Supporting Information

Acharya et al. 10.1073/pnas.1007071107

SI Text

SI Methods. ***Peptide synthesis.*** The two sequences of the A/M2 25-46 wild type and G34A peptides (M2TM and M2TM′, 4-bromobenzoyl-PLVVAASIIAILHLILWILDRL-CONH2) and the corresponding wild-type peptide with Gly at position 34 (designated here M2TM). Both peptides were synthesized using Fmoc chemistry on an Applied Biosystems 433A and a Protein Technology Symphony synthesizer as previously described (1). 4-bromobenzoic acid (Sigma Aldrich) was coupled to the amino terminus of the peptide on resin in N,N-dimethylformamide, with HATU activation. Products were cleaved from the resin and purified as previously described (1). Purity was verified with analytical reverse phase high-performance liquid chromatography (RPHPLC) and matrix assisted laser disorption-ionization/time of flight (MALDI-TOF) mass spectrometry.

Data collection. Data sets for M2TM′ and M2TM were collected from several different crystals, with cryocooling to 100 K during data collection. These include multi-wavelength anomalous dispersion (MAD) datasets recorded at synchrotron beam lines (NSLS) and datasets collected at a home source. Most M2TM′ crystals featured a diffraction resolution between 2.6 Å and 1.65 Å (the resolution of the crystal used for structure refinement). The best M2TM dataset could only be processed up to ~68% before the crystal degraded.

Diffraction images were indexed, integrated using MOSFLM (2, 3) and scaled using SCALA (4). Table S2 shows data collection statistics of the two highest resolution M2TM′ datasets (the one at 1.65 Å resolution, used to determine the M2TM′ structure, and an additional one at 1.85 Å), and of the incomplete M2TM dataset.

Similarity tests between the M2TM and M2TM′ structures. The partial M2TM dataset was subject to molecular replacement, using the backbone of the refined M2TM′ structure to build a polyalanine model. A satisfactory solution [$R_{cryst} = 0.41$, $CC = 0.65$ (polyala)] was obtained, resulting in a map of good quality. Clear density from sidechains not included in the polyalanine molecular replacement model was observed. For example, His37 side chains were observed in the same positions as in the 1.65 Å M2TM′ density map.

We also tried another molecular replacement approach, based on the entire M2TM′ refined structure deprived of the Ala34 methyl group and of all the solvent molecules. Rigid body refinement gave a solution with $R_{cryst}/R_{free} = 0.299/0.333$. Furthermore, density that could be assigned to water molecules was observed in the pore in the vicinity of His37, at the same locations as in the M2TM′ structure. These findings strongly suggest that the overall structures are similar, although a M2TM structure can not be refined ab initio from the incomplete M2TM dataset.

Structure solution, model building, and refinement of M2TM′. Initial attempts to determine the structure by experimental phasing [single-wavelength anomalous disperion (SAD), MAD] were not successful, due to the disorder of the bromine atom positions. Molecular replacement using a single alpha-helix or tetramers from a previously solved structure (PDB code 3BKD) (1) as model probes did not yield a solution. By comparison of cell dimensions, it was intuitively obvious that helices forming the tetramers are more closely packed in the present datasets than they were in 3BKD. Whereas in 3BKD (with cell dimensions $a = 38.753$ Å, $b = 56.557$ Å, and $c = 56.009$ Å) helices are oriented along the **a** crystallographic axis, and two tetramers lie in the **bc** plane, in the present datasets the tetramers lie in the **ac** plane with helices oriented along the **b** axis. Based on this packing knowledge, tetramer models were generated to use as model probes in molecular replacement.

Tetramer model generation. Tetramer models were constructed based on one of the helices of our previously published 2.0 Å resolution structure (3BKD, chain "A"). The tetramers were generated with with various helical orientations; X_{angle} ranged from −30° to −20°, Y_{angle} from 7° to 10°, Z_{angle} from 160° to 190°, and bundle radius from 6.8 to 8.0 Å (X_{angle}, Y_{angle}, and Z_{angle} represents the angle formed by the helix with the X-axis, Y-axis, and Z-axis respectively). During tetramer generation, the helix oriented along the Z-axis was first rotated around the Z-axis (Z_{angle}), then about the X-axis (X_{angle}) and then about the Y-axis (Y_{angle}). Four copies of this oriented helix were created, from which a bundle was constructed by applying four-fold symmetry along the Z-axis and translating each helix to set the bundle radius.

Among tetramer models generated as described above, only those models where Val27 was oriented similarly to that in the 3BKD structure were considered for molecular replacement calculations. The molecular replacement calculations were performed using the program PHASER (5, 6) for data in the resolution range 15–4.0 Å. The datasets used for the initial phasing were processed in the space group $P2_1$ with cell parameters $a = 46.41$ Å, $b = 48.52$ Å, $c = 46.27$ Å and $\beta = 116.97°$ and were collected to a resolution limit of 2.6 Å at our home source. The Matthews coefficient (7) ($V_m = 2.3$ Å^3/Da) indicated the presence of two tetramers in the asymmetric unit. Initially molecular replacement calculations were performed using bundles with parameters: bundle radius = 8.0 Å – 7.5 Å, $X_{angle} = -26°$, $Y_{angle} = 10°$, and $Z_{angle} = 160°$–$190°$. Later, other models were randomly selected for the calculations. In evaluating the quality of the models, weight was given to recognizing an interpretable electron density map. In general, the resulting maps were poor, but one model ($X_{angle} = -24°$, $Y_{angle} = 10°$, $Z_{angle} = 167°$, and radius = 7.0 Å) showed the very clear density of a dimer. The dimer was used as a model and two tetramers were located in the asymmetric unit, resulting in an excellent electron density map.

Of several crystals tested for diffraction at the synchrotron beam line (NSLS, X6A), a crystal belonging to the space group $C222_1$ with the cell parameters $a = 48.67$ Å $b = 79.09$ Å and $c = 48.56$ Å was found to diffract to a maximum resolution of 1.65 Å. The Matthews coefficient ($V_m = 2.4$ Å^3/Da) suggested the possibility of a tetramer in the asymmetric unit, thus the structure was solved by molecular replacement using one of the above-determined tetramers as a model. Iterative refinement and model building were carried out, during which all side chains could be traced in the electron density map. As the refinement progressed water molecules were located in the map. When the refinement converged, the model was refined to $R_{work} = 19.6\%$ and $R_{free} = 20.5\%$ (Table S2), with all the residues in the allowed regions (100% favorable) of the Ramachandran plot. Refinement statistics are shown in Table S2. The refinement was carried out using the CCP4 program suite (8), REFMAC (9), and all model building was done using COOT (10).

In general, solvent molecules were included in the structure only if they were visible at the 3-sigma level in an $F_0 - F_C$ map. All atoms were found in full occupancy with the exception of the Br atoms in the bromobenzoyl group, which were disor-

dered, presumably due to radiation damage and the dimer of water molecules in the outer cluster, which had two alternate conformations that refined with occupancies of 0.6 and 0.4.

Backbone flexibility analysis. Fig. 2*B* was created by the following analysis: Debye–Waller (temperature or B) factors were analyzed using several steps of normalization. First, to focus on the helical movements as well as avoid fluctuations resulting from crystal contacts, only backbone atoms (C, O, Cα, N) were included in the analysis. Thus, hereafter, "all" B-factors corresponds only to the backbone B-factor subset. Second, the standard B-factor normalization was applied—i.e. the difference between each B-factor and the mean of all B-factors was divided by the standard deviation of the B-factors. Third, because due to helix unwinding the B-factors at the helix outer turns are far larger compared to the B-factors at the rest of the helix, the normalized B-factors were percent ranked. Finally, to increase the signal to noise ratio the four B-factors of each residue (same sequence position on each helix) were averaged over the four subunits and the result was placed in the C_α line to enable easy colouring by C_α B-factors as applied in PyMol (11).

To validate the results, the analysis was conducted on two independent sets of B-factors from a pair of similar M2TM′ crystals —those from the 1.65 Å resolution structure and those from the 1.85 Å resolution structure. The independently determined flexibility distributions were strikingly similar.

QM optimization for NMR calculations. Cluster models (as described in the text) were optimized for four protonation states were generated (2_δ, 2_ε, 3_δ, and 3_ε) using the CP2K electronic structure package (12). The optimizations were performed at the density functional theory (DFT) level of theory in the Gaussian plane wave (GPW) approximation (13). Molecular optimized triple-zeta basis sets with two polarization functions (14), Goedecker–Teter–Hutter pseudopotentials (15), and the Becke exchange and Lee–Yang–Parr correlation (BLYP) functional (16, 17) were used. An energy cutoff of 360 Ry was employed in the plane wave representation of the density. A wavelet based Poisson solver (18) was used to remove the spurious interactions of system with its periodic images.

Classical molecular dynamics (MD) simulations details. Periodic boundary conditions were applied to all classical MD simulations, with the electrostatic potential being solved by the particle mesh Ewald (PME) method (19) with a real space spherical cutoff of 12 Å. Lennard–Jones interactions were cut off at 12 Å. The equations of motion were integrated using a timestep of 1.5 fs. The lengths of all bonds involving hydrogen atoms were kept constrained with the SHAKE algorithm (20). The system was run at 310 K temperature and 1 atm pressure using Langevin coupling schemes.

System Setup for Classical MD Simulations. We set up the three sets of simulations with the following details:

i. M2TM′ at different protonation states (0_ε, 1^+_ε, 2^+_ε, 2^+_δ, 3^+_ε, 3^+_δ, and 4^+) was solvated in a $54 \times 54 \times 64$ Å^3 water box. Heavy atoms of the peptide backbone were kept restrained at all times to their experimental positions in the crystal structure. Langevin temperature coupling was applied with a reference temperature of 310 K. Isotropic Langevin piston pressure coupling was applied with a reference pressure of 1 atm.
ii. M2TM′ at different protonation states (1^+_ε, 2^+_ε, 3^+_ε, and 4^+) was embedded into a 80×80 Å^2 bilayer of 1-palmitoyl-2-oleoylphosphatidylcholine lipid molecules, and solvated by adding a 18 Å thick water layer at each side (~10000 water molecules in total) with 150 mM of KCl dissolved at random initial positions. Each system was gently brought to equilibrium by restraining the heavy atoms of the protein backbone and side chains to their initial positions, using, as initial force constants, 20 and 5 kcal/mol/Å^2, for backbone and side-chains, respectively. Restraining potentials were progressively reduced in steps of 1 ns down to 0.01 kcal/mol/Å^2, and completely released after 10 ns. Temperature and pressure coupling were applied as in (*i*), but only the dimension of the periodic box perpendicular to the membrane was allowed to vary under the pressure coupling scheme.
iii. WT (M2TM), G34A (M2TM′), and G34V peptides, modeled based on the crystal structure, were embedded in the already equilibrated lipid bilayer as in (*ii*), followed by a short 1 ns equilibration with position restraints. Density profiles of water oxygens were calculated on the following 15 ns trajectory, using a bin-width of 0.25 Å.

NMR shielding calculations. We performed NMR shielding calculations on the nitrogen atoms of the four histidines using NWChem (21) at the DFT-BLYP level of theory (16, 17) on a gas phase cluster model of the M2 His-box region. We included in these calculations the 4 His37 side chains up to the C_α atom and the 6-water outer cluster, the 4 Ala34-Ile35 peptide groups, and the bridging water cluster between the His-box and the Trp-basket. We added a ninth water molecule above the bridging cluster, observed during all of the classical MD simulations. Four protonation states were generated (2_δ, 2_ε, 3_δ, and 3_δ) and optimized using CP2K (12). To compute the magnetic shielding tensors of all the nuclei in the imidazole moieties, we employed the London orbitals (or gaugeinvariant atomic orbitals, GIAO) approach (22, 23) implemented in NWChem.

SI Discussion. ***NMR chemical shifts.*** Table S1 shows the isotropic chemical shifts (in ppm) of the nitrogen atoms, referenced to NH_3 and obtained using the theoretical-experimental calibration derived by in ref. 24 across a wide range of chemical compounds. In each state, we averaged the chemical shifts across chemically equivalent nuclei (for example, in the 2_ε state, we averaged the two protonated N_δ shieldings).

M2 peptide length. The various experimental and theoretical studies of M2 presented in this paper and elsewhere have used different segments of the M2 protein. The UDorn sequence of the A/M2 protein was considered as the wild-type in all studies. A peptide containing residues 22–46 has been shown to be active and have transport properties similar to the full length M2 (25). For the sake of clarity, we list here the different lengths of M2 peptides used in our study and related studies. The 25-46 peptide of the wild-type protein (M2TM), of the G34A mutant (M2TM′) and of the G34V mutant (PLVVAASIIGILHLILWILDRL, PLVVAASIIAILHLILWILDRL and PLVVAASIIVILHLILWILDRL respectively) were used in the present crystal structure and simulations. The M2TM peptide was in the previous high resolution crystal structure (1). The oocyte electrophysiologic results presented here were conducted with the fulllength M2 wild-type protein and G34A and G34V mutants. The NMR experimental results of Hu et al. were run on the same wild-type peptide (22–46) that was recently shown to be functional (26). The NMR structure solved by Schnell et al. was for a yet longer wild-type peptide spanning residues 18–60.

Comparison to other channels. It is instructive to compare the structure and mechanism of M2 to the gating transitions of cation channels as well as the conduction mechanism of proton pumps. M2 bears several similarities to the K^+ and Na/K channels (27): Both possess a water-filled aqueous pore near the center of the bilayer, which helps provide a water-like environment to stabilize ions near the center of the membrane. Furthermore, both M2 and

cation channels feature a dilating hydrophobic valve or gate that opens and closes as a result of TM helix motions (28–30). This motion appears coupled to the binding of ions to a selectivity filter (31) (His-box in M2 and carbonyllined pore in cation channels) (32, 33) whose affinity is tuned to allow relatively rapid passage of ions at their physiological concentration. The first two protons bound by M2 are held with a very high affinity (pK_a approx. 8), contributing to ion-selectivity and priming the channel for binding a third permeant proton with a pK_a near 6 (26). By matching the pK_a of the third proton with the pH of an acidifying endosome the proton is bound sufficiently well to facilitate diffusion, but not so tightly as to fall into a deep well (too tight binding) or overcome a barrier (too weak binding) that would decrease flux. Similarly, the first K^+ ion bound to a K^+ channel (34) primes the selectivity filter for binding additional, permeant ions with an overall binding constant appropriate for reasonable flux while maintaining a high degree of selectivity.

It is also interesting to compare the proton translocation pathway in M2 with bacteriorhodopsin, which uses light to achieve proton transport against a concentration gradient. Its proton release pathway features two intraprotein water clusters that are anchored and interconnected by polar sidechains (35, 36). As proposed for M2, one of the 6 water clusters stabilizes a protonated water molecule via an extended network of water-sidechain interactions. Moreover, the geometry and strengths of the H-bonds appear optimized to mediate vectorial proton transfer through the protein matrix following photoisomerization of its retinal cofactor (37).

1. Stouffer AL, et al. (2008) Structural basis for the function and inhibition of an influenza virus proton channel. *Nature* 451:596–599.
2. Powell HR (1999) The Rossmann Fourier autoindexing algorithm in MOSFLM. *Acta Crystallogr* D55:1690–1695.
3. Leslie AGW (1999) Integration of macromolecular diffraction data. *Acta Crystallogr.* D55:1696–1702.
4. Evans P (2006) Scaling and assessment of data quality. *Acta Crystallogr* D62:72–82.
5. Read RJ (2001) Pushing the boundaries of molecular replacement with maximum likelihood. *Acta Crystallogr* D57:1373–1382.
6. McCoy AJ, Grosse-Kunstleve RW, Storoni LC, Read RJ (2005) Likelihood-enhanced fast translation functions. *Acta Crystallogr* D61:458–464.
7. Matthews BW (1968) Solvent content of protein crystals. *J Mol Biol* 33:491–497.
8. Collaborative Computational Project N (1994) The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr* D50:760–763.
9. Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr* D53:240–255.
10. Emsley P, Cowtan K (2004) COOT: Model-building tools for molecular graphics. *Acta Crystallogr* D60:2126–2132.
11. DeLano WL (2002) The PyMOL Molecular Graphics System. (http://www.pymol.org).
12. The CP2K developers group (2008) (http://cp2k.berlios.de/).
13. Lippert G, Hutter J, Parrinello M (1997) A hybrid Gaussian and plane wave density functional scheme. *Mol Phys* 92:477–488.
14. VandeVondele J, Hutter J (2007) Gaussian basis sets for accurate calculations on molecular systems in gas and condensed phases. *J Chem Phys* 127:114105.
15. Goedecker S, Teter M, Hutter J (1966) Seperable dual-space Gaussian pseudopotentials. *Phys Rev B* 54:1703–1710.
16. Becke AD (1988) Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys Rev A* 38:3098–3100.
17. Lee C, Yang W, Parr RG (1988) Development of the Colle-Salvetti correlation energy formula into a functional of the electron density. *Phys Rev B* 37:785–789.
18. Genovese L, Deutsch T, Goedecker S (2007) Efficient and accurate threedimensional Poisson solver for surface problems. *J Chem Phys* 127:054704.
19. Darden T, York D, Pedersen L (1993) Particle mesh Ewald—An Nlog(N) method for Ewald sums in large systems. *J Chem Phys* 98:10089–10092.
20. Ryckaert J-P, Ciccotti G, Berendsen HJC (1977) Numerical integration of the cartesian equations of motion of a system with constraints: Molecular dynamics of n-alkanes. *J Comput Phys* 23:327.
21. Bylaska EJ, et al. (2007) *NWChem* (Pacific Northwest National Laboratory, Richland, WA).
22. Ditchfield R (1974) Self-consistent perturbation theory of diamagnetism. I. A Gauge-invariant LCAO (Linear combination of atomic orbitals) method for NMR chemical shifts. *Mol Phys* 27:789–807.
23. Wolinski K, Hinton JF, Pulay P (1990) Efficient implementation of the gauge-independent atomic orbital ethod for NMR chemical shift calculations. *J Am Chem Soc* 112:8251–8260.
24. Dokalik A, Kalchhauser H, Mikenda W, Schweng G (1999) NMR spectra of nitrogen-containing compounds. Correlations between experimental and GIAO calculated data. *Magn Reson Chem* 37:895–902.
25. Ma C, et al. (2009) Identification of the functional core of the influenza A virus A/M2 proton-selective ion channel. *Proc Natl Acad Sci USA* 106:12283–12288.
26. Hu J, et al. (2006) Histidines, heart of the hydrogen ion channel from influenza A virus: toward an understanding of conductance and proton selectivity. *Proc Natl Acad Sci USA* 103:6865–6870.
27. Doyle DA, et al. (1998) The structure of the potassium channel: molecular basis of K^+ conduction and selectivity. *Science* 280:69.
28. Alam A, Jiang Y (2009) Structural analysis of ion selectivity in the NaK channel. *Nat Struct Mol Biol* 16:35–41.
29. Bocquet N, et al. (2009) X-ray structure of a pentameric ligand-gated ion channel in an apparently open conformation. *Nature* 457:111–114.
30. Hilf RJ, Dutzler R (2009) Structure of a potentially open state of a protonactivated pentameric ligand-gated ion channel. *Nature* 457:115–118.
31. Panyi G, Deutsch C (2006) Cross talk between activation and slow inactivation gates of Shaker potassium channels. *J Gen Physiol* 128:547–559.
32. Morais-Cabral JH, Zhou Y, MacKinnon R (2001) Energetic optimization of ion conduction rate by the K^+ selectivity filter. *Nature* 414:37–42.
33. Alam A, Jiang Y (2009) Structural analysis of ion selectivity in the NaK channel. *Nat Struct Mol Biol* 16:35–41.
34. Ray EC, Deutsch C (2006) A trapped intracellular cation modulates K^+ channel recovery from slow inactivation. *J Gen Physiol* 128:203–217.
35. Lanyi JK (2004) Bacteriorhodopsin. *Annu Rev Physiol* 66:665–688.
36. Garczarek F, Gerwert K (2006) Functional waters in intraprotein proton transfer monitored by FTIR difference spectroscopy. *Nature* 439:109–112.
37. Wolf S, Freier E, Gerwert K (2008) How does a membrane protein achieve a vectorial proton transfer via water molecules? *Chemphyschem* 9:2772–2778.

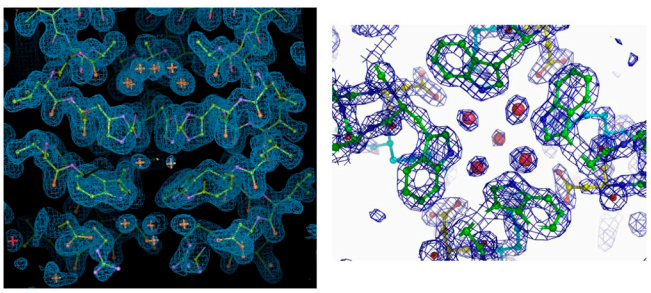

Fig. S1. (*Left*) $2F_0 - F_C$ map contoured at 1.1σ, showing density from solvent molecules above and below the His37 residues. (*Right*) $2F_0 - F_C$ map showing the Trp-basket and the exit water cluster. The protein atoms are contoured at 1.5σ. The density from the solvent molecules is contoured at 1.1σ. Note the poorer density for the middle water, which refined with a higher β-factor (53 Å^2).

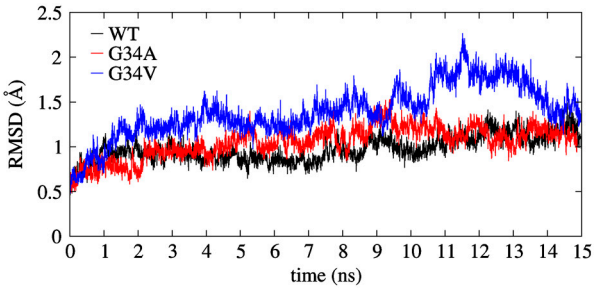

Fig. S2. Root mean square deviation (RMSD, Å) of the C_α atoms with respect to the X-ray structure is plotted as a function of time (ns) for WT (black), G34A (red), and G34V (blue).

Table S1. Chemical shifts of imidazole nitrogen atoms for different protonation states (ppm units, referred to ammonia)

Atom \ state	2_ε	3_ε	2_δ	3_δ	pH 6.5 (22)
N_δ (prot.)	156.50	148.87	148.52	146.08	147, 156
N_δ (unprot.)	226.63	226.76	—	—	(230)
N_ε (prot.)	146.18	153.64	167.40	159.28	144, 153
N_ε (unprot.)	—	—	238.93	229.82	—

Table S2. Data collection, processing, and refinement statistics for M2TM′ and M2TM

X-ray refinement statistics			
Crystal peptide	M2TM'G34A (1.65 Å)	M2TM′G34A (1.85 Å)	M2TM′G34 (partial dataset)
Space group	$C222_1$	$C222_1$	$C222_1$
Cell dimensions: a, b, c (Å)	48.67, 79.09, 48.56	48.36, 48.80, 48.39	48.74, 77.86, 48.61
Resolution (Å)	48.6–1.65 (1.75–1.65)*	48.39–1.85 (1.95–1.85)*	41.3–2.50 (2.64–2.5)*
R_{work}/R_{free}	0.196/0.205	0.197/0.237	0.299/0.333
Number of atoms			
Protein	732†	732†	728†
Ligand/ion	31	0	0
Water	25	28	0
B-factors			
Protein	13.7	12.3	13.2
Ligand/ion	48.4		
Water	24.9	25.2	
Rmsd			
Bond lengths (Å)	0.007	0.008	
Bond angles (°)	0.88	0.88	

*Highest resolution shell is shown in parenthesis.
†Including Bromo benzoyl group.