Free energy and kinetics of cAMP permeation through connexin26 hemichannel with and without voltage

Wenjuan Jiang[#1], Yi-Chun Lin[#1], Wesley Botello-Smith [#1], Jorge Contreras*[2], Andrew L. Harris*[2], Luca Maragliano*[3,4], Yun Lyna Luo*[1]

1. College of Pharmacy, Western University of Health Sciences, 309 E. Second St, Pomona, CA, USA.
2. Department of Pharmacology, Physiology, and Neuroscience. New Jersey Medical School, Rutgers, The State University of New Jersey, Newark, NJ, USA.
3. Department of Life and Environmental Sciences, Polytechnic University of Marche, Ancona, Italy
4. Center for Synaptic Neuroscience and Technology, Italian Institute of Technology, Genoa, Italy.

[#] contribute equally
* corresponding authors

Abstract

The connexin family is a diverse group of highly regulated non-β-barrel wide-pore channels permeable to biological signaling molecules. Despite their critical roles in mediating selective molecular signaling in health and disease, the molecular basis of permeation through these pores remains unclear. Here, we report the thermodynamics and kinetics of binding and transport of a second messenger, adenosine-3',5'-cyclophosphate (cAMP), through a connexin26 hemichannel. Inward and outward fluxes of cAMP were first obtained from 4 μs simulations with voltages and multiple cAMPs in solution. The results are compared with the intrinsic potential of mean force (PMF) and the mean first passage times (MFPTs) of a single cAMP in the absence of voltage, obtained from a total of 16.5 μs of multi-replica Voronoi-tessellated Markovian milestoning simulations. The computed transit times through the pore correspond well to existing experimental data. Both voltage simulations and milestoning simulations revealed two cAMP binding sites with binding constants and dissociation rates computed from PMF and MFPTs. The protein dipole inside the pore produces an asymmetric PMF, reflected in unequal cAMP MFPTs in each direction once within the pore. The free energy profiles under voltages derived from intrinsic PMF provided a unified understanding of directional transition rates with/without voltage, and revealed the unique role of channel polarity and the mobile electrolyte within a wide pore on the total free energy. In addition, we show how these factors influence the cAMP dipole vector during permeation, and how cAMP affects the local and non-local pore diameter in a position-dependent manner.

1

Significance Statement

Connexins are wide-pore channels permeable to cellular signaling molecules. They mediate molecular signaling crucial in physiology, pathology, and development; mutations in connexins cause human pathologies. However, the fundamental structural, thermodynamic, and kinetic determinants of molecular permeability properties are unknown. Using multiple molecular dynamics simulation techniques, we report, for the first time, an in-depth investigation of the free energy and the directional transition rates of an important biological signaling molecule, cAMP, through a connexin channel. We reveal the energetics and binding sites that determine the cAMP flux, and the effects of mobile ions and external electrical field on the process. The results provide a basis for understanding the unique features of molecular flux through connexins and other non-β-barrel wide-pore channels.

Introduction

Connexin proteins form wide channels that mediate electrical and molecular signaling through cell membranes. They can function as plasma membrane channels (called "hemichannels) or as intercellular channels that allow direct transfer of small cytoplasmic molecules between cells. The intercellular channels ("gap junction channels") are formed by end-to-end docking of two hemichannels across the extracellular gap between adjacent cells. The pores are relatively wide and are therefore permeable to atomic ions and small molecules in the size-range of key cellular signaling molecules including cAMP, cGMP, ATP, IP$_3$, and glutathione. Each of the 21 human connexin isoform forms channels with distinct permeability and regulatory properties(*1, 2*). With the exception of electrical signaling in excitable tissues, the primary biological function of connexin channels is to mediate movement of small cytoplasmic signaling molecules between cells and/or to release them into the extracellular environment in a highly regulated manner. Mutations in connexins that alter channel function or expression produce human pathologies(*3, 4*). These mutations ultimately exert their pathological effects by disrupting the proper molecular permeability of junctional and plasma membranes that connexin channels mediate.

A large literature documents that channels formed by the different connexin isoforms have dramatically different permeability properties. Their unitary conductances range from 10 pS to 300 pS, their cation/anion permeability ratios (P_{K+}/P_{Cl-}) range from 8.0 to 0.8 and their

67   permeabilities to fluorescent tracers are highly disparate. Strikingly, none of these parameters

68   correlate with each other (e.g., the connexin channel with the largest unitary conductance is among

69   the most size-restrictive). Permeabilities to biological signaling molecules are strikingly different

70   among the different connexin channels, but are difficult to measure quantitatively, as the molecules

71   are not fluorescent and do not carry significant current. Furthermore, for a given type of connexin

72   channel, there are remarkable degrees of selectivity and relative permeability among biological

73   permeants. These connexin-specific and permeant-specific permeability properties are not

74   reasonably inferred from differences in permeabilities to fluorescent tracers(*5*). This suggests that

75   there are, as yet unknown, structural/energetic determinants of molecular permeation that impart

76   to each type of connexin channel specific, biologically required permeability properties. The

77   underlying mechanisms for this are unknown, in spite of their clear biomedical and therapeutic

78   importance. Investigation of these mechanisms by mutagenesis has not been informative in the

79   absence of an understanding of the behavior, energetics, and interactions experienced by a

80   molecule as it traverses the pore. Computational studies can provide the basis for this

81   understanding.

82   We previously used Hamiltonian replica-exchange umbrella sampling to explore the

83   energetics of uncharged permeant and non-permeant tracer molecules in the connexin 26 (Cx26)

84   hemichannel(*6*). That study indicated that the determinants of molecular permeation differ from

85   those that dominate the permeation of atomic ions through commonly studied ion-selective

86   channels, emphasizing the unique aspects of small molecules with conformational and

87   orientational degrees of freedom in a wide pore. The free energy and calculated relative

88   permeabilities derived from that work were consistent with experimental findings. However, while

89   validating the atomistic system and overall computational approach, the previous work did not

90   provide quantitative kinetic information of the permeation process. Furthermore, the results of

91   neutral tracer molecules do not characterize the biomedically crucial process of permeation by

92   charged biological signaling molecules. The present study explores the binding and transport

93   kinetics of a negatively charged second messenger, adenosine-3',5'-cyclophosphate (cAMP),

94   permeates a connexin pore.

95   A single permeation event with a timescale of sub-microsecond to microseconds is within

96   reach of today's computational power. However, a large number of transition events are required

97   to obtain meaningful statistics, for which simulations of orders of magnitude longer than the mean

98   transition time are required. One solution is to accelerate the permeation of charged molecules by

99   imposing a voltage (cAMP carries a charge of -1e). Theoretically, if the system reaches a steady

100 state under a constant electric field and maintains symmetric concentration on both sides of the

101 channel, a mean flux rate can be estimated from the ensemble of nonequilibrium processes using

102 a large number of permeation events. The accumulated density of cAMP along the channel axis

103 during these events may provide an estimate of the locations of the energetic barrier(s) and binding

104 site(s). Alternatively, one can choose an enhanced sampling method that is suitable for computing

105 free energy and kinetics of the permeation process. In this study, we used both approaches to gain

106 a comprehensive understanding of cAMP permeation with and without voltage, and in presence

107 and absence of multiple permeants.

108   A molecular permeation rate is often dominated by the free energy profile or potential of

109 mean force (PMF) along the channel lumen axis. Previously we estimated the relative transition

110 rates between two sugar molecules using transition state theory (TST) based on PMF profiles(*6*).

111 However, TST requires assumptions such as a single dominant transition state and no re-crossing

112 at the barriers, which are often difficult to satisfy in complex biomolecular systems with rugged

113 free-energy landscapes. To overcome this limitation, enhanced sampling techniques have been

114 developed to calculate transition rates from molecular dynamics (MD) simulations. Of particular

115 interest is the milestoning method introduced by Faradjian and Elber(*7*), which has been developed

116 into several versions and used in many biophysical applications(*8*). Voronoi-tessellated Markovian

117 milestoning is an implementation that allows reconstruction of the long-time dynamics of a system

118 from independent simulations confined within a set of cells spanning the space of the reaction

119 coordinates(*9*). This method has successfully captured the rates of CO entry/exit in myoglobin(*10*),

120 and recently the ligand binding kinetics(*11*). Here, we make use of the "soft-walls" version, which

121 confines the sampling within the Voronoi cells using flat-bottom harmonic restraining

122 potentials(*12*). This approach is easy to implement, allowing us to take advantage of CUDA

123 accelerated MD packages, and has been shown to yield the same results as the original "hard-walls"

124 version, which instead inverts atomic velocity at the cell boundaries. The "soft-walls" version has

125 been used to study nucleation of an ionic liquid(*13*) and ion permeation across a claudin-15

126 paracellular channel(*14*). Here we adopt this approach to explore the permeation of cAMP through

127 a Cx26 hemichannel.

128

Results

cAMP permeation under opposite voltages

We first investigated inward/outward cAMP permeation rates under voltages in the presence of symmetric cAMP concentrations. To obtain a sufficient number of transition events and ensure unidirectional flux, we applied +/-200 mV transmembrane potentials to accelerate permeation. 27 cAMP molecules and 27 Mg^{2+} ions, corresponding to 26.5 mM, were present on each side of the membrane. To mimic physiological salt concentration and also to neutralize the protein charge, 82 K^+ and 163 Cl^- ions were added to the bulk aqueous compartments, corresponding to 80.5 mM K^+ and 160.1 mM Cl^- (**Table S1**). Using z<|50| Å as boundaries of the hemichannel (center-of-mass of protein is at z=0 Å), 9 single-molecule transitions were observed at +200 mV and 12 transitions at -200 mV during each 2 μs simulation (**Figure 1, Table 1,** and raw data in **Table S2**). At each voltage, all transitions were in the same direction.

To estimate the uncertainty due to the small number of events, we calculated the confidence interval for the average transition time ($<t>$) of cAMP through the channel by fitting the transition times to an exponential distribution. Based on a maximum likelihood estimation, MFPT is 448 ns at +200 mV with 95% confidence interval (CI_{95}) of 204-1646 ns, and 510 ns with CI_{95} of 249-1571 ns at -200 mV. The average time between consecutive transition events (τ = length of simulation/number of transits) is 222 ns at +200 mV and 166 ns at -200 mV. Given the confidence intervals, there is no significant difference in the flux in each direction. The ratio $<t>/\tau$ is 2.0 and 3.1, for +200 mV and -200 mV, respectively, which indicates there are on average two to three cAMP molecules in the channel at any given time at each voltage. Inspection of the trajectories shows that this is only due to the accumulation of cAMP molecules at the intracellular entrance (**Figures 2 and 3**).

Although the experimental transition time of cAMP under voltage is not available, a cAMP/K^+ permeability ratio of 0.027 was reported using simultaneous measurements of Cx26 junctional conductance and reporter-based intercellular transfer of cAMP(*15*). We calculated K^+ transition time in presence of 27 cAMP is 13.1 ± 13.0 ns at +200 mV and 6.8 ± 7.8 ns at -200 mV. If we assume the time needed to cross the junctional channel (two hemichannels docked at the

157   extracellular ends) is the sum of the transition time in two opposite directions, we obtain a

158   cAMP/K$^+$ ratio of 0.021, reasonably close to the experimental ratio.

159
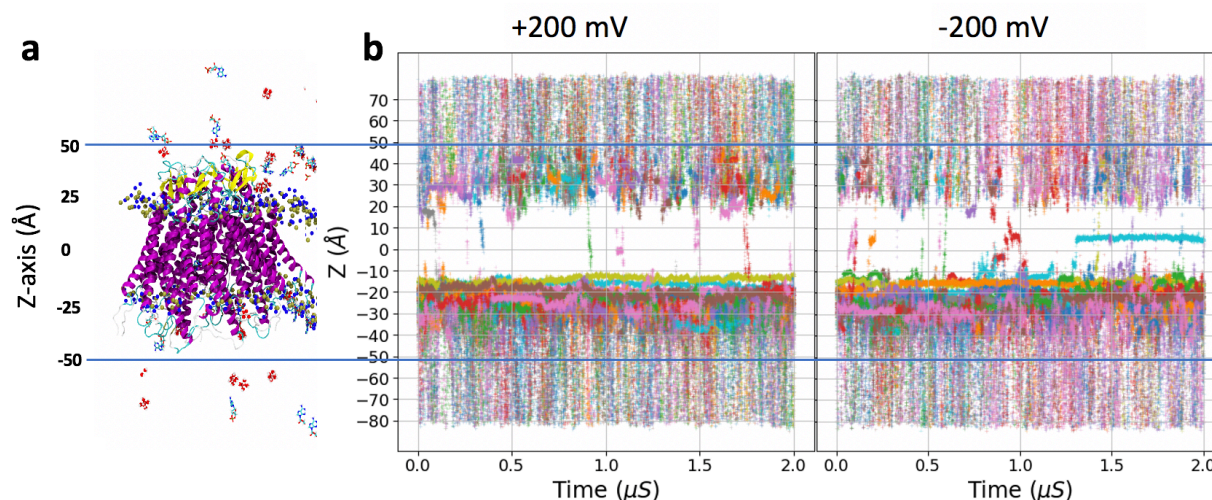
160

161

162   **Figure 1. Simulations of cAMP permeation through Cx26 at +200 mV and -200 mV membrane**
163   **potential. a.** Snapshot of the simulated system. Protein backbone is shown in cartoon mode and colored by
164   the secondary structure (helix in magenta, beta-sheet in yellow, disordered loop in cyan). cAMP molecules
165   are shown in licorice with atoms colored (red oxygen, cyan carbon, blue nitrogen, yellow phosphate). Lipids,
166   ions, and water molecules are not shown. **b.** The z-coordinates of all 27 cAMP molecules in the system
167   during the simulations are shown in different colors. Note that the flux of cAMP is in the direction opposite
168   to the field; thus, cAMP flux is inward (down in this figure) under +200 mV and outward under -200 mV.

169

170

171   **Table 1. Transition time of cAMP through Cx26 hemichannel under voltage.**

	+200mV	-200mV
cAMP flux direction	Inward	Outward
Transition events in 2 μs	9	12
Events per μs	4.5	6.0
Time between transition events	222 ns	166 ns
Mean transition time (CI$_{95}$) *	448 ns (204-1646)	510 ns (249-1571)
Mean dwell time (CI$_{95}$)	305 ns (149-940)	398 ns (205-1086)
Mean barrier crossing time (CI$_{95}$)	143 ns (79-330)	106 ns (54-290)

172
173   *Sample mean and 95% confidence intervals are based on Maximum Likelihood Estimate by fitting the
174   exponential   distributions.   See   code   at   https://github.com/LynaLuo-Lab/MD-data-uncertainty-
175   analysis/blob/master/confidence_interval_exponential.ipynb.

176

177

Pore cAMP density profiles from voltage simulations

The density profiles accumulated from voltage simulations can provide an estimate of the locations of energetic barriers and binding sites. The charge densities of protein, lipids, total ions, and individual ions (K^+, Cl^-, $cAMP^-$, Mg^{2+}) are plotted along the channel z-axis (**Figure 2a**). It can be seen that the channel is largely positive at intracellular entrance where the highest density of Cl^-, $cAMP^-$ ions are located. The accumulation of cAMP at the intracellular entrance can also be seen clearly in **Figure 1.** A volumetric map showing the 3D density of cAMP inside the channel from the -200 mV simulation is illustrated in **Figure 2b**. A smooth indicator of cAMP distribution is shown using Boltzmann inversion ($-kTln(\rho)$) of the cAMP density; we will call these *log-density plots*. **Figure 2c** shows the 2D cAMP log-density along the z-axis and the radial axis *R*. **Figure 2d** shows the 1D cAMP log-density within the pore along z. Note that these density profiles are acquired from nonequilibrium simulations under the influence of voltage. They indicate how the cAMP behaves under the influence of both external voltage and the intrinsic free energy, hence do not represent the equilibrium free energy profiles. The symmetric cAMP density in the bulk regions on each side of the channel is the result of the periodic boundary condition used for MD simulations.

The 1D log-density profiles (**Figure 2d**) have a broad minimum at the intracellular entrance of the channel and a major peak in the middle of the channel. The shapes of the global minimum region are essentially identical under both voltages. There is a second well in the -200 mV density that appears to bisect the central peak. The dwell times of cAMP (**Table 1**) at the broad minimum (-50<z<10 Å) are quite large under both voltages: 305 ns at +200 mV, and 398 ns at -200 mV (**Table 1**). In contrast, the rest of the channel region (-10<z<50 Å), which contains the major barriers, was crossed by cAMP more rapidly: 143 ns for inward flux (at +200 mV), and 106 ns for outward flux (at -200 mV).

The broad binding well (-50<z<-10) present at both voltages is located in the C-terminal regions of the second transmembrane helix (TM2) and the N-terminal helix (NTH). Contact frequency analysis between cAMPs and protein sidechains indicates that cAMP bind to R99/104 and K103 on TM2 over 70% of the simulated time, and bind to K102 and R98 on TM2, and K15 on NTH over 30% of the time (**Figure 3a**), thus providing a basis for the cAMP accumulation at this region. Further investigation revealed that the cAMP forms clusters between -30<z<0 Å, while almost no clusters are found in the rest of the channel (**Figure 3b**). Most of the clusters contain 2

209   cAMPs, which interact through pi-stacking of adenosine rings. Clusters of 3 cAMPs forming

210   hydrogen bonds with each other also exist (**Figure 3c**). These clusters are not seen in the bulk. It

211   appears that the positively charged residues at the intracellular entrance of the pore facilitate cAMP

212   clustering by reducing the translational and rotational entropy of the molecules. To check whether

213   this accumulation produce a "crowding" effect on entry into the pore, we compared the unoccupied

214   lumen radius with and without cAMP molecules present. **Figure S1** suggests that the accumulation

215   of cAMP had little effect on the average available cross-sectional area at the pore entrance. In

216   addition, the peaks in the two log-density profiles (**Figure 2d**) do not closely correlate with the

217   narrow regions of the pore (see **Figure S1 and Figure 8a** for two different radius measurements),

218   indicating that steric hindrance is not a major contributor to the cAMP transition barrier.

219

220
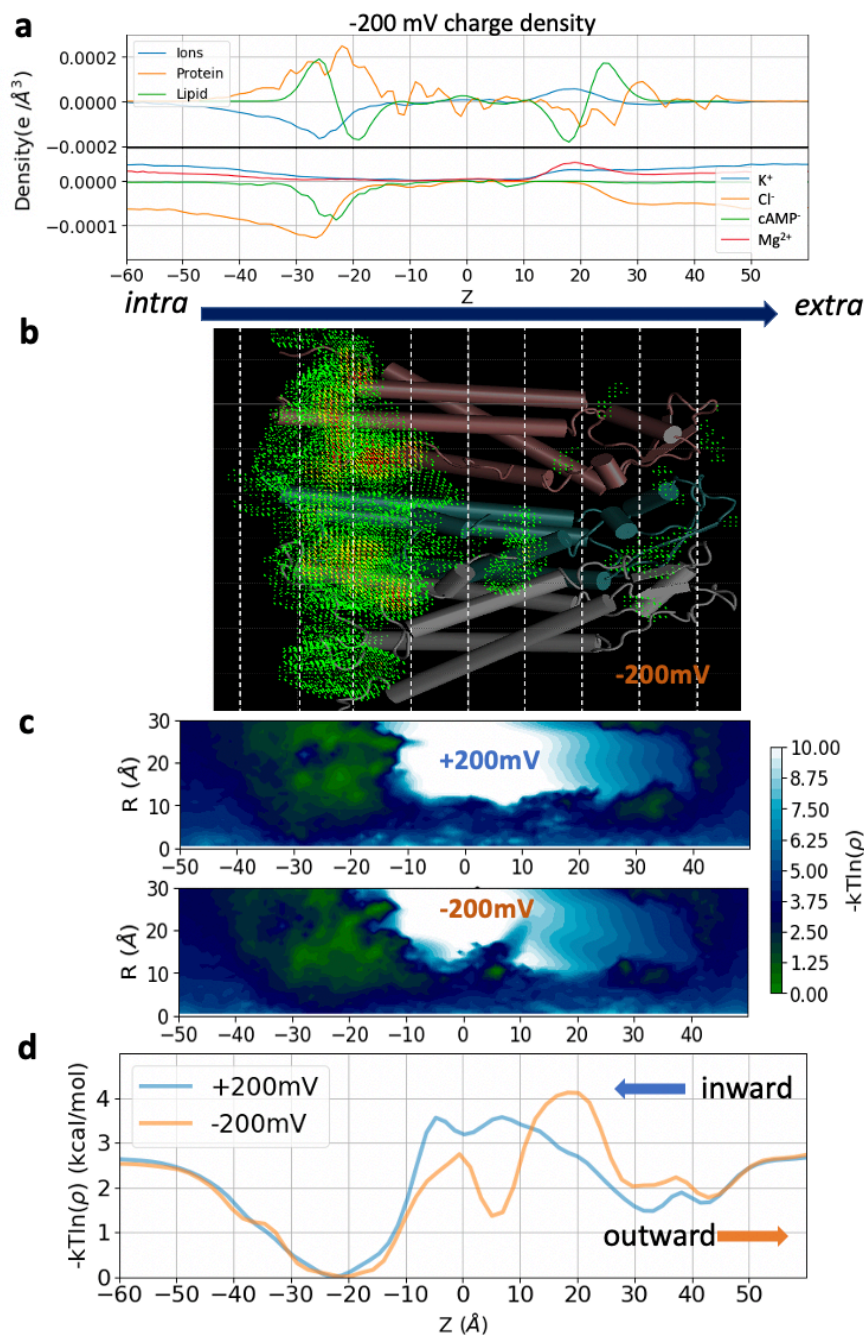
Figure 2. Charge density profile and cAMP log-density profiles in 3D, 2D, and 1D. a. Charge density profiles of protein, lipids, total ions and individual charges groups (K^+, Cl^-, $cAMP^-$, Mg^{2+}) along the channel z-axis obtained from the -200 mV trajectory. **b.** Volumetric map of cAMP 3D density under -200 mV simulation. **c.** Boltzmann inversion $-kTln(p(z))$ of the accumulated cAMP density profile in 2D along the channel z-axis and radial axis R. **d**. 1D log-density plot of the cAMP density profile within the pore along the channel z-axis. Arrows indicate the direction of cAMP flux at each voltage. Note: **a** represents charge density in the entire simulation system, while **b, c,** and **d** depicts the cAMP density only within the pore using the same cylindrical radius cutoff of 30 Å as in the milestoning simulation below.

229

9

230

231
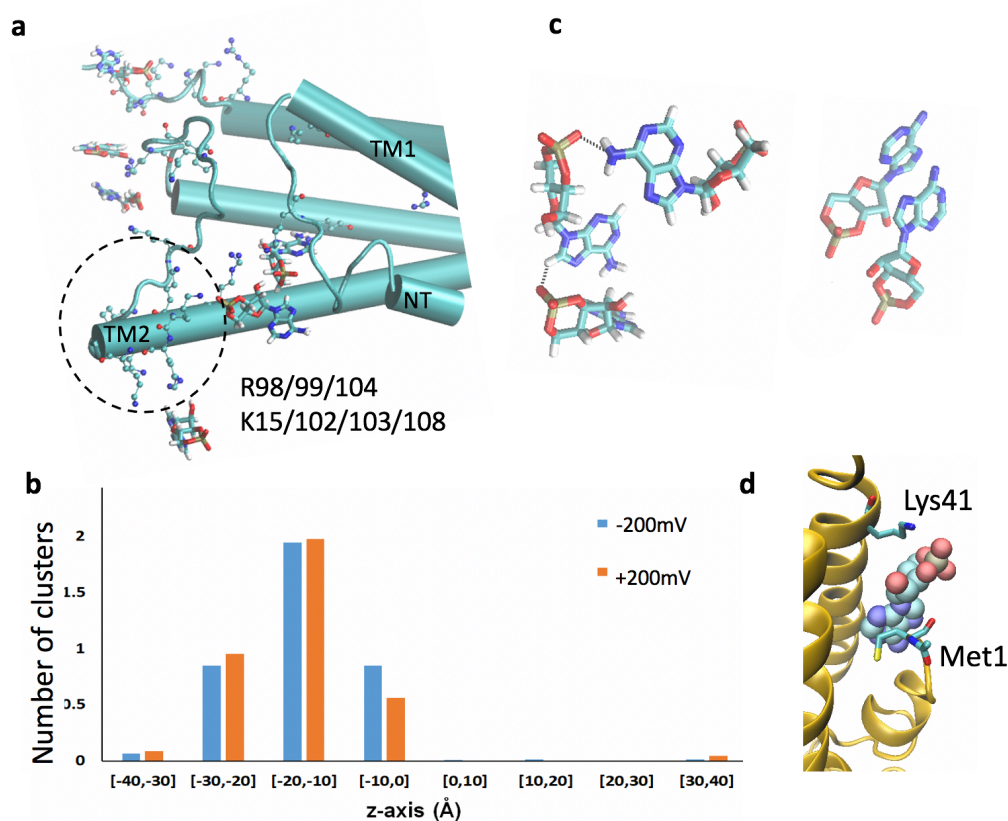
Figure 3. Permeant-permeant interactions and permeant-channel interactions. a. Snapshot showing cAMP molecules (in licorice) near positively charged residues (in CPK mode) at the intracellular entrance. Protein is in cartoon mode. **b.** Number of cAMP clusters in each 10 Å window during +/-200 mV simulations. A cluster is defined as at least two cAMP molecules with centers of mass within 20 Å at any time during the simulations. **c.** Snapshots of two cAMPs clustered via pi-stacking and three cAMPs clustered via hydrogen bonds. **d.** Snapshot showing one cAMP trapped between Lys41 and Met1.

238

239

240   During one of the 12 permeation events at -200 mV simulation, one cAMP was trapped

241 between Met1 in the N-terminal helix (NTH) and Lys41 on the first transmembrane helix (TM1)

242 (**Figure 1b, 3d**). This was reflected in the density profile as a large dip in the peak at z=5 Å (**Figure**

243 **2d**). Both NTH and Lys41 have been suggested to be involved in voltage-sensing in Cx26(*16, 17*).

244 This raised the question of whether the instance of a long residency of a cAMP molecule at this

245 particular position at -200 mV, but not evident at +200 mV or 0 mV (from PMF milestoning below)

246 is a consequence of voltage-driven repositioning of these charged moieties. To evaluate their

247 responses to the local electric field, we plotted the angles between the z-axis and the principal

248 vector of Lys41 or NTH (residues 1 to 11) during 2 µs simulations at +/-200 mV (**Figures S2a**

249 **and S2b**). Except for a clear reduction in Lys41 fluctuation in subunit 5, where the cAMP was

250 trapped, there is no clear preference in the orientation of the Lys41 or NTH in response to the two

251 opposite voltages. Therefore, this trapped cAMP is unlikely to be due to the effect of voltage on

252 the protein. Of course, this result does not indicate that the NTH and Lys41 are uninvolved in

253 voltage-sensing, only that they did not respond to +/-200 mV within the 2 µs simulations.

254

## cAMP permeation free energy using Markovian milestoning

256 The simulations above provide a nonequilibrium view of the cAMP transition process driven by

257 voltage. Two opposite voltages resulted in similar cAMP density profiles and similar ranges of

258 flux rates at each voltage. However, the small number of stochastic events (21 in total) led to the

259 large uncertainty in the mean passage time of the transition event. In contrast to a long trajectory

260 exploring the whole channel, multiple MD simulations confined in intervals partitioning the space

261 can offer sufficient statistics within a shorter running time. Here, we used Voronoi-tessellated

262 Markovian milestoning MD simulations (hereafter referred to as milestoning simulation) on a

263 tessellation along the z-coordinate of the center-of-mass of a single cAMP to estimate the PMF

264 and the kinetics of permeation through the channel at zero membrane voltage.

265 **Figure 4** shows the 1D PMF (in red) obtained from the equilibrium probability of finding

266 cAMP in each milestoning cell. This PMF of a single cAMP permeating without applied voltage

267 has features similar to the log-density plots of the accumulated density of cAMP from the voltage

268 simulations. A free energy barrier spans between $-20<z<30$ Å, with the peak located at $z = +20$ Å,

269 the same as the peak in log-density from the -200 mV simulation. The height of the barrier relative

270 to bulk (1.2 kcal/mol) is slightly lower than in -200 mV log-density (1.8 kcal/mol) and higher than

271 in +200 mV (0.8 kcal/mol). The major well on around $z = -20$ Å from voltage simulations are

272 much broader and more favorable (-2.8 kcal/mol relative to bulk) than the well in the PMF (-2

273 kcal/mol), likely due to the cAMP clustering, which is absent in the single cAMP milestoning. The

274 dip that split the peak at $z = +5$ Å under -200 mV also shows up in the PMF, but with much smaller

275 magnitude. It is thus possible that the magnitude of the dip at -200 mV in the log-density plot is

276 overestimated due to the contribution of the single, rare event of cAMP trapping (evident in **Figure**

277 **1b**).

278   Similar to the voltage simulations, milestoning simulation without voltage reveals two

279   binding sites for cAMP, a major one (site 1) near the intracellular entrance between NT and TM2

280   (-2 kcal/mol relative to bulk value) and a (bisected) much smaller one (site 2) at the extracellular

281   loop (E1) region (-0.7 kcal/mol relative to bulk). The single cAMP dissociation constant, K_D, can

282   be estimated from the single cAMP equilibrium PMF using Eq. 1:

$$1/K_D = \pi R^2 \int_{z_{min}}^{z_{max}} dz \; e^{-W(z)/k_B T}$$

283   Eq. 1

284    $W(z)$ is the PMF with bulk as reference. R is the radius of a cylindrical restraint (30 Å). Integrals

285   over individual energy wells indicate that cAMP will bind to the intracellular site (binding site 1:

286   -43.4<z<18.6 Å) with K_D of 1.8 mM and to the extracellular site (binding site 2: 18.6<z<60.6 Å)

287   with K_D of 7.3 mM (**Figure 4a**). The integral over the entire PMF yields a total K_D of 1.4 mM for

288   the channel. Thus, at the bulk cAMP concentration of the voltage simulations (26.5 mM), both

289   binding sites are likely highly occupied, while at the concentration equivalent to the single cAMP

290   used in milestoning (1 mM), none of the binding sites would be occupied more than 50%.
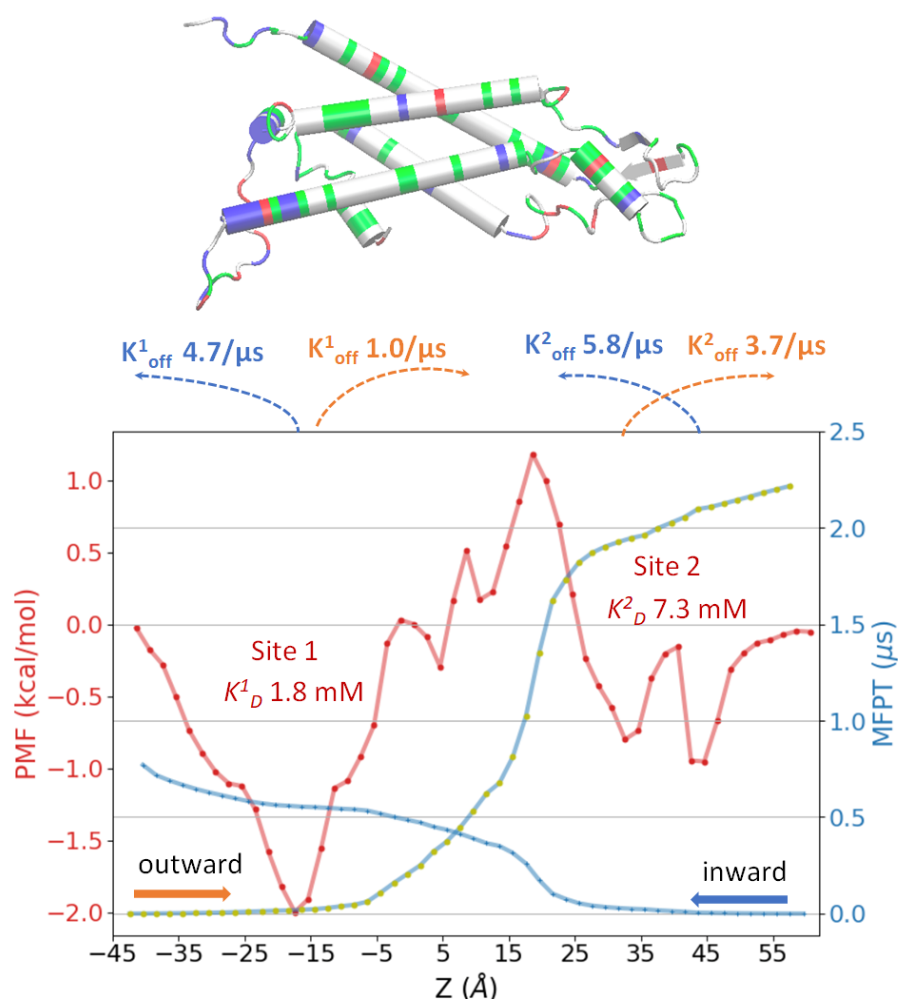
Figure 4. Free energy and kinetics of a single cAMP permeation through Cx26 hemichannel from milestoning simulation (V=0 mV). Potential of mean force (PMF) in red, inward mean first passage times (MFPT) in blue, and outward MFPT in yellow dotted blue line. Binding constants and dissociation rates derived from PMF and MFPT profiles are indicated. The graphic at the top shows the backbone of one Cx26 subunit with z-positions aligned with the plot below (basic residues in blue, acidic in red, polar in green, and nonpolar in white).

cAMP permeation kinetics using Markovian milestoning

According to the MFPT profiles in **Figure 4**, from the first to the last milestone at the boundaries of the channel, it takes about 0.77 µs for inward flux and 2.22 µs for outward flux. Thus, the inward flux is 2.9 times faster than outward flux once a single cAMP enters the channel. The faster inward than outward flux is the direct consequence of the asymmetric channel, reflected in the asymmetric PMF showing a maximum of 3.2 kcal/mol barrier for outward flux, which takes

13

306   1.0 μs to cross, while the two smaller barriers of 2.0 kcal/mol for inward flux only take 212 ns (site

307   1) and 145 ns (site 2) to cross. MFPT profiles can provide direct information about the dissociation

308   rate (k_{off}) of a single cAMP. For outward permeation, k_{off}^1 for binding site 1 is 1.0 μs^{-1} (-17.4

309    $<z<18.6$) and k_{off}^2 for binding site 2 is 3.7 μs^{-1} (32.6 $<z<60.6$). For inward permeation, k_{off}^1 is 4.7

310   μs^{-1} (-41.4 $<z<-17.4$) and k_{off}^2 is 3.7 μs^{-1} (18.6 $<z<42.6$). With the equilibrium constant K_D from

311   PMF, we can also estimate the association rate k_{on}^1 of 0.6 μs^{-1}mM^{-1} and k_{on}^2 of 0.5 μs^{-1}mM^{-1} for

312   outward permeation, and k_{on}^1 of 2.6 μs^{-1} mM^{-1} and k_{on}^2 of 0.8 μs^{-1} mM^{-1} for inward permeation.

313   It should be noted that the MFPT is subjected to the condition that the channel is occupied

314   by only one permeating molecule at any time. Thus, it only represents the transit time of a single

315   permeant traversing an otherwise permeant-free channel. The full kinetics and flux at finite bulk

316   concentrations of the permeant also depend on bulk diffusivity and concentration, and the

317   diameters of the entrance at each end of the pore. For instance, diffusion current to a disk-like

318   adsorber is $\mathbf{I=4DRC}$, where $\mathbf{C}$ is the permeant concentration in the infinite bulk, $\mathbf{D}$ is bulk diffusion

319   constant, and $\mathbf{R}$ is the radius of the disk-shaped absorber[18]. Interestingly, for Cx26 hemichannel,

320   the intracellular entrance (radius $\mathbf{R}$=25 Å) is larger than the extracellular entrance ($\mathbf{R}$=10 Å) (see

321   pore radius in Figure 8a). Taking into account the effective radius of cAMP as r=3 Å (calculated

322   from radius of gyration), the effective pore radii are 22 Å and 7 Å for the intracellular and

323   extracellular pore entrances, respectively, indicating that it is ~3 fold more likely for cAMP to

324   reach the pore by random diffusion from bulk to the intracellular than the extracellular side.

325   Quantitative experimental measurements of the flux of cAMP through Cx26 channels in

326   cells are very complex and subject to a variety of potential confounding factors (5). Two studies,

327   which used different indirect strategies to report cAMP flux through junctional channels in the

328   absence of junctional voltage, yielded estimates of cAMP permeability that differed by nearly a

329   factor of 8 (6.2 and 47 x 10^{-3} um^3 sec^{-1}) (15, 19). Using the volume of the Cx26 channel obtained

330   from the simulations, these permeability predict cAMP transit rates through junctional channels of

331   14 μs and 1.9 μs, respectively, which bracket the ~3 μs transit time inferred from our studies (sum

332   of hemichannel "outward" and "inward" transit times of 0.77 and 2.22 μs).

333

334

335 **Influence of voltages on free energy profile**

336 The milestoning simulations at zero voltage yielded an inward transition time is 2.9 times faster

337 than the outward transition time (0.77 vs 2.22 μs). Interestingly, the mean inward/outward barrier

338 crossing time under voltages are quite similar: 143 ns (CI$_{95}$ 79-330 ns) at +200 mV, and 106 ns

339 (CI$_{95}$ 54-290 ns) at -200 mV (**Table 1**). Below we show that this is likely due to the negative

340 voltage reducing the free energy barrier for cAMP outward flux.

341 It has been shown previously that the total PMF under voltage $W_{tot}(z)$ can be computed

342 as the sum of the intrinsic PMF in the absence of external field $W_{eq}(z)$ (in this case the equilibrium

343 PMF derived from milestoning), and the additional potential introduced by the external field

344 $q\delta\phi(z)$ (Eq. 2)(20-22). This additional potential has two components. One is the constant electric

345 field throughout the entire simulated periodic cell $E= V(z)/L_Z$, where $V(z)$ is the voltage linear to

346 L_Z, the length of the PBC box in the z-direction. The other component is the reaction potential due

347 to the voltage-induced changes of the spatial distribution and orientation of the water dipole and

348 mobile ions, as well as flexible and charged atoms in the protein and membrane(23). This approach

349 presumes that the channel and the permeant do not undergo substantial conformational changes

350 due to the external field within the time of simulation (2 μs in this case). This assumption is

351 supported by our results from +/- 200mV simulations showing highly similar cAMP density

352 distribution (**Figure 2**) and pore radius profile (**Figure S1**).

353

354 $$W_{tot}(z) = W_{eq}(z) + q\delta\phi(z) = W_{eq}(z) + q(V(z) + \phi_v(z) - \phi_0(z))$$

355 Eq. 2

356 The reaction potential introduced by external field may be approximated by the difference

357 in electrostatic potential in presence and absence of the external potential, $\phi_v(z) - \phi_0(z)$. To

358 calculate this difference in electrostatic potential, three additional 100 ns simulations at -200, 0,

359 +200 mV were carried out for the same system but without cAMP. 3D electrostatic potential maps

360 $\phi(r)$ were calculated based on all charged atoms in the simulated system $\rho_i(r)$ by solving the

361 Poisson equation $\nabla^2\phi(r)=-4\pi\sum_i\rho_i(r)$ on a 1 Å resolution grid using the VMD PMEPot plugin(24).

362 **Figure 5a** and **5b** show the 1D potential along the central pore z-axis without voltage (in blue) and

363 under +/-200 mV (in orange).

364

365
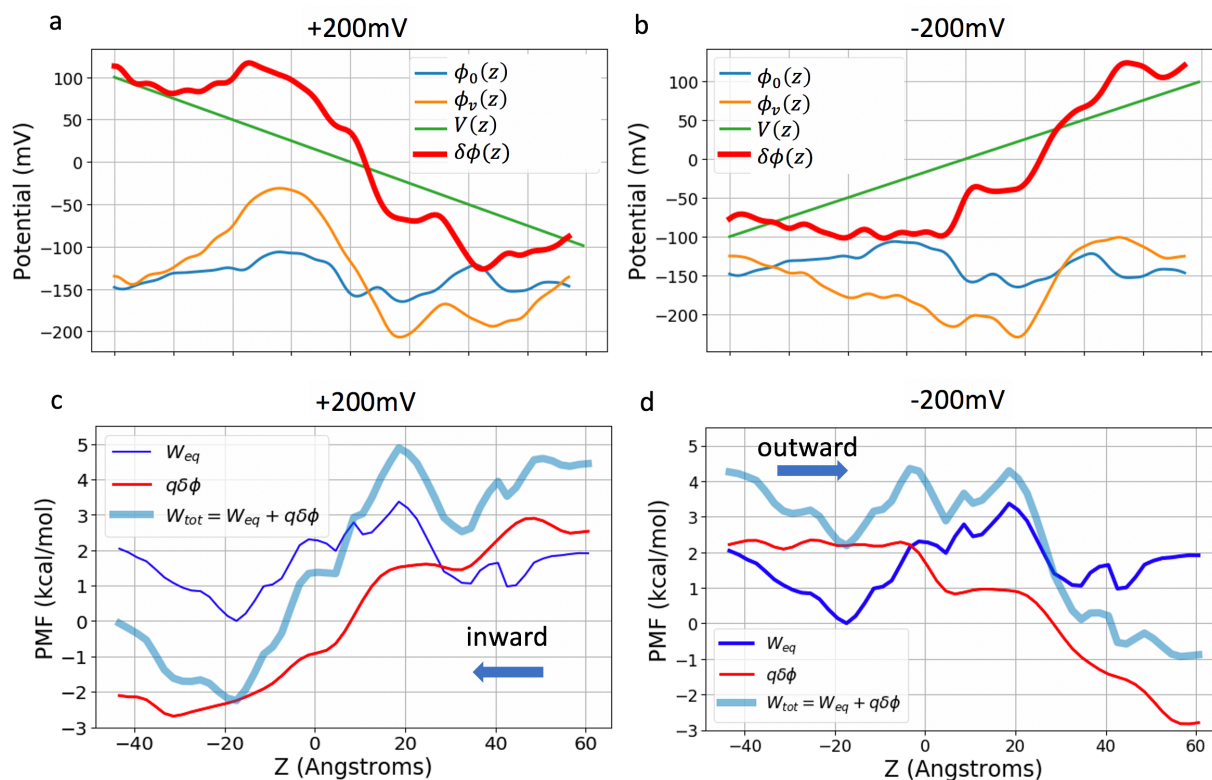
366

367 **Figure 5. Electrostatic potentials and PMFs under voltages. a,b**: applied potential in green (+/-200 mV)
368 and the electrostatic potential under voltage $\phi_v(z)$ in orange, and without voltage $\phi_0(z)$ in blue. The sum of
369 applied potential and reaction potential (see Eq. 2) in shown in red. **c,d**: Total PMF under a constant electric
370 field, $W_{tot}(z)$, from the intrinsic PMF in absence of an electric field $W_{eq}(z)$, and the additional electrostatic
371 potential energy introduced by the applied field $q\delta\phi(z)$.

372

373

374   Cx26 is largely positive in the intracellular side (-40<z<0), negative near the extracellular
375 side (15<z<25), and slightly positive near extracellular entrance (25<z<35) (**Figure 2a**), thus it
376 has an overall dipole vector pointing towards the positive z-direction. Consequently, within the
377 protein-membrane region, positive voltage produces a potential $\phi_v(z)$ that enhances the protein
378 dipole, while the one from negative voltage counters the protein dipole (**Figure 5ab** orange lines).
379 In the bulk region and lipid headgroup region ($z > |15|$), the $\phi_v(z)$ nearly cancels out the external
380 field $V(z)$. The sharp drop of the $\phi_v(z)$ under +200 mV can be visualized on 2D-electrostatic
381 potential maps of the whole system (**Figure 6a**), and the 3D-electrostatic potential overlaid onto
382 the solvent mass density iso-surface (**Figure 6b**) or onto the cAMP density iso-surface (**Figure
383 6c**).

16

384   **Figure 5 cd** show the total PMF under each voltage $W_{tot}(z)$ obtained from Eq. 2. Clearly,

385   +200 mV facilitates the inward flux of the negatively charged cAMP by increasing the free energy

386   on the extracellular side of the protein (z>0) and decreasing the free energy on the intracellular

387   side (z<0). The two inward flux energy barriers remain similar to those of the intrinsic PMF (~2.3

388   kcal/mol). However, the outward $W_{tot}(z)$ at -200 mV significantly reduced the outward barrier

389   from 3.2 kcal/mol of the intrinsic PMF to 2.3 kcal/mol. Thus, the asymmetry of the inward *vs.*

390   outward MFPTs of cAMP within the pore at zero voltage is eliminated by the external voltage.

391   This "voltage-equalizing" effect of permeation kinetics is a unique feature of large-pore channels

392   that have a protein dipole and mobile electrolytes inside the pore.

393

394
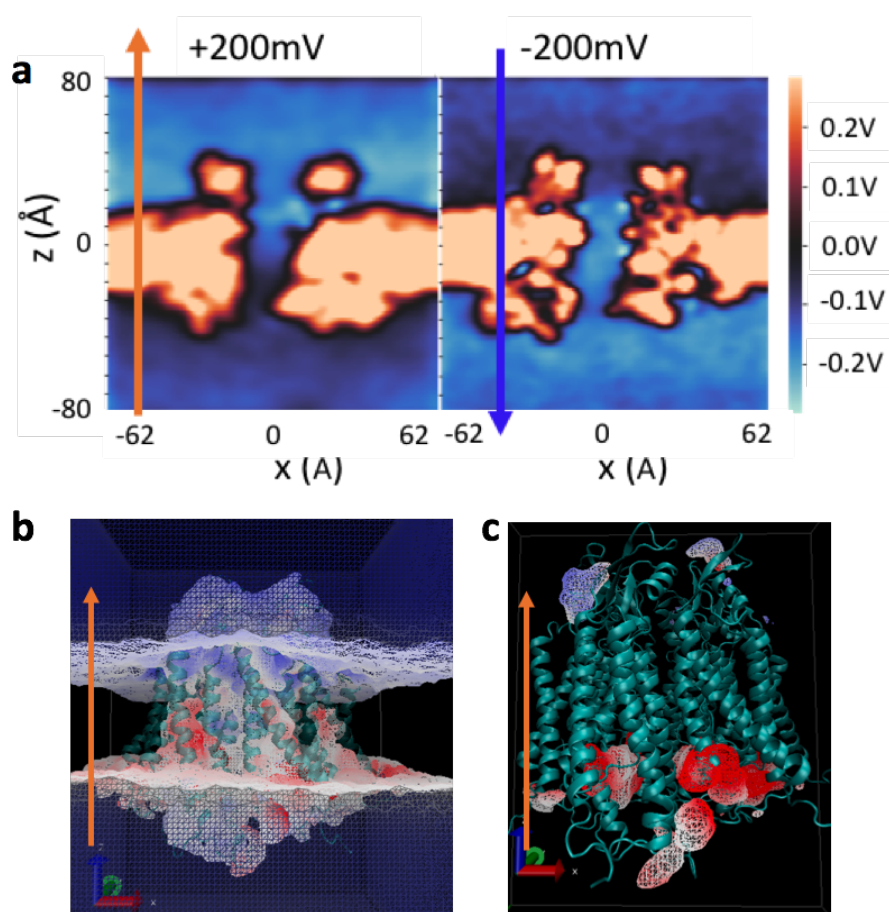
395

396   **Figure 6. Electrostatic potential at +200 mV and -200 mV membrane potential. a.** 2D-electrostatic

397   potential maps (vector 1,0,1 in Cartesian space) at +/-200 mV voltage. **b**. The 3D-electrostatic potential at

398   +200 mV overlaid onto the solvent mass density iso-surface. The color scale is -12.11 to 32.95 kT/e (-323

399   to +880 mV) from blue to red. Only the electrostatic potential at +200 mV is shown here as the color change

17

400   in -200 mV is much less prominent (due to the opposite orientation of the intrinsic dipole as described in
401   the text). **c.** Electrostatic potential overlaid onto the cAMP density iso-surface at +200 mV. The color scale
402   is -267 to +267 mV. The iso-surface contour cutoff is 0.1 amu/Å^3 for panels b and c. All data in this figure
403   are calculated from the 3D electrostatic potential map $\phi(r)$ based on all charged atoms in the simulated
404   system by solving Poisson's equation on a 1 Å resolution grid using the VMD PMEPot plugin. PMEPot
405   approximate point charge by a spherical Gaussian with an Ewald factor of 0.25. $\phi(r)$ is reported as the
406   average of 1000 snapshots from the last 200 ns.

407

408

cAMP dipole orientations during transit with and without voltage

410   cAMP is a fairly rigid molecule with a dipole moment of 37.1 Debye. **Figure 7a** shows the
411   probability distribution of the angle between the cAMP dipole vector and the z-axis from
412   milestoning simulations. Strikingly, the cAMP molecule rotates nearly 180 degrees five times on
413   its way through the pore (-40~-20, -20~0, 0~10, 10~20, and 20~40 Å along the z-axis, also see
414   **Video1**). To understand this, the z-components of the force vector acting on the cAMP from the
415   rest of the system (protein, water, ions, lipids) are decomposed into electrostatics and vdW terms
416   and plotted for each milestoning cell (**Figure 7b**). The reciprocal forces of electrostatics and vdW
417   terms along the channel indicate that cAMP has close interaction with the walls of the pore lumen.
418   The positive forces push cAMP in the positive z-direction, thus facilitating outward permeation,
419   while the negative force facilitates inward permeation towards the negative z-direction. The cAMP
420   dipole rotates as the electrostatic force vector switches the sign. Thus, these re-orientations of
421   cAMP dipole vector in absence of external voltage are due to the local electric field along the pore
422   lumen.
423   What is the cAMP dipole orientation during the voltage induced transition? **Figure 7c**
424   shows scatter plots of dipole vectors of permeating cAMP molecules from the two voltage
425   simulations. Between the intracellular entrance and the major binding site (-40~-15 Å), dipole
426   angles show orientation under both voltages similar to those at zero voltage, likely due to the large
427   magnitude of the channel local field. While the sampling is scarce in the barrier crossing region (-
428   20<z<20 Å), cAMP shows a clear preference in orientation at z~0 and at z~25-30 Å, adopting the
429   opposite orientation under the two voltages. Thus, this position-dependent dipole orientation of
430   the permeant is under the influence of both channel's internal field and external voltage.

431
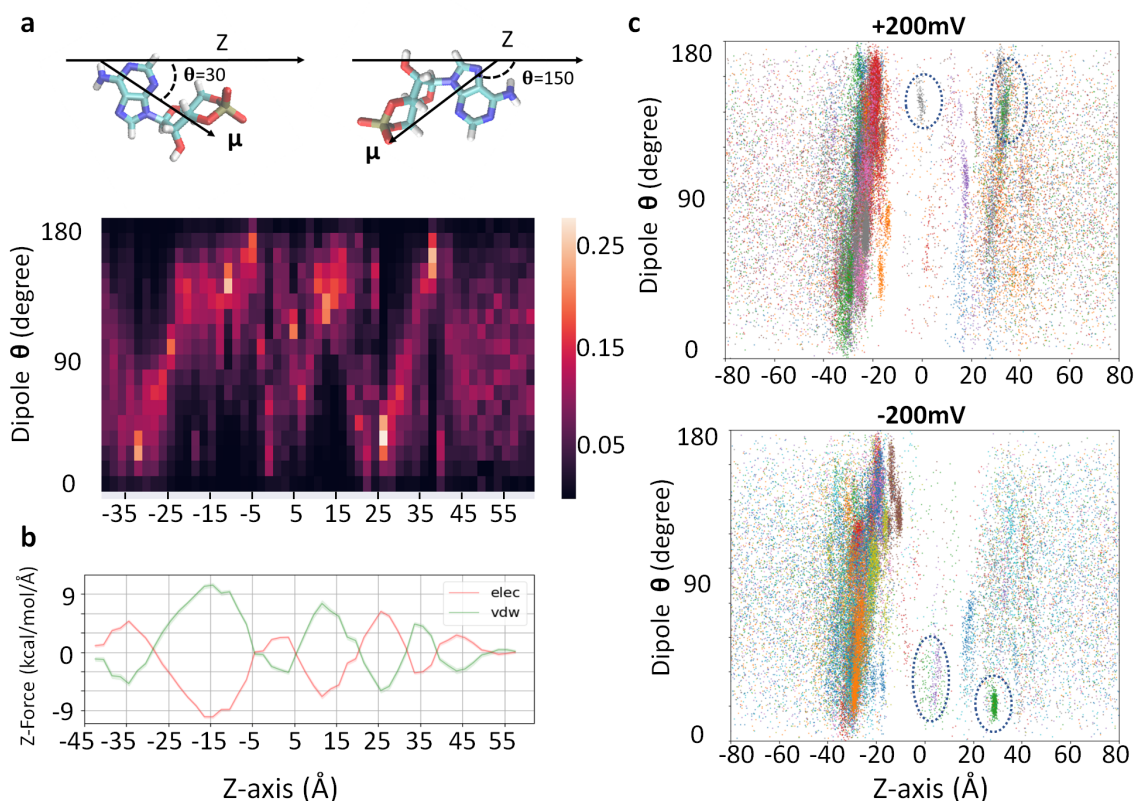
Figure 7. Dipole moment distribution and force decomposition for cAMP molecules along z-axis (Å).
a. cAMP dipole angles are shown as a probability distribution for each milestoning simulation cell along the z-axis. Color bar shows the probability scale, lighter color represents higher probability. Representative dipole angles are illustrated on the cAMP molecule above. **b.** Mean electrostatic and vdW forces along z-axis on cAMP. Lines represent the running average of three milestoning cells and shaded areas represent standard error of mean within each milestoning cell. **c.** Dipole angle scatterplot of permeating cAMP at -200 mV and +200 mV voltage simulations, respectively.

Position-dependent influence of cAMP on pore radius

The presence of cAMP at different locations inside the pore may have local or non-local effect of the pore radius. To investigate the local effect, we first compared the mean and standard deviations of the pore radius along the length of the pore, without cAMP (**Figure 8a** black), and the radius at the position of the cAMP (**Figure 8a** blue). Notice that these pore radii, obtained using grid-based cavity search program trj_cavity(*25*) ranged between 10 to 15 Å, larger than the 7.5 to 10.5 Å radius obtained using the Hole program (**Figure S1**). This is because the Hole algorithm uses a spherical probe of increasing radius, which underestimates the space within an

19

450   irregularly shaped pore. It is clear that the influence on pore radius depends where cAMP resides.

451   For instance, the radius increases when cAMP is around z=-10 Å but decreases when it is around

452   z=30 Å, near the extracellular entrance of the pore.

453   To investigate the non-local effect of cAMP on pore radius, we used Cohen's d-score of

454   pore radius distributions to compare the mean radius with and without cAMP. **Figure 8b** is a heat

455   map of Cohen's d-score, which depicts the changes in pore diameter along its length (x-axis of

456   Figure 8b) caused by the presence of cAMP at each z-position (y-axis of Figure 8b). Cohen's d

457   score was calculated using Eq. 3:

458
$$\text{Cohen's } d(z) = \frac{\langle r'(z) \rangle - \langle r(z) \rangle}{SD_{pooled}}$$

459   Eq. 3

460   where $\langle r(z) \rangle$ is the average pore radius without cAMP (**Figure 8a** black) and $\langle r'(z) \rangle$ is the

461   average pore radius with cAMP at various locations of the pore, obtained from milestoning

462   simulation trajectories. SD_{pooled} is the pooled standard deviation of $r'(z)$ and $r(z)$. For instance,

463   a d score of 6 (blue) means the presence of cAMP increased the mean pore radius by $6SD_{pooled}$,

464   and a score of -6 (green) means the presence of cAMP decreased the mean pore radius by

465     $6SD_{pooled}$.

466   On the heatmap, the d score colors along the diagonal line represent changes in pore radius

467   at the position of the cAMP, corresponding to the **Figure 8a** blue profile. The off-diagonal colors

468   represent non-local changes in pore radius away from where the cAMP is in the pore. It is evident

469   that the presence of cAMP can have effects on lumen radius distant from the position of cAMP

470   itself. Most interestingly, the scores on the upper off diagonal are more populated by green color

471   and lower off diagonal shows more blue color. This trend suggests that when cAMP permeates

472   through the pore, it tends to enlarge the radius on the left side (towards intracellular) and narrow

473   the radius on the right side (towards extracellular).

474
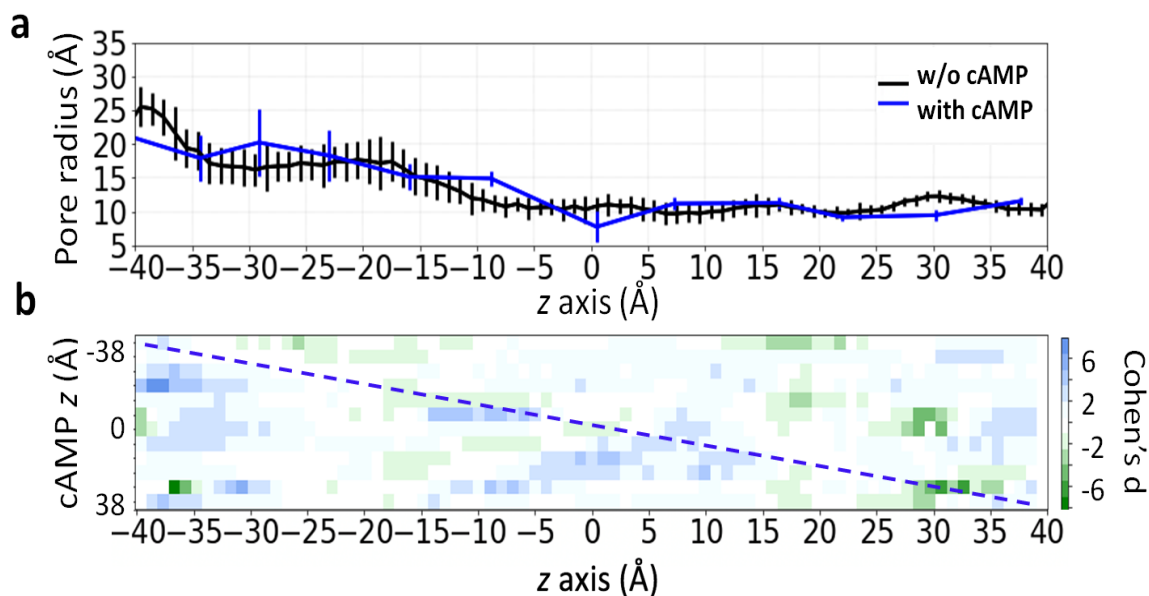
Figure 8. Local and non-local effect of cAMP on Cx26 pore radius. a. The mean and standard deviation of Cx26 pore radius as a function of the pore axis (z) during simulations. Black line is the pore radius without cAMP and the blue line is the radius at the position of the cAMP. **b.** Cohen's d-scores of pore radius distributions (see Eq. 3). The y-axis indicates the z-position of the cAMP, and x-axis is the full-length pore axis. The dashed diagonal line indicates changes in pore radius at the position of the cAMP, while off-diagonal colors indicate cAMP's non-local effect on the pore radius. Blue color indicates cAMP increase the pore radius, and green color indicates cAMP decrease the pore radius.

Discussion

In this work, we use both long-timescale MD simulations and Voronoi-tessellated Markovian milestoning, an enhanced sampling method, to explore how a charged biological signaling molecule, cAMP, permeates a connexin pore. This work builds on our previously developed Cx26 hemichannel model that was validated regarding ionic and molecular permeation properties (*6, 26*). We first obtained the density profile and the inward/outward cAMP flux under +/-200 mV voltages from two 2 μs MD simulations in the presence of multiple cAMPs. These results were compared with the intrinsic potential of mean force (PMF) and the inward/outward mean first passage time (MFPT) of a single cAMP at zero voltage obtained from a total 16.5 μs of multi-replica milestoning. Those two computational approaches – long timescale under voltage (nonequilibrium) and milestoning without voltage (equilibrium) - provided complementary information that allowed detailed analysis of the kinetics of cAMP transit in the absence and presence of voltage. The relation between voltage simulations and milestoning simulations were investigated by deriving

21

498   the PMF under voltage from the intrinsic PMF, which revealed how mobile ions and protein dipole

499   contribute to the resulting free energy landscape. In addition, unbiased simulations within each

500   milestoning cell allowed us to examine the dipole orientation of cAMP through the pore, with and

501   without voltage, and the short-range and long-range effects of cAMP on pore width.

502   Both +/-200 mV simulations in 26.5 mM cAMP and single cAMP milestoning simulation

503   featured a prominent intra-pore binding site, characterized by a high density of positive charge in

504   a wide entrance region of the pore. The PMF and MFPT from milestoning further allow us to

505   estimate the binding constants, dissociation and association rates of cAMP inside the channel.

506   Because this site directly communicates with the bulk aqueous compartment, the "saturability" of

507   this site is expected to be more complex than for a single-occupancy ion binding site in a narrow

508   pore. The results also suggest that while more than one cAMP can be within the pore, this multiple

509   occupancy only occurs at this wide binding site, and that there is single cAMP occupancy in the

510   rest of the pore. Thus, we do not anticipate for cAMP the types of multi-occupancy effects seen in

511   many ion channels that involve interactions between sequential single-occupancy binding sites in

512   the permeation path.

513   At the cAMP concentration used in the voltage simulations, we noted that there were self-

514   interactions among the cAMP molecules at the pore entrance (and not in bulk); cAMP molecules

515   interacted with each other via pi-stacking and hydrogen bonding. This was likely due to the 26.5

516   mM bulk concentration of cAMP used to obtain a measurable number of transits during the voltage

517   simulation, which was greatly increased at the pore entrance. Although Mg^{2+} ions were included

518   in all simulations, they did not play a role in cAMP clustering or permeation. The cAMP self-

519   interactions are unlikely to occur at any physiological cAMP concentrations. However, this finding

520   does point out a caution when a high concentration of ligands is used in computations to speed up

521   the sampling.

522   While the cAMP binding sites and barriers are consistent between nonequilibrium

523   simulation and milestoning simulation, the kinetic features are different. Under equilibrium

524   simulation, the cAMP transit time is ~3 times faster inward than outward. This asymmetric rate is

525   not seen under voltage simulations. By exploring how voltages influence the permeation free

526   energy, we found this is likely due to the negative voltage reducing the free energy barrier for

527   outward cAMP flux. The influence of the voltage of the free energy profile of cAMP permeation

528   was estimated using intrinsic PMF from milestoning and the electrostatic potential change induced

22

529   by the electric field. The PMF profiles under the two opposite voltages highlight how the external

530   voltage alters the thermodynamics and kinetics of cAMP permeation by changing both the relative

531   free energy as well as free energy barriers. One feature that emerged from this investigation was

532   the recognition of that for a wide pore such as connexin26, the effect of voltage on the mobile

533   charges and polarizable elements within the pore produces changes in the electrostatic field within

534   the pore that affect permeation. These voltage-induced modifications of the reaction field alter the

535   energetic landscape in protein-specific ways. For the connexin channel, these changes are imposed

536   on the intrinsic dipole within the pore, which is responsible for the asymmetric effect of symmetric

537   voltage changes on the overall PMF.

538   The primary impact of this work will be to establish a way to generate meaningful

539   hypotheses/understanding about the basis of molecular selectivity of connexin channels and to

540   understand how mutations of connexin proteins alter the selectivity and thereby cause human

541   pathologies. Such hypotheses can be tested experimentally and computationally in a synergistic

542   manner. The broader application will be studies of permeation of other biological molecules (e.g.,

543   ATP, glutathione, IP3) known to permeate Cx26 channels, and eventually to extend the work to

544   other connexin isoforms as validated atomic models are developed. The methods presented in this

545   study can be applied to understand the molecular permeation through a large pore in general. Input

546   files and raw data used to generate each figure, as well as python3 scripts for milestoning analysis

547   are publicly available at https://github.com/LynaLuo-Lab/Connexin-cAMP-milestoning. Long

548   timescale MD trajectories are publicly available on Anton2 supercomputer.

549

550

551 **Methods**

552   **Force Field and cAMP Parameterization**

553   CHARMM36 force field was used for protein(*27, 28*) , POPC lipids(*29*), KCl, and TIP3P water

554   (*30*). Mg^{2+} parameters were from Yoo and Aksimentiev(*31*), in which the van der Waals

555   interaction parameters were fine-tuned to reproduce experimental osmotic pressure. For cAMP,

556   force field parameters were first generated from CHARMM CGenFF(*32*). Additional dihedral

557   fitting between imidazole and pyran groups in the cAMP structure was performed in VMD ffTK

558 plugin(*33*). The final optimized cAMP parameters are provided in **Table S3** and

559 https://github.com/LynaLuo-Lab/Connexin-cAMP-milestoning.

560

561 **System Setup and Equilibrium Protocol**

562 The atomistic model of Cx26 was taken from our previous work(*6*). The system of Cx26 embedded

563 in a solvated 1-palmitoyl-2-oleoylphosphatidylcoline (POPC) bilayer with ions and TIP3P water

564 molecules was built and equilibrated following the step-by-step protocol used in Membrane

565 Builder in CHARMM-GUI website(*34, 35*). The terminal amino acid of each segment was capped

566 using acetylated N-terminus (ACE) and methylated C-terminus (CT1), and three disulfide bonds

567 were added between the residue pairs of C53 and C180, C64 and C169, C60 and C174 respectively

568 per protomer, so the entire channel contained a total of 18 disulfide bonds, consistent with the

569 original crystal structure (PDB ID 2ZW3). Two systems containing a single cAMP for milestoning

570 simulations and 27 cAMP for nonequilibrium simulations were constructed (see **Table S1** for

571 system details for Milestoning simulation and Anton2 simulation). The system was energy

572 minimized and serially equilibrated in NVT and NPT ensembles with positional restraints using

573 AMBER18(*36*). Temperature was maintained at 310.15 K using Langevin thermostat (*37, 38*) and

574 1 atm was maintained by Monte Carlo barostat pressure control (*39, 40*). The time step was 2 fs.

575 Cutoff for calculating van der Waals interactions and short-range electrostatic interactions was set

576 at 12 Å and force-switched at 10 Å. Long-range electrostatic interactions were calculated using

577 the particle mesh Ewald algorithm (*41*).

578

579 **Anton2 Simulation Protocol**

580 After 35 ns equilibrium simulation, the system was run on Anton2 supercomputer with 2.0

581 fs timestep. Lennard-Jones interactions were truncated at 11-13 Å and long-range electrostatics

582 were evaluated using the k-Gaussian Split Ewald method (*42*). Pressure regulation was

583 accomplished via the Martyna-Tobias-Klein (MTK) barostat, to maintain 1 bar of pressure, with a

584 tau parameter of 0.0416667 ps and reference temperature of 310.15 K. The barostat period was set

585 to the default value of 480 ps per timestep. Temperature control was accomplished via the Nosé-

586 Hoover thermostat with the same tau parameter. The mts parameter was set to 4 timesteps for the

587 barostat control and 1 timestep for the temperature control. The thermostat interval was set to the

588 default value of 24 ps per timestep. A 600 ns equilibrium simulation was finished before applying

589   voltage. Constant electric fields of -200 or +200 mV respectively were added for 2μs simulation

590   time with trajectories saved every 200 ps.

591

592   **Milestoning MD Simulation Setup**

593   The AMBER18 CUDA version currently does not support Cartesian coordinates as a

594   collective variable. Thus, we pinned two water molecules using high Cartesian restrain with 6000

595   kcal mol^{-1} Å^{-2} force constant in the top and bottom bulk region (20 Å away from the intracellular

596   entrance and 35 Å away from the extracellular entrance of the channel). We then used the projected

597   distance on the z vector between the nearest pinned water oxygen and center of mass of cAMP to

598   define the Voronoi cells along the z-axis. A cylindrical restraint with a radius of 30Å was applied

599   to cAMP to confine the sampling in the bulk region. To prevent protein drifting, a strong harmonic

600   distance restraint with force constant 2000 kcal mol^{-1} Å^{-2} along *xyz*-axes between the fixed water

601   oxygen in intracellular bulk and the center of mass of the protein was added. The simulation

602   protocol in AMBER18 is the same as above, except all milestoning simulations were run in NVT

603   ensemble. The timestep was 2.0 fs, and each trajectory was saved every 500 frames. Confinement

604   within the Voronoi cells was obtained using flat-harmonic restraint with force constant of 100 kcal

605   mol^{-1} Å^{-2}.

606

607   **Convergence of PMF and MFPT**

608   A total of 53 Voronoi cells B_i are evenly distributed 2 Å apart along the z-axis, and a

609   milestone state S_{ij} is defined as the boundary between two adjacent Voronoi cells B_i and B_j. If $k_{i \to j}$

610   is the rate of attempted escape from cells B_i to B_j, since at statistical equilibrium the total flux in

611   and out of each cell is zero, thus the equilibrium probability π_i for the system to be in cell B_i

612   satisfies a balance equation:

613
$$\sum_{j=1, j \neq i} \pi_j k_{j \to i} = \sum_{j=1, j \neq i} \pi_i k_{i \to j}, \sum_{i=1} \pi_i = 1$$

614   Eq. 4

615   The free energy of each cell can be obtained from the solution of Eq. 4 as $-k_B T ln(\pi_i)$.

616   Running independent simulations in the various cells allow focused sampling on the cells with

617   slow convergence. Here, we monitor the convergence of π_i on the fly by plotting the accumulated

25

618   rate of attempted escape on both sides of the cell B_i, called $k_{i \to j}$ and $k_{i \to k}$, as well as the retention

619   rate inside each cell B_i over time (**Figure 9a**).

620   By defining a milestone S_{ij} as the boundary between two adjacent Voronoi cells B_i and B_j,

621   the dynamics of the system is reduced to that of a Markov chain in the state space of the milestones

622   indices(*9*). The MFPT between any pair of milestones S_{ij} and S_{ik} can hence be calculated from the

623   rate matrix whose elements $q_{ij,ik}$, the rate of moving from milestone S_{ij} to S_{ik}, are given by:

624
$$q_{ij,ik} = \frac{\pi_i n^i_{ij,ik}}{\pi_i r^i_{ij} + \pi_j r^j_{ij}}$$

625   Eq. 5

626   where $n^i_{ij,ik}$ is the number of transitions from S_{ij} to S_{ik}, normalized by the time spend in cell B_i,

627   and r^i_{ij} is the time passed in cell B_i after having hit S_{ij} before hitting any other milestone,

628   normalized by the total time spent in cell B_i. Therefore, after the convergence of π_i, the

629   convergence of MFPT can be monitored directly from the accumulated $n^i_{ij,ik}$ and r^i_{ij} (**Figure 9b**).

630   The final 10 ns windowed relative RMSD is also given at the top of each panel as a measure of the

631   degree of convergence for the corresponding rate matrix entry components. This relative RMSD

632   within 5% (averaged over all replicas) is used as convergence criteria for all Voronoi cells, and in

633   most cases, RMSD is below 2% of the mean value.

634

635

636
637
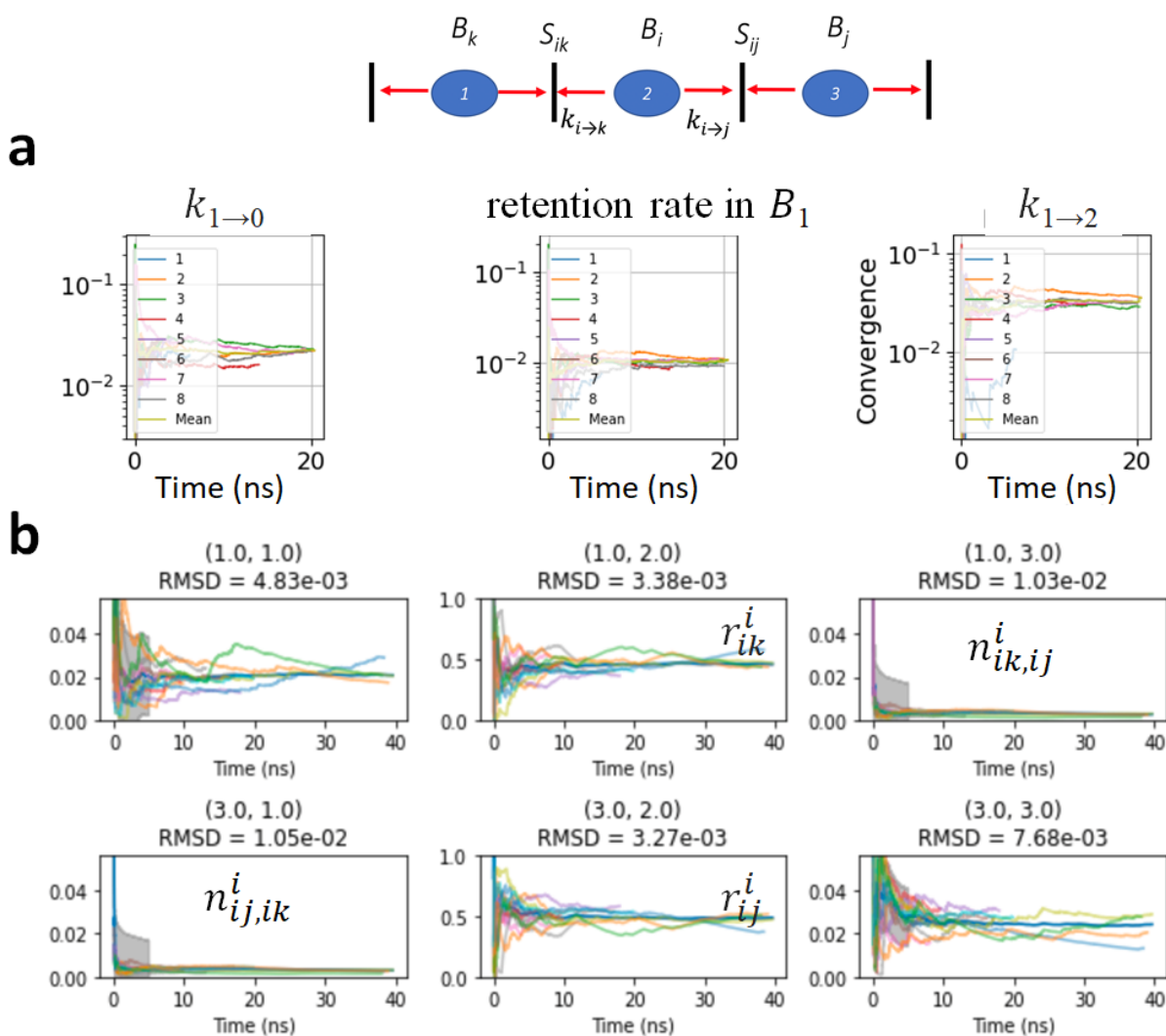
639   **Figure 9. a**) PMF convergence plots of a Voronoi cell for index 01 (B_1) over time. The left and right plots
640   represent the probability of the attempted escape to the left or right milestone states, $k_{1\to0}$ and $k_{1\to2}$. The
641   middle plot is the retention rate inside the cell. Colors represent different replicas and the mean of all the
642   replicas. **b**) convergence plots for MFPT for milestoning cell index 2 (B_2). The upper right and lower left
643   panels correspond to the frequencies of transitions from cell 1 to cell 3, and from cell 3 to cell 1, respectively
644   ($n_{ij,ik}^i$ and $n_{ik,ij}^i$). The upper and lower center panels correspond to the percentage of time spent in cell 2
645   after last touching cell 1 and 3 (r_{ij}^i), respectively. The other two entries are not used in analysis directly but
646   would correspond to re-entering cell 1 or 3 after last visiting that same cell previously. The final 10 ns (last
647   10,000 frames) windowed relative RMSD is also given at the top of each panel as a measure of the degree
648   of convergence for the corresponding rate matrix entry components. The convergence plots of all 53 cells
649   are available on Github repository https://github.com/LynaLuo-Lab/Connexin-cAMP-milestoning.

650

651

652

27

659 **AUTHOR CONTRIBUTIONS**

660 W.J. prepared the system, re-parameterized cAMP force field, and performed nonequilibrium simulations.

661 Y-C.L. performed milestoning simulations. W.J and Y-C.L analyzed both simulation results. W.M.B-S.

662 prepared milestoning analysis scripts in python3 and performed convergence analysis. W.M.B-S. and L.M.

663 supervised milestoning simulation. Y.L.L, A.H, L.M, and J.C designed the project and wrote the paper with

664 input from all authors.

665

666 **REFERENCES**

667 1.   A. L. Harris, Connexin channel permeability to cytoplasmic molecules. *Prog Biophys Mol*
668   *Biol* **94**, 120-143 (2007).
669 2.   A. L. Harris, Emerging issues of connexin channels: biophysics fills the gap. *Q Rev Biophys*
670   **34**, 325-472 (2001).
671 3.   A. Pfenniger, A. Wohlwend, B. R. Kwak, Mutations in connexin genes and disease. *Eur J*
672   *Clin Invest* **41**, 103-116 (2011).
673 4.   I. E. García, P. Prado, A. Pupo, O. Jara, D. Rojas-Gómez, P. Mujica, C. Flores-Muñoz, J.
674   González-Casanova, C. Soto-Riveros, B. I. Pinto, M. A. Retamal, C. González, A. D. Martínez,
675   Connexinopathies: a structural and functional glimpse. *BMC Cell Biol* **17**, S17 (2016).
676 5.   A. L. Harris, Connexin specificity of second messenger permeation: real numbers at last. *J*
677   *Gen Physiol* **131**, 287-292 (2008).
678 6.   Y. Luo, A. R. Rossi, A. L. Harris, Computational studies of molecular permeation through
679   connexin26 channels. *Biophys J* **110**, 584-599 (2016).
680 7.   A. K. Faradjian, R. Elber, Computing time scales from reaction coordinates by milestoning.
681   *J Chem Phys* **120**, 10880-10889 (2004).
682 8.   R. Elber, Milestoning: An Efficient Approach for Atomically Detailed Simulations of
683   Kinetics in Biophysics. *Annu Rev Biophys* **49**, 69 (2020).
684 9.   E. Vanden-Eijnden, M. Venturoli, Markovian milestoning with Voronoi tessellations. *J*
685   *Chem Phys* **130**, 194101 (2009).

686   10.   T.-Q. Yu, M. Lapelosa, E. Vanden-Eijnden, C. F. Abrams, Full kinetics of CO entry, internal
687   diffusion, and exit in myoglobin from transition-path theory simulations. *J Am Chem Soc*
688   **137**, 3041-3050 (2015).
689   11.   B. R. Jagger, A. A. Ojha, R. E. Amaro, Predicting Ligand Binding Kinetics Using a Markovian
690   Milestoning with Voronoi Tessellations Multiscale Approach. *J Chem Theory Comput* **16**,
691   5348-5357 (2020).
692   12.   L. Maragliano, E. Vanden-Eijnden, B. Roux, Free energy and kinetics of conformational
693   transitions from Voronoi tessellated milestoning with restraining potentials. *J Chem*
694   *Theory Comput* **5**, 2589-2594 (2009).
695   13.   X. He, Y. Shen, F. R. Hung, E. E. Santiso, Heterogeneous nucleation from a supercooled
696   ionic liquid on a carbon surface. *J Chem Phys* **145**, 211919 (2016).
697   14.   G. Alberini, F. Benfenati, L. Maragliano, Molecular dynamics simulations of ion selectivity
698   in a claudin-15 paracellular channel. *J Phys Chem B* **122**, 10783-10792 (2018).
699   15.   G. Kanaporis, G. Mese, L. Valiuniene, T. W. White, P. R. Brink, V. Valiunas, Gap junction
700   channels exhibit connexin-specific permeability to cyclic nucleotides. *J Gen Physiol* **131**,
701   293-305 (2008).
702   16.   V. K. Verselis, C. S. Ginter, T. A. Bargiello, Opposite voltage gating polarities of two closely
703   related onnexins. *Nature* **368**, 348-351 (1994).
704   17.   B. I. Pinto, I. E. García, A. Pupo, M. A. Retamal, A. D. Martínez, R. Latorre, C. González,
705   Charged residues at the first transmembrane region contribute to the voltage
706   dependence of the slow gate of connexins. *J Biol Chem* **291**, 15740-15752 (2016).
707   18.   H. C. Berg, *Random Walks in Biology: New and Ezpanded Edition*. (Princeton, 1984).
708   19.   V. H. Hernandez, M. Bortolozzi, V. Pertegato, M. Beltramello, M. Giarin, M. Zaccolo, S.
709   Pantano, F. Mammano, Unitary permeability of gap junction channels to second
710   messengers measured by FRET microscopy. *Nat Methods* **4**, 353-358 (2007).
711   20.   S. Berneche, B. Roux, A microscopic view of ion conduction through the K+ channel. *Proc*
712   *Natl Acad Sci* **100**, 8644-8648 (2003).
713   21.   B. Roux, Statistical mechanical equilibrium theory of selective ion channels. *Biophys J* **77**,
714   139-153 (1999).
715   22.   B. Roux, The membrane potential and its representation by a constant electric field in
716   computer simulations. *Biophys J* **95**, 4205-4216 (2008).
717   23.   J. Gumbart, F. Khalili-Araghi, M. Sotomayor, B. Roux, Constant electric field simulations of
718   the membrane potential illustrated with simple systems. *Biochim Biophys Acta* **1818**, 294-
719   302 (2012).
720   24.   A. Aksimentiev, K. Schulten, Imaging α-hemolysin with molecular dynamics: ionic
721   conductance, osmotic permeability, and the electrostatic potential map. *Biophys J* **88**,
722   3745-3761 (2005).
723   25.   T. Paramo, A. East, D. Garzón, M. B. Ulmschneider, P. J. Bond, Efficient characterization of
724   protein cavities within molecular simulation trajectories: trj_cavity. *J Chem Theory*
725   *Comput* **10**, 2151-2164 (2014).
726   26.   T. Kwon, A. L. Harris, A. Rossi, T. A. Bargiello, Molecular dynamics simulations of the Cx26
727   hemichannel: evaluation of structural models with Brownian dynamics. *J Gen Physiol* **138**,
728   475-493 (2011).

27.   A. D. MacKerell Jr, D. Bashford, M. Bellott, R. L. Dunbrack Jr, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin, M. Karplus, All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J Phys Chem B* **102**, 3586-3616 (1998).

28.   A. D. Mackerell Jr, M. Feig, C. L. Brooks III, Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *J Comput Chem* **25**, 1400-1415 (2004).

29.   J. B. Klauda, R. M. Venable, J. A. Freites, J. W. O'Connor, D. J. Tobias, C. Mondragon-Ramirez, I. Vorobyov, A. D. MacKerell Jr, R. W. Pastor, Update of the CHARMM all-atom additive force field for lipids: validation on six lipid types. *J Phys Chem B* **114**, 7830-7843 (2010).

30.   W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, Comparison of simple potential functions for simulating liquid water. *J Chem Phys* **79**, 926-935 (1983).

31.   J. Yoo, A. Aksimentiev, Improved parametrization of Li+, Na+, K+, and Mg2+ ions for all-atom molecular dynamics simulations of nucleic acid systems. *J Phys Chem Lett* **3**, 45-50 (2012).

32.   K. Vanommeslaeghe, E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, P. Lopes, I. Vorobyov, A. D. MacKerell Jr, CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J Comput Chem* **31**, 671-690 (2010).

33.   C. G. Mayne, J. Saam, K. Schulten, E. Tajkhorshid, J. C. Gumbart, Rapid parameterization of small molecules using the force field toolkit. *J Comput Chem* **34**, 2757-2770 (2013).

34.   S. Jo, T. Kim, W. Im, Automated builder and database of protein/membrane complexes for molecular dynamics simulations. *PloS one* **2**, e880 (2007).

35.   S. Jo, T. Kim, V. G. Iyer, W. Im, CHARMM-GUI: a web-based graphical user interface for CHARMM. *J Comput Chem* **29**, 1859-1865 (2008).

36.   D. Case, D. Cerutti, T. Cheatham, T. Darden, R. Duke, T. Giese, H. Gohlke, A. Goetz, D. Greene, N. J. U. o. C. Homeyer, AMBER reference manual. (2018).

37.   W. G. Hoover, A. J. Ladd, B. Moran, High-strain-rate plastic flow studied via nonequilibrium molecular dynamics. *Phys Rev Lett* **48**, 1818 (1982).

38.   D. J. Evans, Computer ''experiment''for nonlinear thermodynamics of Couette flow. *J Chem Phys* **78**, 3297-3302 (1983).

39.   J. Åqvist, P. Wennerström, M. Nervall, S. Bjelic, B. O. Brandsdal, Molecular dynamics simulations of water and biomolecules with a Monte Carlo constant pressure algorithm. *Chem. Phys. Lett* **384**, 288-294 (2004).

40.   M. P. Allen, D. J. Tildesley, *Computer simulation of liquids*. (Oxford university press, 2017).

41.   T. Darden, D. York, L. Pedersen, Particle mesh Ewald: An N· log (N) method for Ewald sums in large systems. *J Chem Phys* **98**, 10089-10092 (1993).

42.   Y. Shan, J. L. Klepeis, M. P. Eastwood, R. O. Dror, D. E. Shaw, Gaussian split Ewald: A fast Ewald mesh method for molecular simulation. *J Chem Phys* **122**, 054101 (2005).

772

Supporting Information

773

774
775
776
777
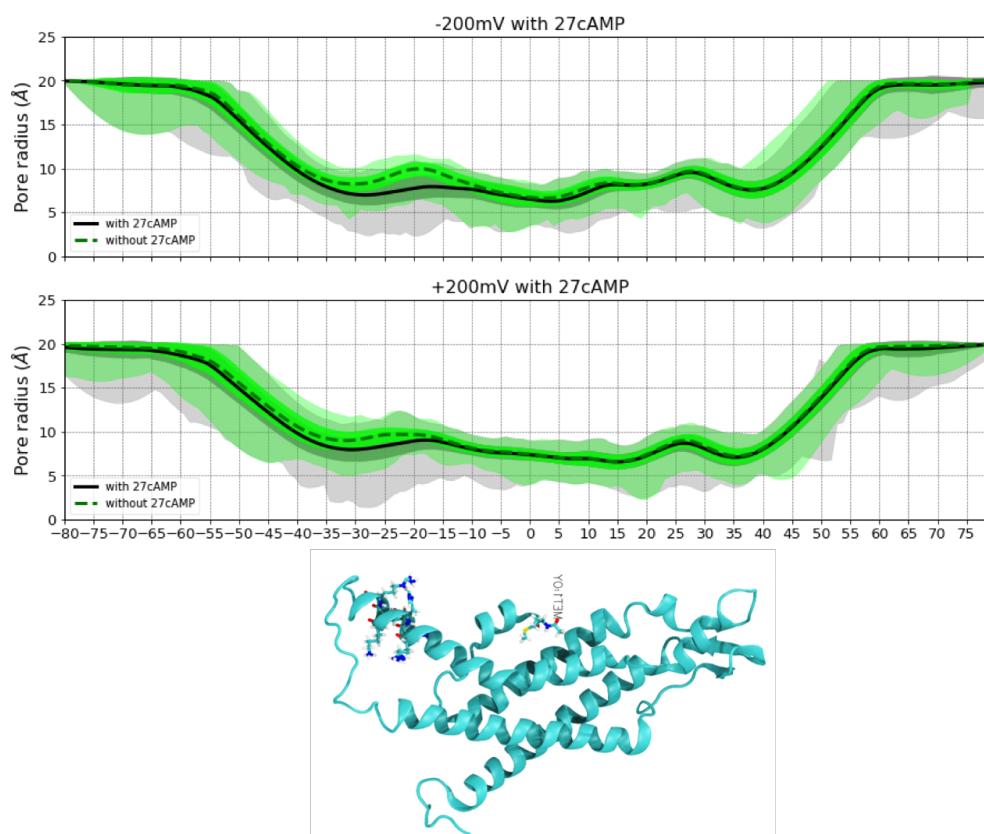
778
779 **Figure S1.** Pore radius profiles from two simulations under +/-200 mV, calculated using Hole
780 program. The black line is the average value over last 1 µs, with 1.2 ns interval between snapshots.
781 The dark grey shade represents the +/- standard deviation, the light grey shade represents the
782 minimum and maximum radius values. The dashed line and green shade are the pore radius profile
783 with protein only.
784
785
786
787
788
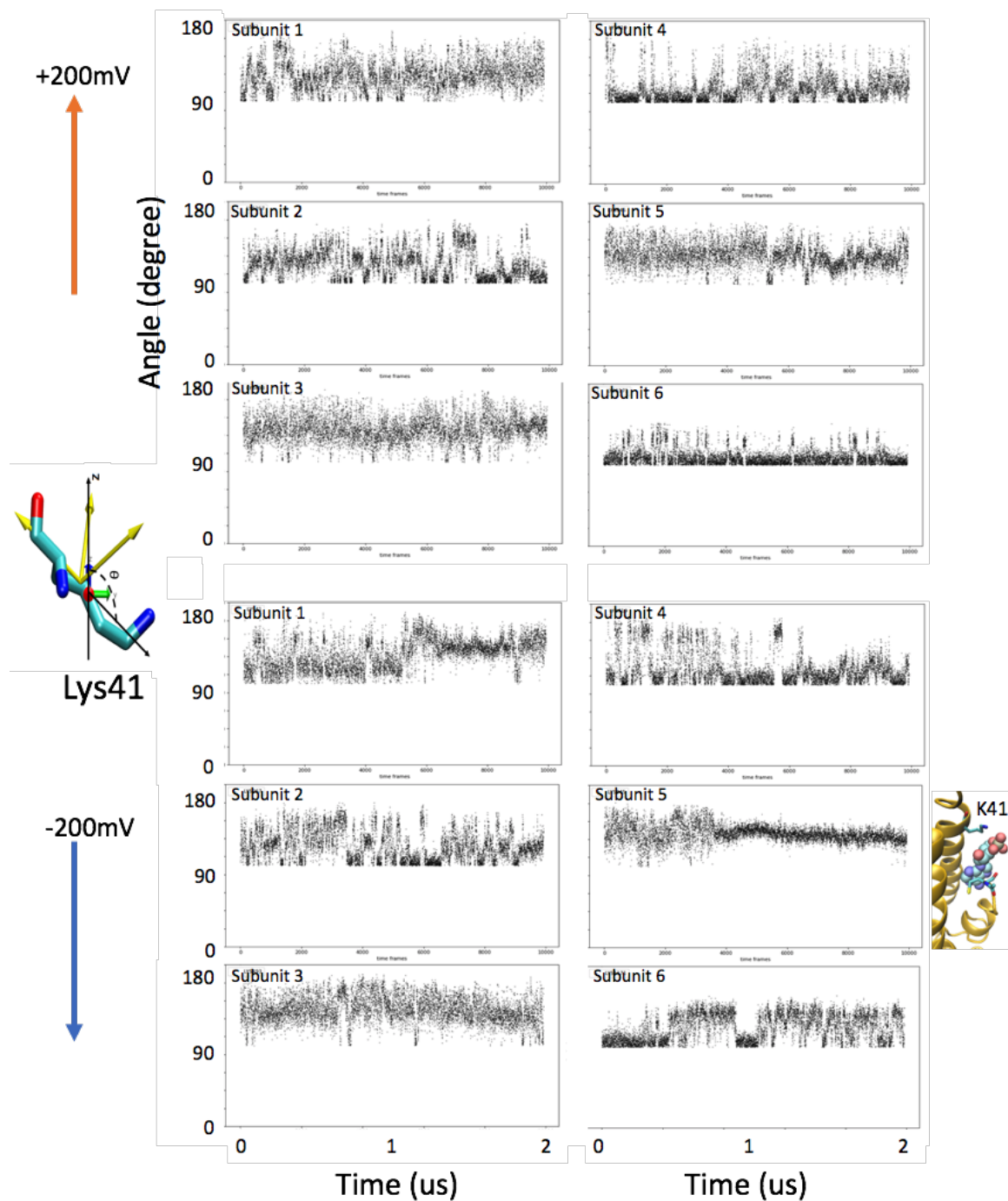
Figure S2a. Fluctuation of the angles between the principal axis of Lys41 and z-vector in each subunit under +/-200 mV. Representative Lys residue shown in the licorice model with its principal axes and the angle plotted.
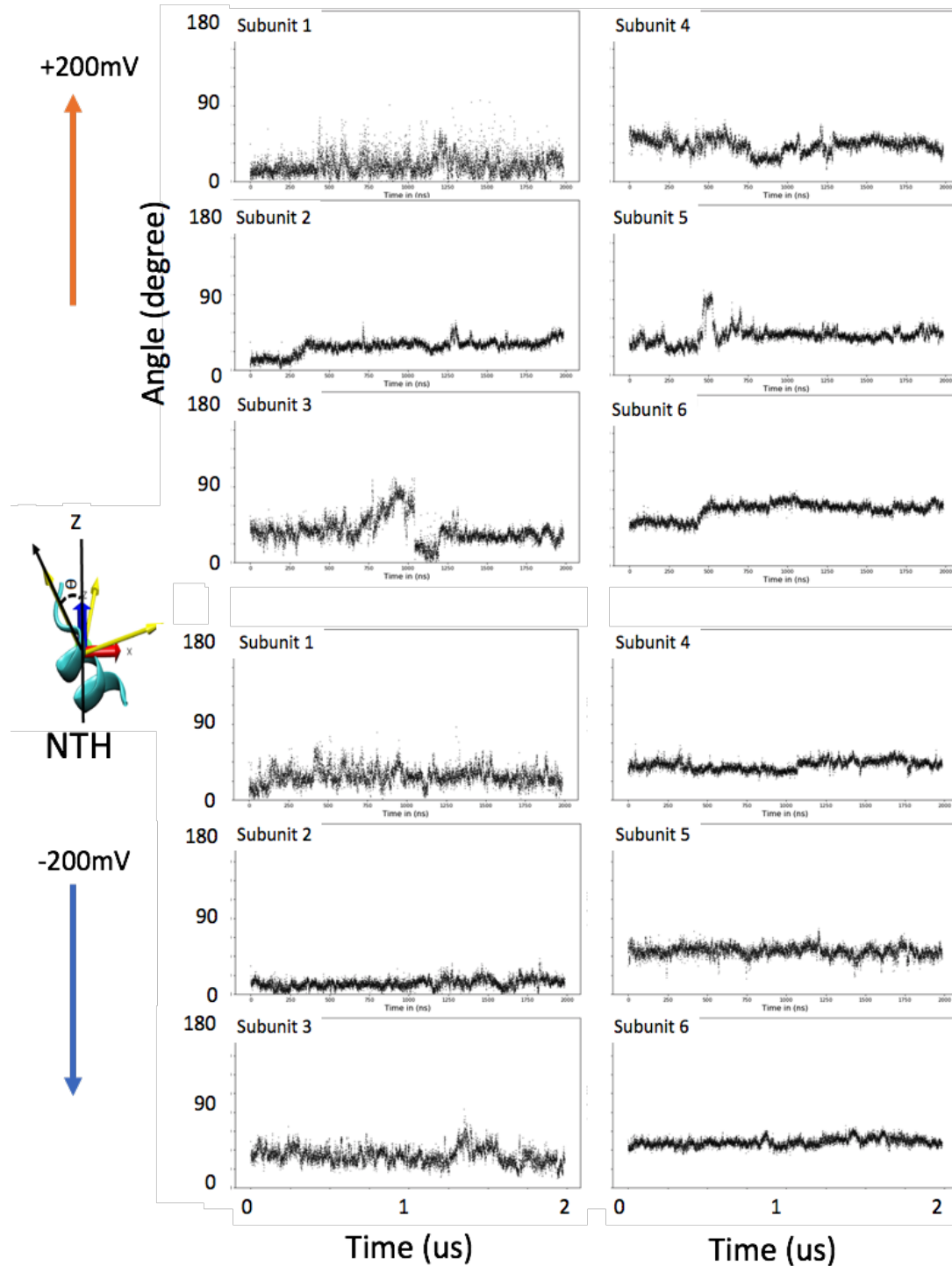
Figure S2b. Fluctuation of the angles between the principal axis of NTH (1-11) and z-vector in each subunit under +/-200 mV. Representative NTH shown in the new cartoon model with its principal axes and the angle plotted.

33

807 **Table S1.** Description of the computational systems

	Voltage	term patching	Disulfide bonds	Box (A^3)	Number (concentration*)			
					atoms	cAMP	K+/Mg2+/Cl-	POPC/water
Anton2	+/- 200mV	ACE CT1	C53-C180 C60-C174 C64-C169	122× 122x 154	237030	27 (26.5 mM)	82/27/163 (80/26/160 mM)	323/56569
Milestoning	0 mV	ACE CT1	C53-C180 C60-174 C64-C169	121x 121x 155	234652	1 (1 mM)	82/1/137 (81/1/135 mM)	324/56191

808 *Concentrations were calculated based on number of water molecules.
809
810
811 **Table S2. Transition time and dwell time (ns) from 2 µs trajectories at +200 mV and -200**

812 **mV.**

Transition Event	+200mV (inward flux)			-200mV (outward flux)		
	Transition Time -50<z<50	Dwell time -50<z<-20	Barrier crossing -20<z<50	Transition Time -50<z<50	Dwell time -50<z<-20	Barrier crossing -20<z<50
1	222	24	198	160	110	50
2	1218	1058	160	242	210	32
3	416	308	108	290	50	240
4	282	248	34	1496	1400	96
5	542	498	44	230	162	68
6	290	149	140	630	570	60
7	192	120	72	384	350	34
8	454	222	232	152	100	52
9	426	150	276	1410	1228	182
10				198	76	122
11				390	150	240
12				445	350	95
Raw mean	449	308	140	502	396	105
Sample mean*	448	305	143	510	398	106
95% CI*	204-1646	149-940	79-330	249-1571	205-1086	54-290

813 *The sample mean and confidence intervals are based on Maximum Likelihood Estimate by fitting the
814 exponential distributions (see code at https://github.com/LynaLuo-Lab/MD-data-uncertainty-
815 analysis/blob/master/confidence_interval_exponential.ipynb).
816
817
818
819 **Table S3.** Optimized cgenff force field parameters for cAMP

BONDS						
CG3RC1	OG3R60	334.3	1.411			
OG3R60	PG1	237	1.61			

		ANGLES				
CG3RC1	CG321	OG3R60	75.7	110.1		
CG321	CG3RC1	OG3C51	45	111.5		
CG3C51	CG3RC1	OG3R60	58	106.5	8	2.561
CG3RC1	CG3RC1	OG3R60	53.35	111	8	2.561
OG3R60	CG3RC1	HGA1	45.9	108.5		
CG3C51	OG3C51	CG3RC1	170	109		
CG321	OG3R60	PG1	20	120	35	2.33
CG3RC1	OG3R60	PG1	20	120	35	2.33
OG2P1	PG1	OG3R60	98.9	107.5		
OG3R60	PG1	OG3R60	80	104.3		
		DIHEDRALS				
OG3R60	CG321	CG3RC1	CG3RC1	0.6	1	0
OG3R60	CG321	CG3RC1	CG3RC1	0.45	2	0
OG3R60	CG321	CG3RC1	CG3RC1	0.7	3	0
OG3R60	CG321	CG3RC1	OG3C51	3.4	1	180
OG3R60	CG321	CG3RC1	HGA1	0.195	3	0
HGA2	CG321	CG3RC1	OG3C51	0.16	3	0
CG3RC1	CG321	OG3R60	PG1	0.002	1	180
CG3RC1	CG321	OG3R60	PG1	1.14	2	180
CG3RC1	CG321	OG3R60	PG1	0.086	3	0
HGA2	CG321	OG3R60	PG1	0.599	3	0
CG3RC1	CG3C51	CG3C51	NG2R51	0	3	0
CG3RC1	CG3C51	CG3C51	OG3C51	0	3	0
CG3C51	CG3C51	CG3RC1	OG3R60	2	3	180
CG3C51	CG3C51	CG3RC1	OG3R60	0.4	5	0
CG3C51	CG3C51	CG3RC1	OG3R60	0.8	6	0
OG311	CG3C51	CG3RC1	OG3R60	0	3	0
HGA1	CG3C51	CG3RC1	OG3R60	0.195	3	0
CG3C51	CG3C51	OG3C51	CG3RC1	0.5	3	0
NG2R51	CG3C51	OG3C51	CG3RC1	0	3	0
HGA1	CG3C51	OG3C51	CG3RC1	0.3	3	0
CG321	CG3RC1	CG3RC1	OG3R60	0.15	3	0

OG3C51	CG3RC1	CG3RC1	OG3R60	1.2	3	0
OG3R60	CG3RC1	CG3RC1	HGA1	0.15	3	0
CG321	CG3RC1	OG3C51	CG3C51	0.3	3	0
CG3RC1	CG3RC1	OG3C51	CG3C51	0	3	0
HGA1	CG3RC1	OG3C51	CG3C51	0.3	3	0
CG3C51	CG3RC1	OG3R60	PG1	2.553	1	180
CG3RC1	CG3RC1	OG3R60	PG1	1.5	1	180
CG3RC1	CG3RC1	OG3R60	PG1	0.625	2	180
CG3RC1	CG3RC1	OG3R60	PG1	0.488	1	180
CG3RC1	CG3RC1	OG3R60	PG1	0.972	3	180
HGA1	CG3RC1	OG3R60	PG1	1.903	1	0
HGA1	CG3RC1	OG3R60	PG1	0.089	3	180
CG321	OG3R60	PG1	OG2P1	0.1	3	0
CG321	OG3R60	PG1	OG3R60	2.248	1	180
CG321	OG3R60	PG1	OG3R60	0.261	2	0
CG321	OG3R60	PG1	OG3R60	2.64	3	180
CG321	OG3R60	PG1	OG3R60	0.929	1	180
CG321	OG3R60	PG1	OG3R60	1.149	3	0
CG321	OG3R60	PG1	OG3R60	2.138	2	0
CG321	OG3R60	PG1	OG3R60	1.101	1	0
CG321	OG3R60	PG1	OG3R60	1.793	2	0
CG321	OG3R60	PG1	OG3R60	0.54	3	0
CG3RC1	OG3R60	PG1	OG2P1	0.229	3	180
CG3RC1	OG3R60	PG1	OG3R60	0.799	1	0
CG3RC1	OG3R60	PG1	OG3R60	0.137	2	0
CG3RC1	OG3R60	PG1	OG3R60	1.1	3	0

820