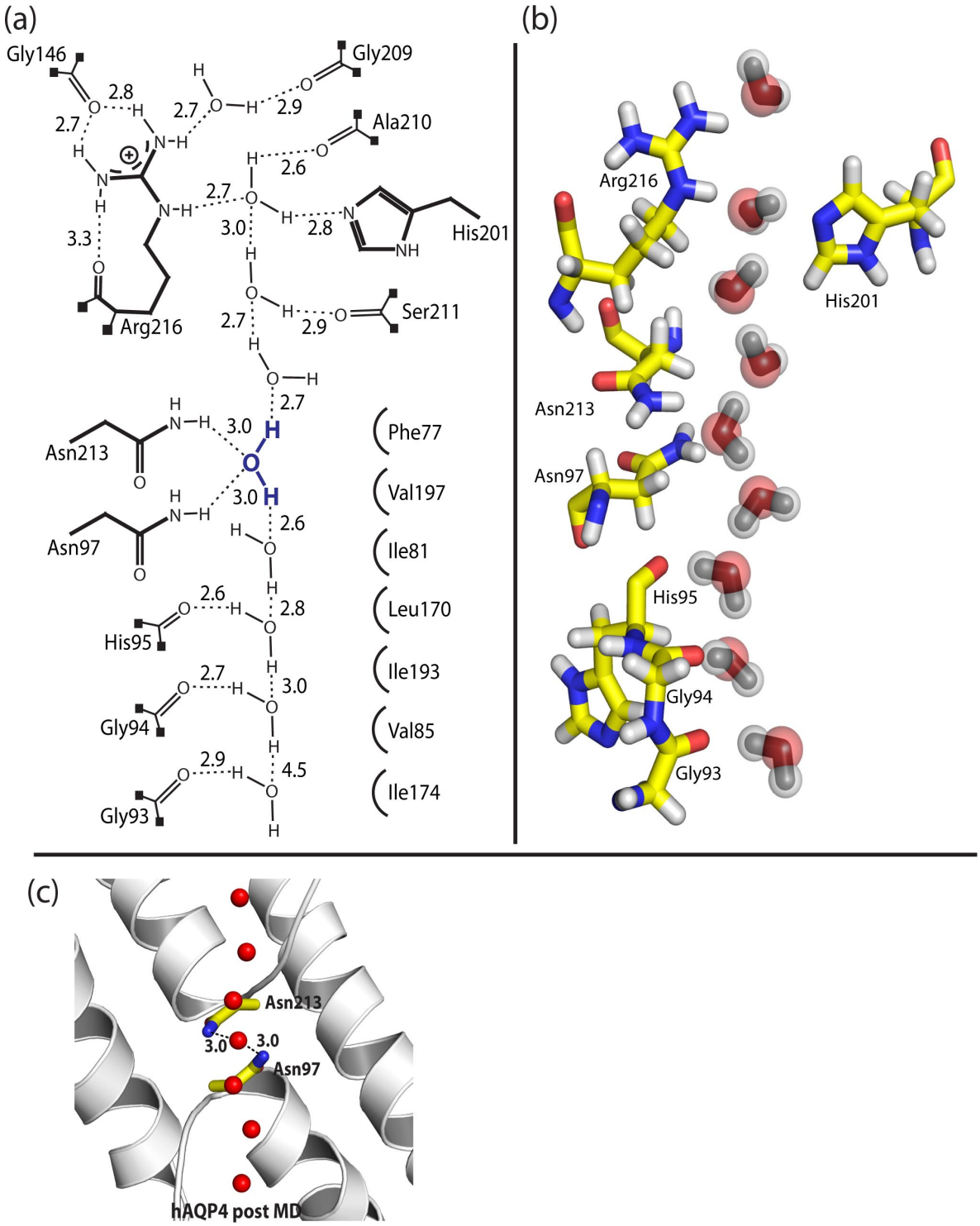


Fig. S7. AQP4: post MD simulation from the third experiment (see *Discussion*) (*a*) Schematic representation of the hydrogen bonding network through the channels. The distances are between heavy-atom to heavy-atom. (*b*) Stick representation of the conducting channel with hydrogen atoms from the simulation. (*c*) Another view of the NPA motifs and the central water.

Table S1. Summary of TLS refinement trials

TLS Groups	R_{fact}	R_{free}	FOM	LLG	rmsBOND	rmsANGLE	rmsCHIRAL
1	0.181	0.194	0.898	118283.2	0.01	1.217	0.081
2	0.18	0.191	0.9	117942.1	0.01	1.216	0.082
3	0.178	0.189	0.902	117673.9	0.01	1.209	0.082
4	0.178	0.189	0.904	117552.8	0.01	1.201	0.081
5	0.178	0.19	0.903	117593	0.01	1.206	0.081
6	0.177	0.189	0.905	117338	0.01	1.189	0.08
7	0.177	0.188	0.905	117267.7	0.01	1.183	0.079
8	0.176	0.187	0.906	117085	0.01	1.171	0.079
9	0.176	0.187	0.906	117078	0.01	1.166	0.078
10	0.175	0.186	0.906	116935.9	0.01	1.17	0.078
11	0.175	0.188	0.905	116933	0.01	1.173	0.078
12	0.174	0.188	0.906	116838.4	0.01	1.174	0.079
13	0.173	0.186	0.909	116610.8	0.01	1.172	0.078
14	0.173	0.185	0.909	116547	0.01	1.174	0.078
15	0.173	0.186	0.91	116453.9	0.011	1.182	0.079
16	0.173	0.186	0.91	116440.4	0.01	1.178	0.079
17	0.173	0.186	0.91	116440.4	0.01	1.178	0.079
18	0.172	0.184	0.911	116299.9	0.011	1.177	0.079
19	0.172	0.184	0.911	116269.7	0.011	1.176	0.078
20	0.172	0.184	0.91	116285.3	0.011	1.17	0.078
19 + H ride	0.164	0.174	0.922	114509	0.008	1.141	0.068
20 + H ride	0.166	0.174	0.92	114859.7	0.008	1.149	0.068

Different numbers of TLS groups were tried, and the best result was with 19 TLS groups (see Table S2). Statistics were taken from the Refmac5 refinement output log. The TLS server (http://skuld.bmsc.washington.edu/~tlsmd/) was used to generate the different numbers of TLS groups. For comparison purpose, riding hydrogen was applied to refinement with 19 and 20 TLS groups.

Table S2. The 19 TLS groups used for refinement

Group no.	Residue range
1	Q32 to K36
2	A37 to I47
3	F48 to W59
4	G60 to P65
5	L66 to V71
6	L72 to S92
7	G93 to V102
8	A103 to S115
9	V116 to I132
10	L133 to T148
11	M149 to L154
12	T155 to V171
13	F172 to S177
14	C178 to V185
15	T186 to V197
16	A198 to A210
17	S211 to I232
18	Y233 to Y248
19	E249 to P254