Supporting Information for

Proton Induced Conformational and Hydration Dynamics in the Influenza A M2 Channel

Laura C. Watkins,[1] Ruibin Liang,[1] Jessica M.J. Swanson,[1,*] William F. DeGrado,[2] Gregory A. Voth[1,*]

[1]Department of Chemistry, Institute for Biophysical Dynamics and James Franck Institute, The University of Chicago, Chicago, IL 60637, USA

[2]Department of Pharmaceutical Chemistry, University of California, San Francisco, CA 94158, USA

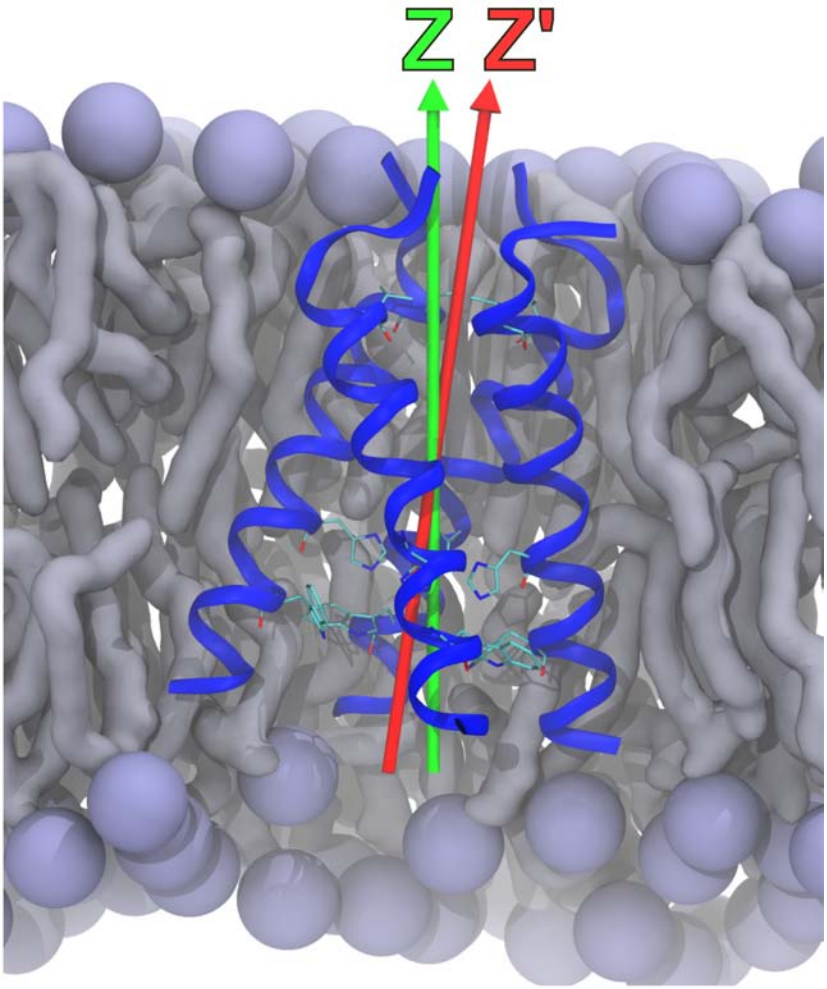

Figure S1. Illustration of the axis Z' used throughout this study. The coordinate Z (green arrow) is the same as the Z-axis of the simulation box. The system is setup such that the membrane normal aligns with Z. Z' is the principal axis of the protein (red arrow), calculated for each frame. This is used to redefine the reference frame in terms of the protein's central axis for analysis. The protein helices for an example configuration are shown in blue, and the pore lining residues Val27, His37, and Trp41 are additionally shown as sticks. The lipids are light gray-blue.

* jmswanson@uchicago.edu, *gavoth@uchicago.edu

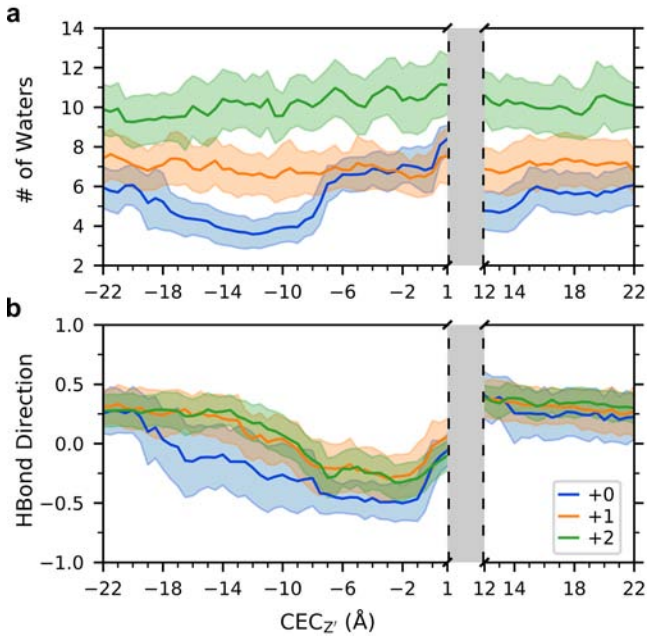

Figure S2. Analysis of water in the region above the His37 tetrad and below Gly34, see SI Methods for detailed definition. (a) Average and fluctuations (standard deviation) of the number of waters in this this region as a function of excess proton CEC position. (b) Average and standard deviation of the direction of water-water hydrogen bonds in this region as a function of CEC position.

Protein Sidechain Atoms Involved in Hydrogen Bonding

The sidechain atoms also play an important, but dynamic, role in stabilizing water structures through hydrogen bonding. As the hydrated excess proton moves through the channel, these interactions transiently shift both in strength and orientation. To understand the protein atoms' role in PT, we focus here on the sidechain hydroxyl group of Ser31, which is the only polar residue in the central portion of the channel.

The Ser31 hydroxyl oxygen atoms form hydrogen bonds as acceptors with water molecules (Figure S3a). The frequency of these bonds in the +1 and +2 His37 states is relatively constant as the proton moves through the channel, with +2 having a slightly higher frequency. The +0 state exhibits an increase in frequency when the proton is below the sidechain, near $CEC_{Z'}$= 0.0 Å, at which point all three states have the same frequency. Interestingly, the orientation of these hydrogen bonds is clearly altered as the proton passes through the Val27 gate. When the gate is closed, there is a water hydrogen bond with the Ser31 sidechain ~71, 76, and 78% of the time (averaged over the four helices) for the +0, +1, and +2 states, respectively. These water molecules form the top layer of water in the middle of the channel – one hydrogen bonded to each Ser31 sidechain in a ring-like structure, situated below Ser31 (Figure S3a, left panel). The opening of the gate, and subsequent formation of a continuous water wire to the top of the channel, disrupts and shifts the four water molecules such that they interact with Ser31 with an altered orientation as illustrated in the right panel of Figure S3a and calculated in Figure S3b. In the +0 state, all four tend to interact from a slightly higher position, whereas in the +1 and +2 state the change is predominantly in one water-Ser31 interaction, pointing to the asymmetry in the channel's response to PT. The Ser31 hydroxyl group also donates a strong hydrogen bond to the Val27 backbone carbonyls on the same chain, present ~95% of the time when the excess proton is out of the channel. These interactions are also impacted by the Val27 gate opening, which distorts the geometry and decreases the frequency of backbone hydrogen bonding. These Ser31 interactions indicate how residues can act as a scaffold for PT in M2.

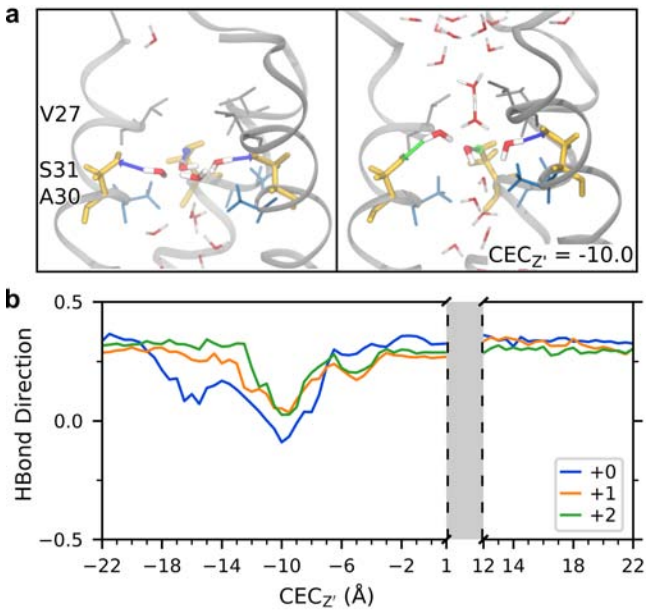

Figure S3. Protein-water hydrogen bonds affected by the CEC position. (a) Typical water structures in the +0 state when the excess proton is outside the channel (left) and above the Ser31 tetrad (right). Ser31 is shown in yellow, Ala30 in blue, hydrogen bonds are shown as arrows colored blue and green for bonding from below and above, respectively. Front helix does not have a Ser31-water hydrogen bond and is removed for clarity. (b) Ser31: Average direction of hydrogen bonds between water as donors and Ser31 sidechain oxygen atoms as acceptors.

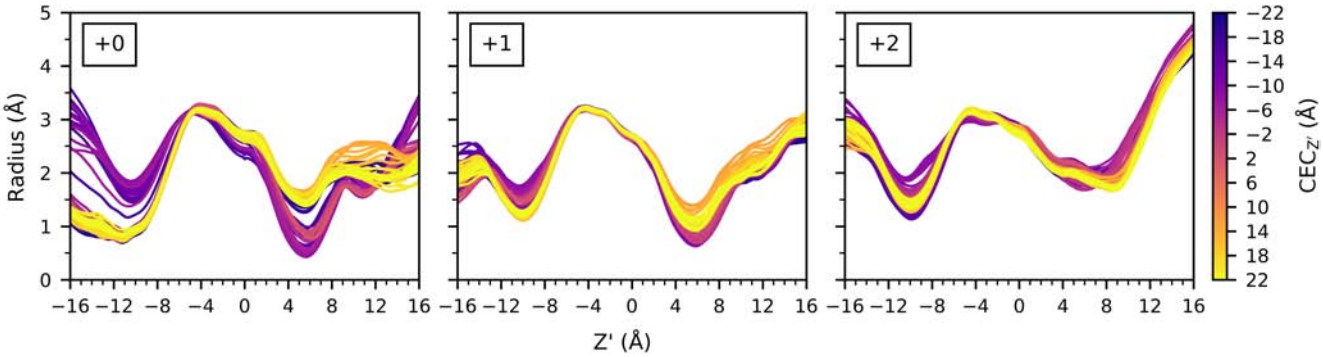

Figure S4. Radii profiles for each proton position as in Figure 5, but here separated by charge state and colored by excess proton position to highlight differences based on CEC$_{Z'}$ more clearly.

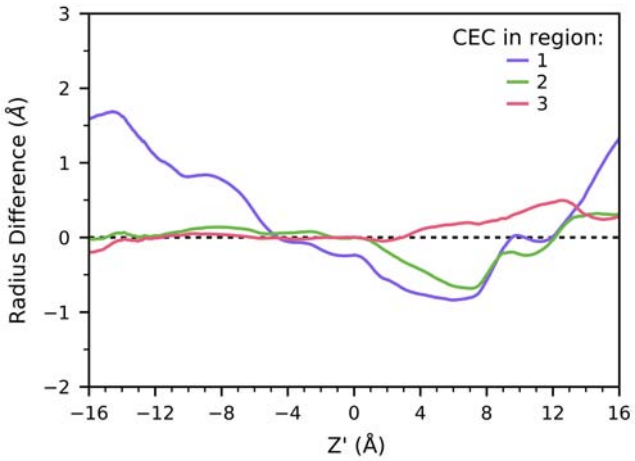

Figure S5. Difference plot of radii profiles for CEC in the three different regions of the +0 channel indicated in Figure 6. Region 1 (purple) corresponds to the range [-18.0, -7.0], from the top of the channel to just below Val27. Region 2 (green) corresponds to the range [-7.0,1.0], from below Val27 to just below Gly34. Region 3 (pink) corresponds to CEC$_{Z'}$ in the range [10.0, 18.0], from Trp41 to the bottom of the channel. The difference is calculated between the radius profile averaged over all CEC positions in a given region and the average radius profile when the CEC is out of the channel.

Channel Water Asymmetry

Table S1 shows the percent of frames, averaged over all proton positions, that have hydrogen bonds between water and the Gly34 backbone carboxyl oxygen of each helix. The disparity between the helices is highest in the +0 state – while helix C is hydrogen-bonded (74%) of the time, helix B only has such interactions (25%) of the time, indicating that the water structure is not four-fold symmetric in the dynamic system.

		Helix			
		A	B	C	D
His37 State	+0	55 %	25 %	74 %	41 %
	+1	69 %	56 %	63 %	86 %
	+2	68 %	66 %	72 %	78 %

Table S1. Gly34 backbone carbonyl and water hydrogen bond prevalence

Asymmetric Proton Path

As the excess proton approaches the channel, the water is more bulk-like until ~$CEC_{Z'}$ = -14.0 Å, and thus the CEC distribution above this point is more spread out and potentially out of the channel. After this point, the distribution narrows and is more focused in the center of the channel. The proton approaches and passes through the Val27 gate near $CEC_{Z'}$ = -12.0 through -8.0 Å. The CEC distribution is the narrowest in this region due to the small opening afforded by the Val27 gate – the sidechains fluctuate and shift up enough to let one water molecule bridge between the top-outer portion and the interior of the channel.

Around $CEC_{Z'}$ = -5.0 Å the distribution becomes ring-like, most prominent in the +0 and +1 states, as there are no waters directly in the center of the pore and the proton stays around the edge of the pore. The channel pore radius here is the largest of the internal portion of the channel (between Val27 and Trp41). In the +0 state, the proton has propensity towards one side of the channel near helix C. As the proton moves down, this propensity moves to the opposite side of the pore. This asymmetry can be attributed to the proximity of Ala30 backbone carbonyls and the water structure they influence. Past this, the channel narrows, the distribution again becomes more centralized. As the proton passes through the Gly34 region in the +2 state, approaching the His tetrad, there is slightly greater propensity for the proton to be on the sides of the channel above the uncharged His residues. This is likely due to reoriented His-water hydrogen bonds on protonated His residues that would not easily accept an excess proton.

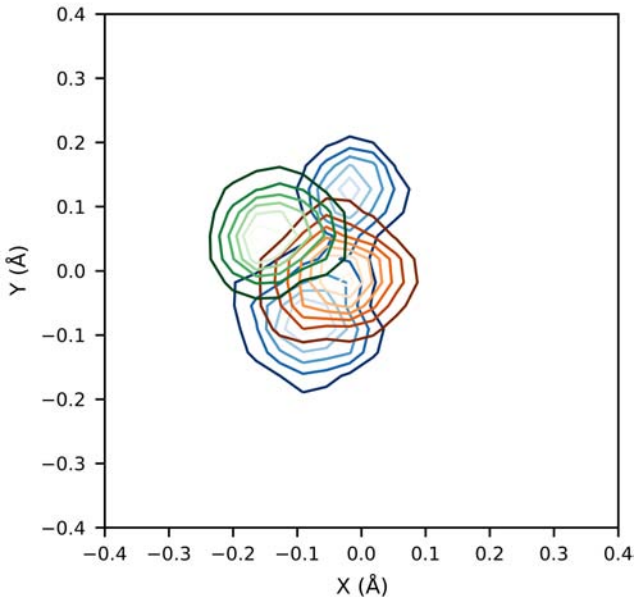

Figure S6. Contour plot of unit vectors describing the protein principal axis projected onto xy-plane. The principal axis and its unit vector were found for all frames and the histogram was calculated with bins of length 0.04 Å. The contour plots of the histograms for the +0, +1, and +2 charge states in blue, orange, and green respectively.

SI Movies 1-3: Same as in Figure 7, for charge states +0, +1, and +2 respectively.

SI Movie 4: 2D-histograms for $CEC_{Z'}$ positions every 0.2 Å, as in Figure 8.

METHODS

Principal axis alignment. In order to calculate properties with respect to the protein's principal axis, the atom positions in all trajectories were rotated such that the protein principal axis was aligned with the z-axis. The principal axis was calculated from all non-hydrogen protein atoms, and the rotation matrix about the center of mass of Gly34 alpha-carbons was calculated from the resulting vector. The hydrated excess proton CEC positions were also rotated to get $CEC_{Z'}$.

Radii. Radii profiles were calculated using the program HOLE[1] for each frame of the aligned trajectories. Profiles were binned according to $CEC_{Z'}$ every 0.5 Å, and the average and standard deviation of the profile were calculated for each bin.

Hydrogen-bonding and water analysis. Frames from all aligned trajectories were binned according to $CEC_{Z'}$ every 0.5 Å.

Hydrogen bonds were identified by the following standard criteria, as has been used in previous M2 studies[2-4]: the distance between the donor and acceptor atoms was less than 3.5 Å, and the angle between the donor, hydrogen, and acceptor was greater than 150°.

Region assignment. The water was divided into four non-overlapping internal sections spanning the area between Val27 and Trp41, chosen in part based on experimental work showing layers of water in the crystal structure.[2] Hydrogen bonds were assigned to regions based on the donor's coordinates.

Abbreviations:

VALSC = geometric center of Val27 sidechain carbons atoms

SER = center of mass of 4 Ser31 alpha carbon atoms

GLY = center of mass of 4 Gly34 alpha carbon atoms

HISSC = geometric center of His37 sidechain atoms

TRPSC = geometric center of Trp41 sidechain atoms

The region definitions:

1. The region below VALSC and at or above SER.
2. The region below SER and at or above GLY.
3. The region below GLY and at or above HISSC, to account for water directly above the His37 tetrad.

4. The region below HISSC and at or above TRPSC, i.e., water between the His37 tetrad and Trp41 gate.

H-bond direction. To determine the direction of a hydrogen bond, the vector between the acceptor and donor atoms was found. The z-component of the unit vector (i.e., the cosine of the angle between the hydrogen bond vector and the z-axis) of this was taken as a measure of the hydrogen bond's alignment with the protein principal axis (using aligned trajectories), yielding values between (-1.0,1.0).

Alpha-carbon distances. The backbone alpha-carbon distances from the center of the channel were calculated for each frame of each trajectory, for all residues on each helix. The center of the channel in a given simulation for a given residue was defined as the center of mass of the four alpha-carbons of that residue (one from each helix) averaged over that trajectory. Thus, these distances capture any outward-inward fluctuations at a given point along a helix and can distinguish the movements of individual helices. Frames were then binned according to $CEC_{Z'}$, and the averages and standard deviations of the distances for each helix were calculated. These averages and standard deviations were then averaged (this is the weighted average of all distances, where each frame is weighted such that each $CEC_{Z'}$ bin has equal contribution to the overall average).

Excess Proton CEC Histograms. The CEC xyz-coordinates were collected every 100 fs. The CEC position in the rotated frame of reference (with respect to Z') was calculated based on the frame nearest in time. Coordinates were binned according to $CEC_{Z'}$ every 0.2 Å. The two-dimensional histogram was then calculated using the Python package NumPy for each bin, and images were generated with Gaussian interpolation to smooth over bin edges while preserving the main data features.

Figures of molecular structures were generated using Visual Molecular Dynamics (VMD) software.[5]

Supporting References

1. Smart, O. S.; Neduvelil, J. G.; Wang, X.; Wallace, B. A.; Sansom, M. S. P., HOLE: A program for the analysis of the pore dimensions of ion channel structural models. *Journal of Molecular Graphics* **1996,** *14* (6), 354-360.
2. Thomaston, J. L.; Woldeyes, R. A.; Nakane, T.; Yamashita, A.; Tanaka, T.; Koiwai, K.; Brewster, A. S.; Barad, B. A.; Chen, Y.; Lemmin, T.; Uervirojnangkoorn, M.; Arima, T.; Kobayashi, J.; Masuda, T.; Suzuki, M.; Sugahara, M.; Sauter, N. K.; Tanaka, R.; Nureki, O.; Tono, K.; Joti, Y.; Nango, E.; Iwata, S.; Yumoto, F.; Fraser, J. S.; DeGrado, W. F., XFEL structures of the influenza M2 proton channel: Room temperature water networks and insights into proton conduction. *Proc Natl Acad Sci U S A* **2017,** *114* (51), 13357-13362.
3. Thomaston, J. L.; Alfonso-Prieto, M.; Woldeyes, R. A.; Fraser, J. S.; Klein, M. L.; Fiorin, G.; DeGrado, W. F., High-resolution structures of the M2 channel from influenza A virus reveal dynamic pathways for proton stabilization and transduction. *Proc Natl Acad Sci U S A* **2015,** *112* (46), 14260-5.
4. Gianti, E.; Carnevale, V.; Degrado, W. F.; Klein, M. L.; Fiorin, G., Hydrogen-bonded water molecules in the m2 channel of the influenza a virus guide the binding preferences of ammonium-based inhibitors. *Journal of Physical Chemistry B* **2015,** *119* (3), 1173--1183.
5. Humphrey, W.; Dalke, A.; Schulten, K., {VMD} -- {V}isual {M}olecular {D}ynamics. *Journal of Molecular Graphics* **1996,** *14*, 33-38.