Acid activation mechanism of the influenza A M2 proton channel

Rubin Liang a,b,c, Jessica M. J. Swanson a,b,c, Jesper J. Madsen a,b,c, Mei Hong d, William F. DeGrado e,f, and Gregory A. Voth a,b,c,1

aDepartment of Chemistry, The University of Chicago, Chicago, IL 60637; bInstitute for Biophysical Dynamics, The University of Chicago, Chicago, IL 60637; cJames Franck Institute, The University of Chicago, Chicago, IL 60637; dDepartment of Chemistry, Massachusetts Institute of Technology, Cambridge, MA 02139; and eDepartment of Pharmaceutical Chemistry, University of San Francisco, San Francisco, CA 94158

Contributed by William F. DeGrado, September 23, 2016 (sent for review July 13, 2016; reviewed by José D. Faraldo-Gómez and Wei Yang)

The homotetrameric influenza A M2 channel (AM2) is an acid-activated proton channel responsible for the acidification of the influenza virus interior, an important step in the viral lifecycle. Four histidine residues (His37) in the center of the channel act as a pH sensor and proton selectivity filter. Despite intense study, the pH-dependent activation mechanism of the AM2 channel has not been completely understood at a molecular level. Herein we have used multiscale computer simulations to characterize (with explicit proton transport free energy profiles and their associated calculated conductances) the activation mechanism of AM2. All proton transfer steps involved in proton diffusion through the channel, including the protonation/deprotonation of His37, are explicitly considered using classical, quantum, and reactive molecular dynamics methods. The asymmetry of the proton transport free energy profile under high-pH conditions qualitatively explains the rectification behavior of AM2 (i.e., why the inward proton flux is allowed when the pH is low in viral exterior and high in viral interior, but outward proton flux is prohibited when the pH gradient is reversed). Also, in agreement with electrophysiological results, our simulations indicate that the C-terminal amphipathic helix does not significantly change the proton conduction mechanism in the AM2 transmembrane domain; the four transmembrane helices flanking the channel lumen alone seem to determine the proton conduction mechanism.

The influenza A M2 channel (AM2) is activated when the pH of the endosomally entrapped virus is lowered, resulting in proton flux into the virus interior, a crucial step in viral replication (5–7). M2 has a single transmembrane helix, which tetramerizes to form the conducting pore. The His37 tetrad (Fig. L4), located near the center of the M2 transmembrane helix (M2TM), serves as a pH sensor that activates the channel as pH is lowered to near 6.0. His37 can also act as a shuttle that is sequentially protonated and deprotonated as an excess proton transits the activated channel. The Trp41 tetrad, next to the His37 tetrad toward the C terminus and virus interior, functions as a pH-dependent gate for proton conductance (8). The interplay between the protonation state of the His37 tetrad and conformations of the protein backbone and Trp41 side chains is essential for the pH activation mechanism of the channel (9–12). Under high-pH conditions the Trp41 side chains close the C-terminal pore below His37, forming a gate that blocks proton flow through the channel (referred to below as the “Cclosed” state).

As the pH decreases to ~6, the protonation state of the His37 tetrad reaches a critical level [likely Q2 (7)] and the channel becomes activated, able to enter the primary proton conduction cycle (4, 9, 13–17). (Herein the notation Qn denotes the state of the His37 tetrad, where n is the number of biprotonated histidines, ranging from 0 to 4.) In this cycle, a proton diffusing down its concentration gradient through the activated M2 channel transiently protonates His37, generating the Q3 state (assuming Q2 as the activated resting state). The high density of positive charge in the channel destabilizes the Cclosed state sufficiently to open an aqueous conduction path (8, 9, 16, 18). Diffusion of a

Significance

The influenza A M2 channel (AM2) transports protons into the influenza virus upon acid activation. It is an important pharmacological target as well as a prototypical case to study proton conduction through biological channels. The current work provides the most complete computational characterization to date of the physical basis for the acid activation mechanism of the AM2 proton channel. Our results show that lowering the pH value gradually opens the Trp41 gate and decreases the deprotonation barrier of the His37 tetrad, leading to channel activation. Our result also demonstrates that the C-terminal amphipathic helix does not significantly change the proton conduction mechanism in the AM2 transmembrane domain.


Reviewers: J.D.-F.G., NIH/National Heart, Lung, and Blood Institute; and W.Y., Florida State University.

The authors declare no conflict of interest.

1To whom correspondence may be addressed. Email: gavoth@uchicago.edu or william.degrado@ucsf.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1615471113/-/DCSupplemental.
proton into the interior of the virus could return the His37 tetrad back to the Q2 state, with accompanying conformational relaxation and reequilibration of the His37 tautomers, thereby completing the cycle (8, 13, 18).

Structures determined using X-ray crystallography, solution NMR, and solid-state NMR (SSNMR) support this acid-activated transporter-like mechanism (9, 10, 18–26). Near neutral pH the channel is closed in the C_{closed} conformation (Fig. 1) with tight packing of His37 and Trp41 (10, 19, 23, 26). At lower pH the channel is more dynamic (26–28), which has precluded structure determination by solution or SSNMR. However, a series of crystallographic structures of the C_{open} states (Fig. 1) (23–25) have been determined, including high-resolution structures of the TM domain crystallized from lipidic cubic phase (24, 25). The C_{open} forms seen in these crystal structures (23–25) show more structural heterogeneity than the C_{closed} forms, despite the C_{closed} form’s having been crystallized over a range of pH values. These findings suggest that the allosteric coupling between protonation of His37 and opening of the C-terminal region is not absolute. In other words, a significant fraction of the protein in solution exists in multiple conformations, which can be captured within the crystal lattice over a range of pH values. This conclusion is in agreement with NMR measurements, which show a pH-dependent population of multiple conformational states (10–12, 26, 29, 30), as well as molecular dynamics (MD) simulations, which show graded increases in dynamics with increasing protonation state (16, 18, 31–38).

Classical MD has been very useful for examining the conformational relaxation that occurs following a change in the protonation state of the His37 tetrad during activation and conduction. However, a significant assumption of these classical MD analyses has been that the conformational changes are both computationally reversible and can be captured on the 100-ns to microsecond time scales that are readily accessible by MD (18, 31–37). Given the availability of a high-resolution structure for both the C_{open} and C_{closed} forms of the protein it is now possible to determine the reversibility of the calculations and to determine whether the starting configurations substantially alter the conformational ensemble and computed energetics of proton translocation through the channel.

The major rate-limiting step for conduction (defined as conductance × [H^+]^{-}) involves diffusion of protons to His37 at low permeant ion concentration (high pH). As the pH is lowered below the midpoint of the pH/flux curve, the chemical step of proton dissociation from His37 becomes rate-limiting (4, 16, 39). Because the conduction process involves bond making and breaking steps associated with excess proton diffusion (Grotthuss hopping) through water wires as well as protonation and deprotonation of His37, it is not possible to address the conduction mechanism directly using classical MD. However, as noted earlier the proton conduction process can be studied via multiscale simulation. For example, in our recent multiscale simulation study (16) we reported that during the transition from intermediate to low pH the C-terminal helices of the channel expand and the hydration level in the C-terminal pore is increased. This decreases the His37 deprotonation barrier and increases the proton conductance at low pH, which partially explains the acid activation mechanism of the channel. However, the acid activation mechanism for the transition from high to intermediate pH was not investigated in ref. 16. We note that recent experimental work by Cross and coworkers (27) confirmed (but did not cite) some of the results from ref. 16, such as the gating role of Trp41 and rapid equilibration of protons between His37 and the viral exterior originally observed electrophysiologically (8, 18).

Herein, we build upon our previous simulations of explicit PT in AM2 (16) and use recently published crystal structures (24) to probe the coupled scales and detailed mechanism of the AM2 acid activation behavior for the transition from high to intermediate pH. We have performed multiscale simulations starting from the high-pH room-temperature crystal structure in ref. 24 [Protein Data Bank (PDB) ID code 4QKL]. Our results show that, in lipid bilayers, the channel closes its C-terminal helices under high-pH conditions. Using permeation theory, we show that the resulting large deprotonation barrier of the His37 tetrad blocks the channel. Lowering the pH value increases the charge state of the His37 tetrad, which gradually opens the Trp41 gate and increases its hydration. This increases the channel conductance and leads to activation of the AM2 channel.

We also study here the importance of including the amphipathic helices of the AM2 channel by performing simulations starting from an earlier NMR structure (PDB ID code 2L0J) (10), which includes the transmembrane domain (MT2M; residues 26–46) and the amphipathic helices (residues 22–62 in total). The C-terminal amphipathic helix of M2 is essential for endosomal sorting complex required for transcription (ESCRT)-independent virus scission and filamentous virus formation (40), acting to stabilize the negative Gaussian curvature associated with the budding intermediate (41). Although the sequence and...
amphipathic character of this helix are essential to these processes, mutations that affect these parameters do not have a measurable effect on the conductance of the channel (41). Indeed, this C-terminal helix can be deleted entirely without greatly influencing the PT properties of the channel (42). This finding is also consistent with SSNMR structures of the channel (10, 26, 43), which show the amphipathic helix radiating away from the pore of the channel. For this reason our previous simulations (16, 34) and many other computational and experimental studies (9, 11, 30, 32, 44–46) have focused on the M2TM domain to investigate the AM2 proton conduction mechanism. However, it has recently been suggested, based on changes in the NMR signal, that the M2TM and full-length AM2 protein have different structure, dynamics, and function (including PT mechanism) (27). This rationale apparently led those authors to disregard previous work on the M2TM (16), despite the fact that it predicted several of their findings. Moreover, because the influence of the amphipathic helices has not yet been explicitly tested in simulations, we sought to verify whether they are or are not important for PT in the AM2 channel. Here we provide computational evidence confirming the structural and functional measurements (9, 11, 30, 32, 44–46) and demonstrating that the presence of these amphipathic helices does not significantly change the predicted conductivity of the channel.

Results
Our results are organized as follows. First, we discuss the structure of the AM2 channel under high-pH conditions obtained from classical MD simulations starting from the 4QKL crystal structure. Second, we discuss the pH activation mechanism of the AM2 channel based on the free energy profiles for the explicit PT across the channel obtained from a multiscale bridging of multiscale reactive molecular dynamics (MS-RMD) and quantum mechanics/molecular mechanics (QM/MM) MD simulations (Methods and Simulation Details).

AM2 Has Closed C-Terminal Helices Under High-pH Conditions. The M2TM crystal structure 4QKL (residues 22–46), resolved at room temperature and high pH (pH 8.0) in the lipidic cubic phase, features open C-terminal helices. In contrast, the M2TM G34A mutant crystal structure 3LBW (residues 25–46) was resolved at an intermediate pH (pH 6.5) in detergent and features more closed C-terminal helices (Fig. 1D). Previous wild-type simulations starting from the 3LBW structure showed that the protein structure is stable in lipid bilayer at room temperature for the Q1 and Q2 states, which corresponds to the high to intermediate pH range (9). It was also shown by previous simulations that the C-terminal helices start to open in the Q3 state, which corresponds to an intermediate to low pH range (9, 16). Therefore, it was necessary to also test whether or not the 4QKL structure remains stable in the lipid bilayer under high to intermediate pH conditions. To do so, we inserted the 4QKL crystal structure into a 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) lipid bilayer and equilibrated the protein structure for ∼500 ns for the Q0 state. Following this simulation, we increased the charge state of the His37 tetrad and equilibrated the system again for the Q1 and Q2 states, which mimics the activation of the channel upon lowering the pH to an intermediate value.

In the Q0 state, the C-terminal helices transition from the C
\text{open} to a canonical C
\text{closed} conformation, with the protein backbone resembling that of the 3LBW crystal structure (Figs. S14 and S24). The closing of the C-terminal helices brings the His37 and Trp41 side chains closer together and dehydrates the region between His37 and Trp41. The dehydroxylation is dramatic when we compare the water density profile for the equilibrated Q0 state structure with that in which the backbone C\alpha atoms are restrained to the 4QKL structure (Fig. 1 A and E). In the C-terminal half of the channel (3 Å < z < 16 Å) the pore hydration level is significantly reduced for the unrestrained equilibrated protein, compared with that with protein backbone restrained to the 4QKL structure (Fig. 1A vs. Fig. 1E). In the Q1 and Q2 states, the pore hydration is slightly increased, which induces larger backbone conformational flexibility (Fig. 1 B and C). However, the hydration level in the C-terminal pore is still lower than when the backbone is restrained to the 4QKL structure, especially in the His37 and Trp41 regions (Fig. 1E). Thus, 4QKL structure may represent a limiting case of a fully diluted C
\text{open} conformation that was trapped in the lipid cubic phase during crystallization. The differences in the Q2 water density profile near the N-terminal helices between the 3LBW and 4QKL equilibrated structures (black and purple curves in Fig. S2B, respectively) are likely due to the fact that the 4QKL structure has three more residues (22–24) in the N-terminal region than the 3LBW. This addition causes the N-terminal helices in the equilibrated 4QKL structure to be more closely packed than in the equilibrated 3LBW structure in the Q2 state, resulting in lower water density in the N terminal for the former. This difference in the N-terminal packing propagates to the C-terminal helices, making the C-terminal helices in the equilibrated 4QKL structure slightly more open than those in the equilibrated 3LBW structure in the Q2 state (Fig. 1C vs. Fig. 1D).

Regardless, the low hydration level in the C-terminal pore obtained after equilibration in a lipid bilayer at the high-pH condition (Q0 and Q1 states), as well as the close packing of the Trp41 side chains, contribute to large free energy barriers for His37 deprotonation, as discussed below.

AM2 Channel Has a Low Conduction Rate in the Inactivated Conformations Present at High pH. We computed the free energy profile (or potential of mean force, PMF) for the charge defect associated with an excess proton to travel through the M2 channel (Fig. 2), including (i) diffusion from the N-terminal end of the pore into the outward-facing aqueous channel, (ii) protonation of the His37 tetrad, (iii) dissociation from the His37 tetrad, and (iv) diffusion to the C-terminal end of the pore. Thus, the +0 PMF describes the process where the Q0 state His37 tetrad is protonated by an excess proton from the viral exterior to transiently form the Q1 state, followed by a deprotonation of the excess proton to the viral interior to regenerate the Q0 state. The PMF was calculated as a function of a collective variable (hereafter referred to as the “CV”) that tracks the progress of the excess proton center of excess charge (CEC) during the PT process (Methods and Simulation Details). As the proton leaves the pore and enters into bulk solution on either side of the membrane, a cylindrical restraint (47) is used to quantify the free energy obtained by releasing the excess proton to the bulk proton concentration. Thus, the proton activity in bulk solution (i.e., pH), whether symmetric or asymmetric, does not enter into the calculation. Instead, these variables are considered via electropermeation theory following the construction of the PMF.

The PMF profiles for the Q0 and Q1 states, which predominate at high pH, show that the protein is not sufficiently activated in these states to enable efficient proton flux. A proton passing through the channel in the Q1 protonation state is influenced by a deep local free energy minimum, indicating that there is a strong preference for the proton to reside on His37 vs. other regions of the pore. The barrier for His37 deprotonating the excess proton to the viral interior (CV from 6 Å to 11 Å) is more than 15 kcal/mol. Application of permeation and kinetic theories of proton conduction to this PMF predicts a maximal possible conduction at a hypothetical saturating proton concentration for a given initial charge state of His37 (Q0 and Q1 for +0 and +1 PMFs, respectively). They are well below the experimentally measured conductance at intermediate pH when the AM2 channel is activated (Table 1). Thus, this conformational ensemble is not fit for proton conduction. Similar observations are also true for the +1 PMF. In both cases (i.e., the Q0 and Q1
proton in high-pH conditions (i.e., Q0 or Q1 states) than at low pH (i.e., Q2 state). This is in agreement with the higher pK_a values for the first and second protonation of the His37 tetrad than the third and fourth (7, 30, 48, 49). Charge delocalization of the excess proton among the four His37 residues contributes to the stability of protonation for the Q0 and Q1 states.

The maximum proton conductance, calculated from Poisson–Nernst–Planck (PNP) electrodiffusion theory (Methods and Simulation Details), and the rates for His37 deprotonation in both directions of the channel, estimated from transition state theory (Supporting Information), are listed in Table 2. Differences between the total conductance and rate-limiting His37 deprotonation steps are predominantly due to the barrier for PT from the viral exterior to the His37 tetrad. An additional factor, though its impact is less, is the influence of varying proton diffusivity across the regions of the PMF, which is included in the conductance calculation but not included in rate calculations. Additional insight is gained by considering the rates of His37 deprotonation to just the first solvation shell, which are also listed in Table 2. For the Q0 and Q1 states, the rate constants for the forward proton deprotonation to the first solvation shell are much slower than those for the overall forward deprotonation. This is because the overall forward deprotonation process also includes PT from the hydronium below His37 through the Trp41 residues. This step results in additional barrier and much slower forward deprotonation rate in the Q0 and Q1 states (Tables 1 and 2).

Mechanism for Rectification. A particularly important and interesting feature of the PMFs in Fig. 2 is the pronounced asymmetry in the +0 PMF. The barriers for PT between the His37 tetrad and the viral exterior and interior are quite different. In particular, the forward protonation barrier (PT from the viral exterior to His37) is small (∼−4.7 kcal/mol) relative to the forward deprotonation barrier (PT from His37 to the viral interior; ∼18.0 kcal/mol), whereas the backward protonation barrier (PT from viral interior to His37) is larger (∼8.9 kcal/mol) relative to the backward deprotonation barrier (∼12.9 kcal/mol). Including the approximate time for diffusion of an excess proton to the channel entrance slows down the k_on (protonation) rates. This can be equivalently expressed in the context of transition state theory by increasing the protonation barriers. For example, at a pH of 4.9 [the measured pK_a of the third His37 residue (30), which should correspond to the pH at which the channel activates] diffusion would effectively add 6.7 kcal/mol (ΔG = 2.3RTpH) to the protonation barrier in either direction. Thus, the forward effective

states) the PMFs and calculated conductances indicate that the channel is inactive under high-pH conditions.

The large deprotonation barrier of His37 for the Q0 and Q1 states is caused by the lack of stable hydration between the His37 and Trp41 tetrad (Fig. 1A and B). The close packing of the Trp41 side chains at high pH (Q0 and Q1 states) reduces the number of water molecules accessible for His37 deprotonation, causing a substantial barrier for the deprotonation of His37 to the first solvation shell of water molecules toward the viral interior (CV from 6 Å to 8 Å in Fig. 2A). Steric hindrance of the Trp41 side chains further increases the barrier when the hydrated excess proton moves from the first solvation shell to the second shell below His37 (CV from 8 Å to 11 Å). In addition to the high deprotonation barrier, the free energy minima in the His37 tetrad region are deep relative to bulk (∼−10 kcal/mol) in the +0 and +1 PMFs, which is consistent with the estimated pK_a values for the first two His37 residues (∼7.6 and 6.8 according to ref. 30). Thus, it is more energetically favorable for the His37 tetrad to bind the incoming
protonation barrier (∼11.4 kcal/mol) is lower than the forward protonation barrier (∼18.0 kcal/mol), such that \(k_{\text{on}}\) outcompetes \(k_{\text{off}}\) and the His37 tetrad jumps from the +0 to the +1 charge state. The equivalent happens for the +1 to +2 transition and the channel activates when the pH outside of the virus is low. In the opposite direction, however, the backward protonation barrier (∼15.6 kcal/mol) is larger than the deprotonation barrier (∼12.9 kcal/mol), such that \(k_{\text{off}}\) outcompetes \(k_{\text{on}}\) and the channel remains in the +0 charge state. Although this introduces outward proton flow, it would be below the detectable limit. Thus, the channel activates and conducts inward current when \(\text{pH}_{\text{int}}\) is low and \(\text{pH}_{\text{out}}\) is high but stays inactive and does not conduct outward current when \(\text{pH}_{\text{int}}\) is low and \(\text{pH}_{\text{out}}\) is high. This property is called rectification (8, 50). Although the +0 PMF qualitatively explains the inward rectification mechanism, we note that the calculated difference between the backward \(k_{\text{on}}\) and \(k_{\text{off}}\) could be larger. This could be due to the calculated backward protonation barrier in the +0 PMF (∼8.9 kcal/mol) being underestimated, possibly due to an incompletely equilibrated transmembrane structure and/or from errors in the force-field representation of the Trp41 side-chain interactions. A Q0 state with an even more tightly packed Trp41 tetrad would likely increase the backward protonation barrier, thereby more effectively locking the system in the Q0 state and blocking outward conduction. Nevertheless, the asymmetry in the calculated +0 PMF is clearly evident and is qualitatively consistent with previous experimental results regarding the rectification behavior (8, 50).

It is also interesting to note that the M2 channel in the Weybridge strain of influenza virus (WM2) exhibits the rectification mentioned above, but the M2 in the Rostock strain (RM2) does not (50). This is because the RM2 has a D44N mutation compared with WM2. This mutation has the effect of breaking the hydrogen bond between Trp41 and Asp44 in WM2, opening the Trp41 gate at high pH value (18). Therefore, in RM2 the proton can reach His37 more easily from viral interior, eliminating the rectification.

**Mechanism of Proton Conduction Through the Activated, Q2 State.** In the +2 PMF, the doubly protonated His37 tetrad is transiently protonated by another excess proton from the viral exterior to generate the Q3 state, then it deprotonates a proton to the viral interior to regenerate the Q2 state. The His37 deprotonation barrier for the Q2 state is decreased to ∼12 kcal/mol, (Fig. 2A) and the proton conductance is increased accordingly (Table 1). The rates for His37 deprotonation to the first solvation shell calculated from the +2 PMF are in good agreement with the proton exchange rate between water and histidine in intermediate pH condition measured by NMR experiment (30) (Table 2). The rates are faster than those in the Q0 and Q1 state, which is consistent with the experimental finding that high pH suppresses the histidine–water proton exchange rate in the S31N mutant (51). The conductance is in reasonable agreement with our previous result (16) and the experimental values (10, 52, 53). The difference between the present +2 PMF compared with our prior result is likely due to the three extra residues in the N-terminal helices in the new structure, which cause conformational changes in the backbone structure in both the N- and C-terminal regions (Fig. 1C vs. Fig. 1D), as discussed above.

The decrease in the His37 deprotonation barrier in the +2 PMF can be attributed to increased solvation between His37 and Trp41 (Fig. 1C), as well as increased flexibility of the Trp41 side chains. The more expanded C-terminal helices provide more space for the rotational motion of the Trp41 residues, and as a result the barrier for the excess proton to pass through the Trp41 is reduced (Fig. 2A, CV from 8 Å to 11 Å). This is consistent with the experimental finding that the Trp41 side-chain rotations are less restricted at lower pH (12). Importantly, as seen in Fig. 1C the free energy minimum in the His37 tetrad region for the Q2 state is raised compared with the +0 and +1 PMFs, which is consistent with the lower \(p_{K_a}\) value for the third His37 residue. The rise of the free energy minimum at the His37 tetrad is caused by the less favorable electrostatic interactions between the His37 tetrad and the excess proton in the Q2 state compared with Q0 and Q1 states. Specifically, upon increasing the charge of His37 tetrad from +0 to +2 the barrier from the viral exterior to His37 increases due to larger electrostatic repulsion between the excess proton and the His37 tetrad. At the same time, the barrier from His37 to the viral interior is decreased due to less charge delocalization in the more positively charged His37 tetrad. The combined effect is a shallower free energy minimum at His37 in the Q2 state.

**Effect of Amphipathic Helices.** Because the full-length M2 protein includes both M2TM and the amphipathic helices, as noted earlier we also investigated the influence of the amphipathic helices on the proton conductance mechanism for M2TM with the helices included in the calculations.

The +2 PMF for the system with amphipathic helices equilibrated from the modified NMR structure 2LOJ (10) (discussed below) is, in general, very similar to that of the systems without amphipathic helices equilibrated from the 4OKL and 3LBW structures (Fig. 2B). Specifically, the His37 deprotonation process, which is the rate-limiting step, has a similar barrier in both cases, resulting in similar calculated conductances. Therefore, the existence of the amphipathic helices does not significantly influence the proton conductance mechanism. This is in line with the experimental measurements showing that the proton conductance is similar for the protein construct having only the M2TM and the construct having both the M2TM and the amphipathic helices (42).

### Table 2. Estimated rates of deprotonation of the His37 to the first solvation shell in the various states in the two possible directions, from transition state theory estimates (Supporting Information)

<table>
<thead>
<tr>
<th>PT process</th>
<th>Forward</th>
<th>Backward</th>
</tr>
</thead>
<tbody>
<tr>
<td>+0</td>
<td>4.1 × 10^4</td>
<td>3.8 × 10^1</td>
</tr>
<tr>
<td>+1</td>
<td>5.0 × 10^4</td>
<td>2.0 × 10^1</td>
</tr>
<tr>
<td>+2</td>
<td>3.1 × 10^4</td>
<td>1.8 × 10^2</td>
</tr>
<tr>
<td>+2 (with amphipathic helices)</td>
<td>6.7 × 10^1</td>
<td>1.4 × 10^2</td>
</tr>
<tr>
<td>+2 (3LBW)</td>
<td>2.2 × 10^3</td>
<td>1.7 × 10^2</td>
</tr>
<tr>
<td>Experiment (intermediate pH)*</td>
<td>≈ 4.5 × 10^2</td>
<td></td>
</tr>
</tbody>
</table>

*From ref. 30.
Rotamer of the Trp41 Residue. For the equilibration simulation starting from the NMR structure 2L0J (10), which includes both the M2TM and amphipathic helices, the structure stabilized only after adjusting the Trp41 side chains to a t90 rotamer configuration (Fig. S3) (the equilibration starting from the t-105 rotamer of Trp41 because in the NMR structure 2L0J does not yield a stable protein structure within 200-ns simulation). Our decision to change to the t90 rotamer of Trp41 is supported by crystal structures (9, 24, 55), SSNMR measurements (12, 56), as well as long all-atom MD simulations (32), the latter showing that all four Trp41 residues flipped from the t-105 to the t90 rotamer after 500 ns. To further address this issue, we also calculated the PMF (Fig. 3C) for the rotation of the χ2 dihedral angle of one of the four Trp41 residues in the lipid bilayer, restraining the backbone of the M2TM to the initial 2L0J structure. The PMF in the figure indicates that the minimum at the t90 rotamer is more stable than that at the t-105 rotamer, which further supports our choice of the t90 rotamer. Once the t90 rotamers were in place, our simulations in the membrane were stable (Fig. 3A and B). In addition, after removing the amphipathic helices from 2L0J structure and equilibrating only the M2TM in the membrane, the M2TM structure remained very similar to the initial 2L0J structure, suggesting that the amphipathic helices do not significantly change the M2TM structure.

Molecular-Level Behavior of the His37 Tetrad. Finally, we comment on the hydrogen bonding in the His37 tetrad before and during proton permeation. In the QM/MM simulations of the +0 and +1 PMFs there are no direct hydrogen bonds between the His37 residues, as suggested by Cross and coworkers (10). This is consistent with the experimental finding of Hong et al. (57) that the His37 residues form hydrogen bond only with water instead of with one another. Typical snapshots of the configuration of the His37 tetrad and solvating water molecules are shown in Fig. S4 A and B. In the QM/MM simulations of the +2 PMFs we again did not observe direct hydrogen bonds between the His37 residues. In the 2L0J structure, the four His37 residues in the Q2 state have a dimer of dimers configuration via direct hydrogen bonds between His37 residues (10). However, it should be emphasized that this geometry resulted from hypothetical distance restraints, and there is no direct experimental spectroscopic evidence for the interaction. After only 100 ns of classical MD equilibration with the t90 rotamer of Trp41 all direct histidine–histidine hydrogen bonds are broken, and only water–histidine hydrogen bonds remain. Typical snapshots of the configuration of the His37 tetrad and solvating water molecules are shown in Fig. S4 C and D. The difference in the His37 hydrogen-bonding patterns between the simulation and NMR structures could be due to the difference in the Trp41 rotamer in the two structures, as discussed above.

As the proton shuttles through the His37 tetrad, we observe in the QM/MM MD simulations the His37 residues undergo interconversion among Τ, Π, and cationic states via the protonation of a delta-N and deprotonation of an epsilon-N. Tautomer exchange within a His37 residue would require a very high energy unprotonated/negatively charged intermediate and is not expected to happen. However, interresidue tautomerization is likely given the stable and similar to the initial structure, both in the presence and absence of the amphipathic helices. This suggests that the amphipathic helices do no significantly change the structure of the M2TM structure in a POPC membrane. (C) The free energy profile for the rotation of the χ2 angle of one of the four Trp41 residues in the 2L0J structure in a lipid bilayer. The regions of the PMF corresponding to the t90 rotamer (in the crystal structures 3LBW and 4QKL) and t-105 rotamer (in the NMR structure 2L0J) are shown. The free energy for the t90 rotamer is ∼3 kcal/mol lower than that of the t-105 rotamer with a ~2 kcal/mol barrier for conversion, indicating that the t90 rotamer is energetically more favorable than the t-105 rotamer in the 2L0J structure.
water-mediated hydrogen bonding between $\tau$ and $\pi$ tautomers. This is in agreement with the mechanism proposed by Williams et al. (figure 10 in ref. 51) based on their SSNMR measurements.

**Conclusions**

We find that under high-pH conditions the AM2 C-terminal helices adopt a $C_{\text{closed}}$ conformation. The closed Trp41 tetrad dehydrates the His37 tetrad and increases the His37 deprotonation barrier, thus blocking proton conduction at high pH. In contrast, the low charge state of the His37 tetrad at high pH reduces the electrostatic repulsion with the excess proton, lowering the barrier for proton permeation from the viral exterior. Thus, the protonation of His37 occurs more readily from the viral exterior than from the viral interior in the low charge state of His37 tetrad, and this asymmetry at least partly explains the rectification of proton flow observed in experiments (8, 50). As the pH decreases to intermediate values the positive charge on the His37 tetrad increases and the C-terminal helices are opened and more hydrated. This lowers the His37 deprotonation barrier and increases the proton conductance, thus activating the channel. Our previous multi-scale simulation (16) has shown that as the pH further decreases from intermediate to low value (Q2 to Q3 state) the channel expansion and increased pore water dynamics further increase the proton conductance, despite the increased electrostatic repulsion between the His37 tetrad and the incoming proton. In combination with our previous work (16), the present results provide the most complete picture to date of the physical basis for the acid activation mechanism of the AM2 proton channel. The activation mechanism is summarized in Fig. 4. Additionally, our results indicate that the amphipathic helices do not significantly influence the proton conductance mechanism of the AM2 channel or its quantitative value.

**Methods and Simulation Details**

**Classical MD Simulations.** The crystal structure resolved at room temperature and high pH (4QKL) (24) and the NMR structure obtained at intermediate pH (2L0J) (10) were used as initial structures for classical MD simulations. For the simulations starting from crystal structure the protein was inserted into a lipid bilayer membrane consisting of 207 POPC molecules and then solvated by a 25-Å-thick layer of water molecules on each side (Fig. 1A). Although a pure POPC bilayer is lacking the lipid heterogeneity of a real membrane, it is a reasonable choice for the sake of consistency (POPC has been has been used in most of the recently published M2 simulations (9, 16, 24)) and a widely used component of virus-mimetic membranes in experiments (12, 42, 49, 51)). The protein’s principal axes were aligned with the x z plane. For the equilibration of Q0 state all four His37 residues were singly protonated at the N atom (10, 11). In the first stage of equilibration, 100 kcal mol$^{-1}$ Å$^{-2}$ harmonic restraints were first applied to the protein backbone C$^\alpha$ atoms and then gradually reduced to 0.01 kcal mol$^{-1}$ Å$^{-2}$ over 30 ns, while the lipid and water molecules relaxed to stabilize the protein–lipid bilayer packing. Next, the harmonic restraints were removed, allowing all molecules in the system to relax for another 5 ns. Equilibration was followed by ~500 ns of MD simulation. Equilibrations of the Q1 and Q2 state systems were initiated from the final structure of the Q0 state production simulation by protonating of one or two of the His37 residues, respectively. For the Q2 state, the two charged His37 residues were diagonal to each other (32). For both states, 1 kcal mol$^{-1}$ Å$^{-2}$ harmonic restraints were applied to proton backbone C$^\alpha$ atoms and gradually reduced to 0.01 kcal mol$^{-1}$ Å$^{-2}$ over ~10 ns. Then all harmonic restraints were released and a production trajectory of ~180 ns was sampled for each system.

For the simulations starting from the NMR structure that includes the amphipathic helices the protein was inserted into lipid bilayer membrane consisting of 207 POPC molecules and then solvated by a 30-Å-thick layer of water molecules on each side. The protein’s principal axes were aligned with the x-axis, and the lipid bilayer was aligned with the yz-plane. Only the Q2 state was simulated, where the two doubly protonated His37 residues were in a diagonal position to each other (32), and the singly protonated His37 residues were protonated at the N atoms (10, 11). Importantly, the equilibration starting from the t-105 rotamers of the Trp41 residues, as found in the 2L0J NMR structure, did not yield a stable protein structure for over 200 ns. Moreover, one of the Trp41 residues flipped to the t90 conformer, as found in the crystal structures (9, 24, 55), during this simulation. This is consistent with another very long simulation study in which all four Trp41 residues eventually flipped to the t90 conformation (32), experimental support of the t90 conformation (9, 12, 24, 55, 56), and our own PMF calculation of the Trp41 $\chi_2$ dihedral angle rotation discussed earlier in this paper. Therefore, to facilitate the equilibration, the remainder of the Trp41 side chain was manually aligned from the t105 to the t90 rotamer. In the first stage of the equilibration, 100 kcal mol$^{-1}$ Å$^{-2}$ harmonic restraints were first applied to the all of the protein backbone’s C$^\alpha$ atoms for ~30 ns, while the lipid and water molecules relaxed to stabilize the protein–lipid bilayer packing. In the second stage, the restraints on the C$^\alpha$ atoms of amphipathic helices were gradually reduced to 0.01 kcal mol$^{-1}$ Å$^{-2}$ over 20 ns, while keeping the restraints on the M2TM C$^\alpha$, at 100 kcal mol$^{-1}$ Å$^{-2}$. In the third stage, the amphipathic helices were allowed to move freely, while the restraints on the C$^\alpha$ atoms of M2TM helices were reduced to 0.01 kcal mol$^{-1}$ Å$^{-2}$ over 65 ns. Then, all of the harmonic restraints were totally removed, allowing all molecules in the system to relax for another 10 ns. Equilibration was followed by ~220 ns of MD simulation.

The protein and lipid were modeled using the CHARMM36 force field (58–61) and the water molecules were described using the TIP3P model (62). The electrostatic interactions were treated with particle mesh Ewald, using a spherical cutoff of 12.0 Å and an accuracy threshold of $10^{-6}$ (63). The Lennard-Jones (LJ) interaction cutoff was 12.0 Å using a switching function starting at 10 Å. Simulations were performed in the constant number, pressure, temperature ensemble at 308 K and 1 atm. For simulations starting from the crystal structure the classical MD simulations were performed with the NAMD package (64). The temperature and pressure were controlled by the Langevin temperature coupling and Langevin piston pressure coupling schemes (65), respectively. The equations of motion were integrated using the velocity Verlet integrator, with time step of 2.0 fs (66). For the simulations starting from NMR structure, the classical MD simulations were performed with GROMACS 5.0.4 software package (67, 68) to take advantage its graphics processing unit acceleration. The temperature and pressure were controlled by the V-rescale algorithm (69) and the Berendsen pressure coupling algorithm (70), respectively. The equations of motion were integrated using the leap-frog integrator, with time step of 2.0 fs.

**QM/MM MD Simulations.** The free energy profile (or PMF) for PT through the His37-Trp41 tetrad region (Fig. 1A) was calculated with hybrid QM/MM umbrella sampling simulations (16). The PMF was calculated as a function of a collective variable (hereafter referred to as the “CV”) defined as the z coordinate difference between the center of mass of C$^\alpha$ atoms of Gly34 and that of the excess proton CEC. (The CV here is in essence a “reaction” or permeation coordinate for the PT process.) The excess proton CEC coordinate used in the QM/MM simulation was defined as (71)

\[
\xi = \sum_{i=1}^{N_O} \sum_{j=1}^{N_O} \sum_{k=1}^{N_O} \sum_{l=1}^{N_O} \sum_{m=1}^{N_O} \sum_{n=1}^{N_O} \sum_{p=1}^{N_O} \sum_{q=1}^{N_O} \sum_{r=1}^{N_O} \sum_{s=1}^{N_O} \sum_{t=1}^{N_O} \sum_{u=1}^{N_O} \sum_{v=1}^{N_O} \sum_{w=1}^{N_O} \sum_{x=1}^{N_O} \sum_{y=1}^{N_O} \sum_{z=1}^{N_O} \chi_{\text{position}}
\]

where the $X_i$‘s are the histidine nitrogen atoms and the water oxygen atoms in the QM region, and $H_i$‘s are the hydrogen atoms bound to those heavy atoms in the QM region. The $w_i$‘s are the hydrogen coordination number of...
a molecule in the least protonated state during the PT process. Therefore, for water oxygen atoms they were set to 2, and for His37 nitrogen atoms they were set to 0.50 in +0 PMF (four hydrogens shared by eight nitrogens in the Q0 state). Similarly, the w0-ws for the His37 nitrogen atoms were set to 0.625 in +1 PMF (five hydrogens shared by eight nitrogens in Q1 state) and 0.75 in +2 PMF (six hydrogens shared by eight nitrogens in the Q2 state). The dx, dy is the distance between atom X and atom H, and fsw(dx, dy) = 1/(1 + exp(dx + dy))/dx, dy reflects the coordination number of H to X, with the parameters chosen as dx, dy = 0.04 Å, rw = 1.25 Å (72).

In addition, a correction term fcorr was introduced to the excess proton CEC definition in Eq. 1 (16) to correct for the contribution due to the presence of multiple protons around multiple protonatable sites in the His37 tetrad, such that

\[ \xi_{\text{correct}} = \sum_{i=1}^{8} \sum_{j=1}^{8} m_i R^X_{i,j}, \]

where X denotes the eight nitrogen atoms of the His37 tetrad, and \( m_i \) gradually switches from 1 to 0 as the nitrogen atom i is deprotonated. The excess proton CEC as defined in Eq. 1 describes the delocalized nature of the excess proton and has been shown to accurately locate the position of the excess proton charge defect in previous QM/MM simulations of biological PT channels (16, 71–73).

The last snapshots of Q0, Q1, and Q2 classical MD simulation trajectories were used as the initial structures for the QM/MM simulations for the +0, +1, and +2 PMFs, respectively. An excess proton was then added to a water molecule ~6 Å above the His37 residue. The QM atoms included His37 side chains and up to three solvation shells of water molecules above and below His37. In all simulations, the QM box size was chosen to be 8 Å larger than the actual size of the QM atoms in each dimension. A quadratic confining potential was applied to restrain the QM atoms within the QM box, with the wall skin thickness being 2 Å × 2 Å × 2 Å. The QM region was treated by Becke–Lee–Yang–Parr level density functional theory (74, 75) with empirical dispersion corrections (76), under the Gaussian plane wave scheme (77). The Goedecker–Teter–Hutter pseudopotentials (78) were used and the Kohn–Sham orbitals were expanded in the Gaussian TZV2P basis set. The electron density was expanded by auxiliary plane wave basis set up to 360 Ry. The Gaussian expansion of the electrostatic potential scheme was used to treat the QM/MM electrostatic coupling with periodic boundary conditions (79, 80), and the spurious QM/MM periodic image interactions were decoupled as described in ref. 81. The equation of motion of the nuclei was integrated using a time step of 0.5 fs, and the wavefunction was optimized to the Born–Oppenheimer surface by an orbital transformation method (82) with a convergence criterion of 10^-6. The temperature was controlled at 308 K by a Nosé–Hoover thermostat with a relaxation time constant of 0.1 ps. The initial configurations of all windows were prepared from the previous equilibrated window with smaller z value. With a ~0.25 Å window spacing and force constants of 40 kcal mol^-1 Å^-2, ~40 umbrella windows were simulated for each of the QM/MM PMFs, respectively. All windows in this narrow region of the Q2 channel had an equilibration time of ~2 ps and production sampling of ~15 ps. The excess proton CEC position was collected every step (0.5 fs). All QM/MM simulations were performed using the CP2K package (83).

**Reactive MD Umbrella Sampling.** The PT PMF in the regions outside the His37–Trp41 tetrad region (Fig. 1A) was calculated using the MS-RMD method (84–87) implemented with the LAMMPS MD package (lammmps.sandia.gov) (88). Umbrella sampling was used and the same CV was used, except that in the MS-RMD simulation the excess proton CEC is given by (89)

\[ \xi_{\text{EC}} = \sum r^2_{\text{foc}}, \]

where \( r_{\text{foc}} \) is the coordinates of the center of charge of the i-th diabatic state.


It was verified that the MS-RMD and QM/MM CEC definitions were essentially identical in the overlapping regions (i.e., when the CEC is nearly three solvation shells away from the His37–Trp41 tetrad). The MS-RMD model 3.1 version (83) was used with SPC-Fw water model to describe the hydrated excess proton (90). The cutoffs for LJ and real-space electrostatic interactions were 12 Å, using a switching function starting at 10 Å for the LJ interactions. Long-range electrostatics were treated by the particle–particle particle–mesh method (91) with an accuracy threshold of 10^-10. The integration time step was 1 fs. The temperature was controlled at 308 K by Nosé–Hoover thermostat with a relaxation time constant of 0.1 ps.

The initial configuration for each umbrella window was prepared by replacing a water molecule with a hydronium close to the center of the window. With a ~0.5 Å window spacing and force constants of 10 kcal mol^-1 Å^-2, 75–80 windows were simulated for each PMF. The replica-exchange umbrella sampling (REUS) method (92) was used to facilitate convergence, with an exchange attempt frequency of 1 ps^-1. All windows were equilibrated for ~100 ps and sampled with REUS for ~12 ns. It should be noted that in the wider regions of the M2 channel longer sampling times are required in the umbrella windows. The CV value was collected every 10 fs.

**Calculation of Proton Conductance.** The final PMFs for PT across the entire channel were generated by combining the umbrella windows in the MS-RMD and QM/MM zones (Fig. 1A) using the weighted histogram analysis method (93–95), similar to what was done in ref. 16. The statistical errors were estimated by a block-average analysis.

The maximum proton conductance, \( g_{\text{max}} \), for proton permeation through M2 in different charge states was calculated from NPN electrodiffusion theory (96–98) via the formula

\[ g_{\text{max}} = \frac{e^2}{k_B T C} \int (z_2 - z_1) \rho(z) dz, \]

where \( e \) is the elementary charge, \( k_B \) is Boltzmann’s constant, \( T \) is the temperature, \( L \) is the length of the M2 channel (\( L = 45 \) Å for systems setup with crystal structure and \( L = 54 \) Å for system setup with NMR structure because the amphipathic helices of the latter extends the proton pathway within the protein), \( \rho(z) \) is the density of states, \( C \) is the diffusion coefficient and the PMF as a function of the CV for proton translocation through the channel. The brackets in Eq. 5 denote spatial averaging over the length of the channel in the direction of the channel axis (z axis). The position-dependent diffusion coefficient for the excess proton was evaluated by the Woff-Roux equation (99):

\[ D(z) = \lim_{\Delta z} \frac{z}{\Delta z} \frac{\rho(z, z + \Delta z)}{2} \int_{C(z_1)}^{C(z_2)} \frac{\rho(z) dz}{C(z)}, \]

where \( C(z_1) = C(z_2) = 0 \), is the velocity autocorrelation function for the CV calculated based on the data from window \( i, z_i \) is the reference point for the harmonic restraint potential in window \( i \), and \( C(z) = \int_0^{\infty} e^{-\gamma t} C(t) dt \) is the Laplace transform of this function. To estimate the value of the limit as \( \Delta z \rightarrow 0 \), we linearly extrapolate from the range \( 5 \leq z \leq 15 \) to obtain more converged diffusion constants in the His37–Trp41 region, the computationally more efficient MS-RMD simulation was used instead of the more expensive and time scale-limited QM/MM simulation. Due to space constraints, we refer to our previous paper (16) for the discussion on the limitations of the method used to calculate the PMFs and conductances.

**ACKNOWLEDGMENTS.** This research was supported by National Institutes of Health Grants R01-GM053148 (to G.A.V., J.M.J.S., and R.L.), R01-GM056432 (to C.E.H. and C.M.C.), and a Carlsberg Foundation International Fellowship (J.M.J.). The researchers used computing facilities provided by the Extreme Science and Engineering Discovery Environment, which is supported by National Science Foundation Grant OCI-1053575, as well as by the University of Chicago Research Computing Center and the US Department of Defense High Performance Computing Modernization Program.


