

Supplementary Information for: Nanopore Sensing of Protein Folding

Wei Si^{†,‡} and Aleksei Aksimentiev^{*,†}

*Center for the Physics of Living Cells, Department of Physics, University of Illinois at
Urbana-Champaign, 1110 West Green Street, Urbana, Illinois 61801, United States, and
Jiangsu Key Laboratory for Design and Manufacture of Micro-Nano Biomedical
Instruments and School of Mechanical Engineering, Southeast University, Nanjing, 210096,
China*

E-mail: aksiment@illinois.edu

*To whom correspondence should be addressed

[†]Center for the Physics of Living Cells, Department of Physics, University of Illinois at Urbana-Champaign, 1110 West Green Street, Urbana, Illinois 61801, United States

[‡]Jiangsu Key Laboratory for Design and Manufacture of Micro-Nano Biomedical Instruments and School of Mechanical Engineering, Southeast University, Nanjing, 210096, China

Table of Contents:

Additional systems used for MD simulations of ionic current blockade (SI Figure S1) 3

Physical properties of proteins considered in this work. (SI Tables S1 and S2) 4

Calculation of ion concentration and mobility distributions (SI Methods 1, SI Figure S2) 5–6

Fraction of the nanopore volume occupied by the protein (SI Figure S3) 7

Theoretical model of ionic current blockades. (SI Methods 2, SI Figures S4 and S5) .. 8–10

Calculations of rotation-averaged ionic currents. (SI Methods 3, SI Figures S6 and S7) 11–13

Theory *versus* simulation at variable electrolyte conditions (SI Figure S8) 14

The effect of temperature on the ionic current difference (SI Figure S9) 15

Real-time detection of folding-unfolding transitions (SI Methods 4, SI Figures S10-S16) 16–24

The effect of protein conformation on ionic current blockades (SI Figure S17) 25

Trajectory-averaged blockade currents (SI Figure S18) 26

Characterization of protein folding intermediates (SI Figures S19 and S20) 27–28

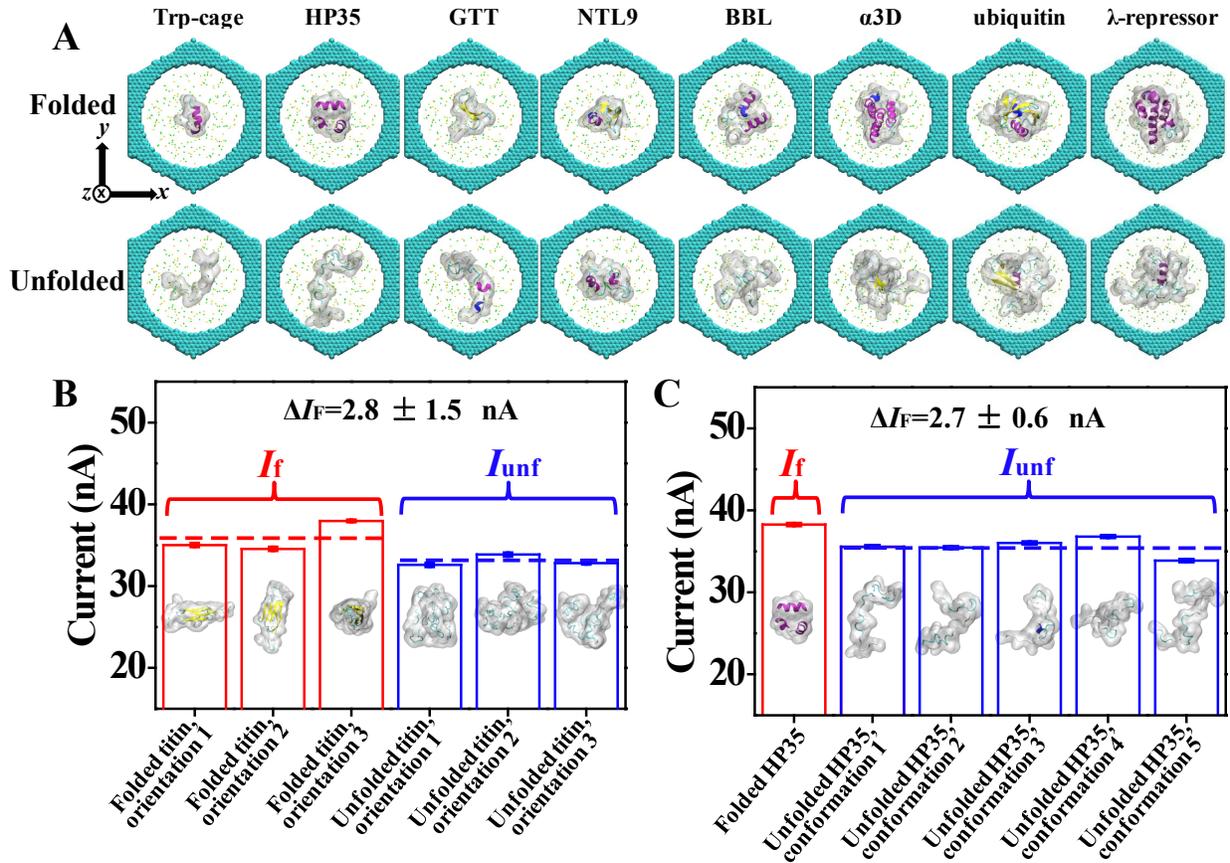


Figure S1: All-atom MD simulations of the ionic current blockade difference produced by the folding state of a protein. (A) Eight additional systems used for the all-atom MD simulations. The ninth system is shown in Figure 1A (main text). Each simulation system contains a protein in either folded (top) or unfolded (bottom) conformation at the center of a 6 nm-diameter solid-state nanochannel (cyan) filled with 2 M KCl solution. Electric field (74.67 mV/nm) was applied along the axis of the channel, the protein is restrained to maintain its initial conformation. The protein is colored according to the secondary structure of its folded conformation: purple indicates α -helix, yellow indicates β -sheet, white and cyan indicate coil and turn, respectively; the semitransparent molecular surface represents the excluded volume of the protein. Green and yellow spheres depict potassium and chloride ions, respectively; water is not shown. Coordinates of the unfolded proteins were randomly selected from the respective all-atom MD trajectories (see Methods); the Q factor of each unfolded conformation was less than 0.1. SI Table S2 characterizes the average shape of the unfolded conformations. The simulated ionic current blockade difference is shown in Figure 1C (main text) (B) Simulated blockade currents for three orthogonal orientations of the titin protein in both folded and unfolded states. This set of simulations was carried out using a bulk electrolyte system shown in SI Figure S2A. (C) Simulated blockade currents for five distinct conformations of unfolded HP35 protein and one folded conformation of HP35. These simulations were carried out using a nanopore system shown in panel A. In panels B and C, the error bars represent the standard errors of mean each computed from a 150 ns ionic current trace averaged in 1.2 ns blocks.

Table S1: Physical properties of the proteins considered in this work. For each protein, N_P , N_N and N denote the number of positively charged, the number of negatively charged and the total number of residues, respectively; q indicates the total net charge of the protein whereas $(N_P+N_N)/N$ is the fraction of charged residues in the protein. MW denotes the protein’s molecular weight whereas J_1 , J_2 and J_3 denote the three principle moments of inertia of the protein’s folded conformation. The ratio of the smallest to the largest moments, J_{\min}/J_{\max} , reports on the shape of the protein. The last column lists PDB ID for folded structure of the protein. The \star , \dagger and $+$ subscript symbols indicate the use of a particular protein in nanochannel, bulk electrolyte and protein folding simulations, respectively.

Protein	N	N_P	N_N	q	$(N_P + N_N)/N$	MW kDa	J_1 Da*nm ²	J_2 Da*nm ²	J_3 Da*nm ²	J_{\min}/J_{\max}	PDB ID
Trp-cage $\star,+$	20	1	2	-1	0.15	2.028	882.0	736.8	475.0	0.539	2JOF
HP35 $\star,+$	34	5	5	0	0.294	3.905	2994.2	2344.3	1427.8	0.477	2F4K
GTT $\star,+$	34	5	5	0	0.294	4.068	3219.6	2601.7	1696.3	0.527	2F21
NTL9 $\star,+$	39	8	5	3	0.333	4.29	2699.1	2262.1	1842.3	0.683	2HBA
BBL $\star,\dagger,+$	47	8	7	1	0.319	5.002	4433.1	3918.7	2668.5	0.602	2WXC
protein G $\star,+$	56	4	8	-4	0.214	6.142	5589.8	4974.0	3078.7	0.551	1MIO
α 3D $\star,+$	73	12	13	-1	0.342	8.107	11885.8	11231.0	4459.6	0.375	2A3D
ubiquitin $\star,+$	76	12	12	0	0.316	8.565	9373.4	8371.4	6160.7	0.657	1UBQ
λ -repressor $\star,\dagger,+$	80	11	10	1	0.263	8.784	10018.8	9831.7	6137.0	0.613	1LMB
BBA \star	28	8	4	4	0.429	3.543	3034.6	2480.9	1164.7	0.384	1FME
titin \dagger	119	12	17	-5	0.244	13.05	37450.0	35769.3	8681.4	0.232	1TIT

Table S2: Physical properties of the unfolded proteins used for all-atom MD simulations. The first seven columns of the table are identical to those of SI Table S1. J_1 , J_2 and J_3 denote the three principle moments of inertia of the protein’s unfolded conformation. The ratio of the smallest to the largest moments, J_{\min}/J_{\max} , reports on the shape of the protein. Five different unfolded conformations were used for the simulations of the HP35 protein reported in SI Figure S1C. The last column indicates the source of the atomic coordinates.

Protein	N	N_P	N_N	q	$(N_P + N_N)/N$	MW kDa	J_1 Da*nm ²	J_2 Da*nm ²	J_3 Da*nm ²	J_{\min}/J_{\max}	Source
Trp-cage $\star,+$	20	1	2	-1	0.15	2.028	3803.0	3208.9	858.3	0.226	Ref. 1
HP35,Conf 1 $\star,+$	34	5	5	0	0.294	3.905	8369.7	8009.3	2486.4	0.297	Ref. 2
HP35,Conf 2 $\star,+$	34	5	5	0	0.294	3.905	10306.7	9233.4	1800.2	0.175	Ref. 2
HP35,Conf 3 $\star,+$	34	5	5	0	0.294	3.905	4791.4	4315.2	2243.9	0.468	Ref. 2
HP35,Conf 4 $\star,+$	34	5	5	0	0.294	3.905	5970.6	5378.9	2030.9	0.34	Ref. 2
HP35,Conf 5 $\star,+$	34	5	5	0	0.294	3.905	8816.8	8154.2	2271.0	0.258	Ref. 2
GTT $\star,+$	34	5	5	0	0.294	4.068	6695.1	5771.7	2564.0	0.383	Ref. 2
NTL9 $\star,+$	39	8	5	3	0.333	4.29	4382.2	3862.0	2096.2	0.478	Ref. 1
BBL $\star,\dagger,+$	47	8	7	1	0.319	5.002	7527.7	5034.9	3980.1	0.529	Ref. 1
protein G $\star,+$	56	4	8	-4	0.214	6.142	8785.2	7869.8	3765.2	0.429	Ref. 1
α 3D $\star,+$	73	12	13	-1	0.342	8.107	23436.3	22635.9	7517.9	0.321	Ref. 1
ubiquitin $\star,+$	76	12	12	0	0.316	8.565	12992.3	10528.4	9910.9	0.763	Ref. 3
λ -repressor $\star,\dagger,+$	80	11	10	1	0.263	8.784	12426.5	11482.0	8694.1	0.7	Ref. 1
BBA \star	28	8	4	4	0.429	3.543	5911.5	4498.2	2039.2	0.345	Ref. 1
titin \dagger	119	12	17	-5	0.244	13.05	35986.0	28284.5	23544.9	0.654	Ref. 4

Supplementary Methods 1: Calculation of the ion concentration and mobility distributions. The distributions of ion concentration and mobility were determined from simulations of bulk electrolyte systems, such as the systems shown in Figure S2A. The volume of a bulk electrolyte system was partitioned into 1 Å-thick shells of complex shapes defined by the minimum and maximum distance from the protein surface, $r_{n,\min} < r \leq r_{n,\max}$, where $r_{n,\max} = n \times 1 \text{ \AA}$, $r_{n,\min} = r_{n,\max} - 1 \text{ \AA}$ and n labels consecutive shells. Then the numbers of ions and water molecules, $N_{\text{ion},n}$ and $N_{\text{water},n}$, were computed for each shell. The ion concentration versus distance from protein surface was determined as $C_{\text{ion}}(r_n) = 55.5 \text{ M} * N_{\text{ion},n}/N_{\text{water},n}$, with $r_n = r_{n,\max} - 0.5 \text{ \AA}$. Similar to the method described in a previous study,⁵ the ion mobility was determined as $\mu_{\text{ion}}(r_n) = (v_{\text{ion}}(r_n) - v_{\text{water}}(r_n))/E$, where $v_{\text{ion}}(r_n)$ and $v_{\text{water}}(r_n)$ are the average components of ion and water velocities along the external electric field E in the n 's shell. The instantaneous velocity of each ion or water molecule in each shell bin was calculated as $v = (z_{j+1} - z_j)/\delta t$, where z_{j+1} and z_j indicate the z coordinates of the ion or water molecule in the $(j+1)$ th and j th frame of the trajectory and δt denotes the time interval between the consecutive trajectory frames. In all simulations, external electric field was applied along the z axis. Then the average velocity in each shell bin was determined by averaging over all ions or water molecules present in that bin and over all frames of the MD trajectory. Figure S2B shows the resulting dependences of the relative ion concentration and ion mobility on distance from the protein surface.

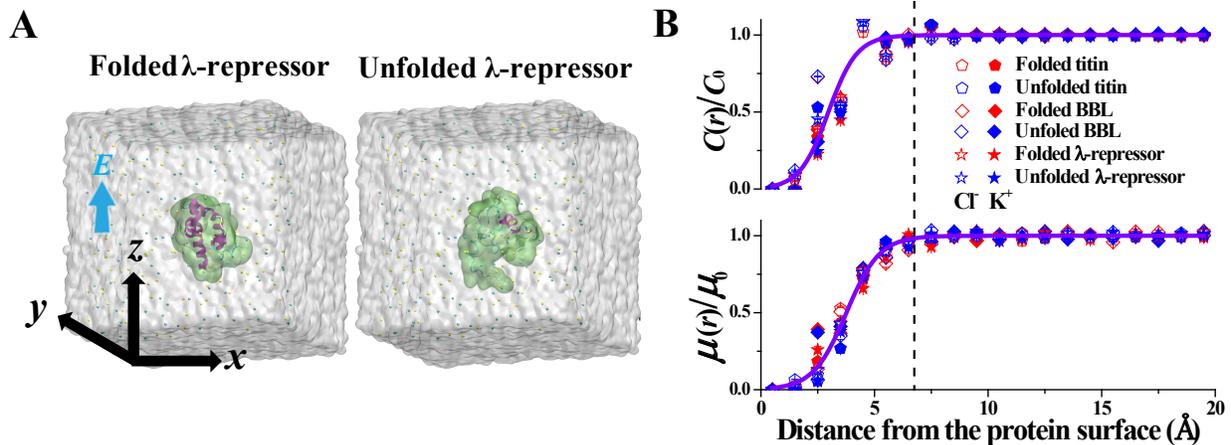


Figure S2: Ion concentration and mobility distributions in bulk electrolyte systems. (A) Typical simulation systems containing a protein (λ -repressor is shown) in either folded or unfolded conformation surrounded by a rectangular volume of 2 M KCl electrolyte solution. The volume occupied by solution is approximately shown as a white semitransparent surface. Cyan and yellow spheres indicate potassium and chloride ions, respectively. The protein molecules are shown using a NewCartoon representation colored according to the secondary structure of the folded state and semitransparent molecular surface (green) indicating the excluded volume of the protein. (B) Concentration $C(r)$ and mobility $\mu(r)$ of K^+ and Cl^- ions *versus* distance from protein surface r normalized by the corresponding bulk values, C_0 and μ_0 . Data are shown for three proteins, BBL, λ -repressor and titin, simulated in both folded and unfolded conformations. Unfolded conformation of titin protein was taken from Ref. 4. The dashed lines in panel B indicate distance r^* where the $C(r)$ and $\mu(r)$ dependences reach their corresponding bulk values.

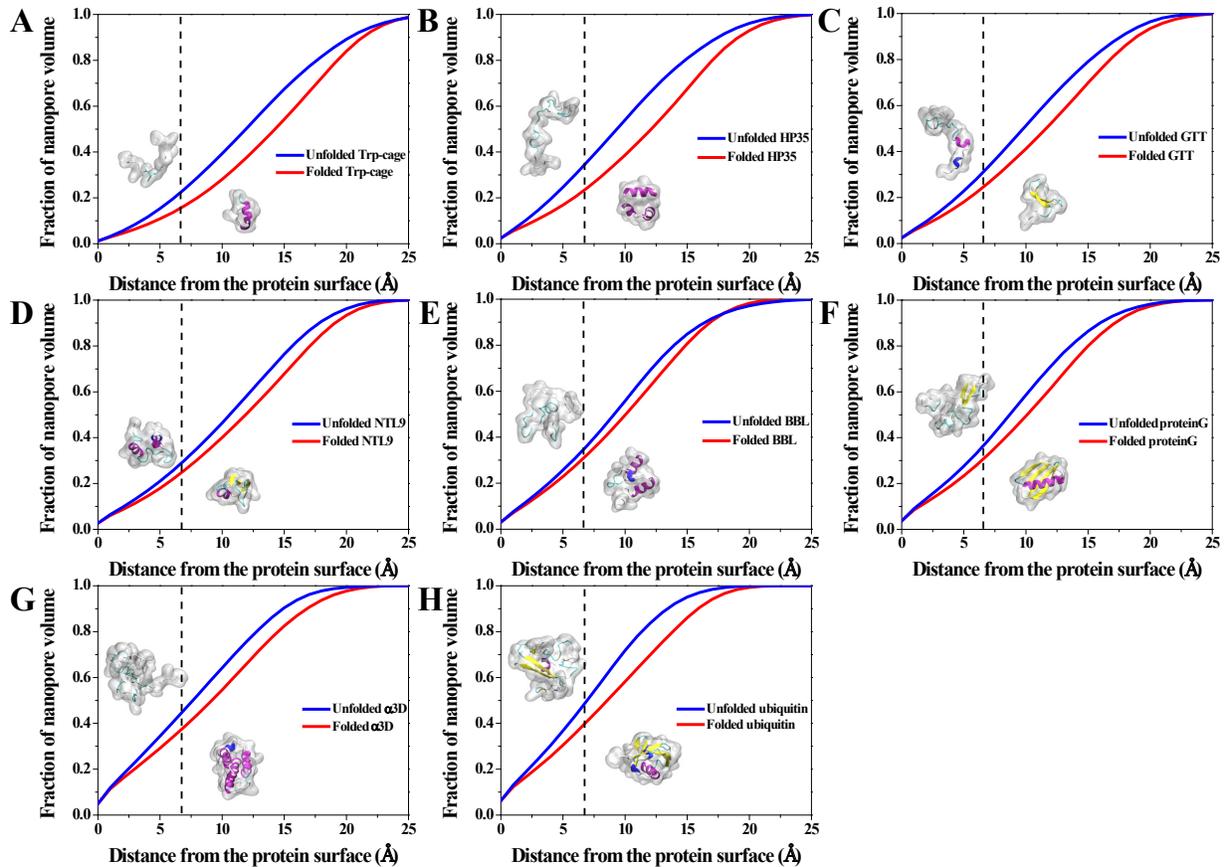


Figure S3: The fraction of the nanopore volume occupied by the protein and the solvent surrounding the protein within distance r of the protein surface as a function of r . The data are shown for the folded and unfolded conformations of the following eight proteins: Trp-cage (A), HP35 (B), GTT (C), NTL9 (D), BBL (E), protein G (F), α 3D (G) and ubiquitin (H). Insert images illustrate the folded (below the red line) and unfolded (above the blue line) conformations of the proteins. The dashed lines schematically indicate the distance $r^* = 6.6 \text{ \AA}$ at which $C(r)$ and $\mu(r)$ attain their bulk values.

Supplementary Methods 2: Theoretical model of ionic current blockades.

As the local conductivity $\sigma(r)$ is proportional to both local ion mobility $\mu(r)$ and ion concentration $C(r)$, the normalized conductivity as a function of distance to the protein surface is obtained as

$$\sigma(r)/\sigma_0 = \frac{\sum \mu(r) * C(r)}{\sum \mu_0 * C_0}, \quad (1)$$

where C_0 , μ_0 and σ_0 are the bulk ion concentration, mobility and conductivity and the sum runs over the two ions types (K^+ and Cl^-). The dependence of the normalized conductivity on distance to the protein, Figure 3A of the main text, can be approximated by a smooth step function

$$\sigma(r)/\sigma_0 = \frac{1}{2}(1 + \tanh(\frac{r - a}{b})) \quad (2)$$

where parameters a and b determine the location and the steepness of the step function. The least square fit to the bulk electrolyte MD data yielded the following numerical values: $a = 3.942 \text{ \AA}$ and $b = 0.809 \text{ \AA}$. For a given protein conformation, the ionic current I was computed by first partitioning the volume of the system into 1 \AA^3 cubes and then determining the distance from the center of each cube to the nearest protein surface using the VolMap plugin (version 1.1) of VMD (version 1.9.1). Knowing the distance of each cube to the protein surface, the local conductivity $\sigma(r_{ij})$ of the entire system can be determined by applying Eq. 2, from which the ionic current I at a voltage bias V is determined following the method described in Figure 3A and the main text. In practice, however, parameter a of Eq. 2 needed adjustment because the $\sigma(r)$ dependence obtained from analysis of the MD trajectories, Figure 2B, was determined relative to the atomic coordinates of the proteins whereas the VolMap plugin returns the distance map relative to the molecular surface of the protein. Therefore, the final value of parameter $a = 3.3 \text{ \AA}$ was determined by least square optimization of the model's predictions with respect to results of all-atom MD simulations of nanopore and bulk electrolyte systems. Optimization against only the bulk electrolyte systems yielded a very similar value ($a = 3.2 \text{ \AA}$). Figure S4 compares predictions of the

model with results of all-atom MD simulations for the bulk electrolyte systems. With small modifications, the model is able to predict ionic current blockades at arbitrary electrolyte concentration and temperature (see SI Figure S8 and Figure S9).

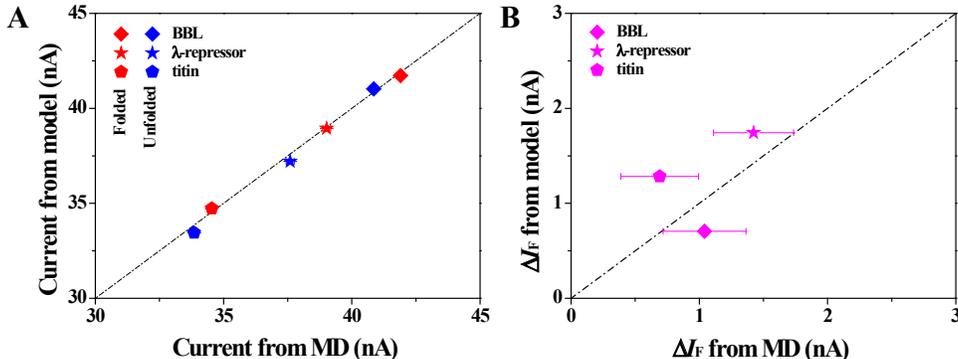


Figure S4: Theory *versus* simulation for ionic current prediction in bulk electrolyte systems. (A) The ionic current predicted by the model *versus* the ionic current measured in MD simulations for three proteins in folded and unfolded conformations. SI Figure S2A illustrates a typical simulation system. In all cases, electric field of 71.43 mV/nm was applied along the z axis; the KCl concentration was 2 M. The protein conformations used for the theoretical model calculations was the same as in the all-atom MD simulations of the ionic current. The blockade current from our model was calculated by the same method for nanopore systems described above but in the absence of the nanopore. Dashed line indicates perfect agreement between model and simulation. Error bars represent standard errors. (B) Same as in panel A but for ΔI_F .

Because our theoretical model assumes the same average ion distribution around a protein regardless of the protein charge, it is expected to fail in the case highly charged proteins. To find the limits of our model, we simulated a moderately charged protein BBA in our nanochannel setup along with GTT and NTL9 proteins in four different conformations, Figure S5. Comparison between the predictions of the model and results of all-atom MD simulations, Figure S5A–C, show that as the charge of the protein increases, our model starts to overestimate slightly the ionic current. Figure S5 (panels D and E) plots the relative error of the model with respect to all-atom MD simulation as a function of the average protein charge, $|q|/N$, and the fraction of charged residues, $(N_P + N_N)/N$. The relative error increases with $(N_P + N_N)/N$, although specific protein conformations are also seen to influence the agreement.

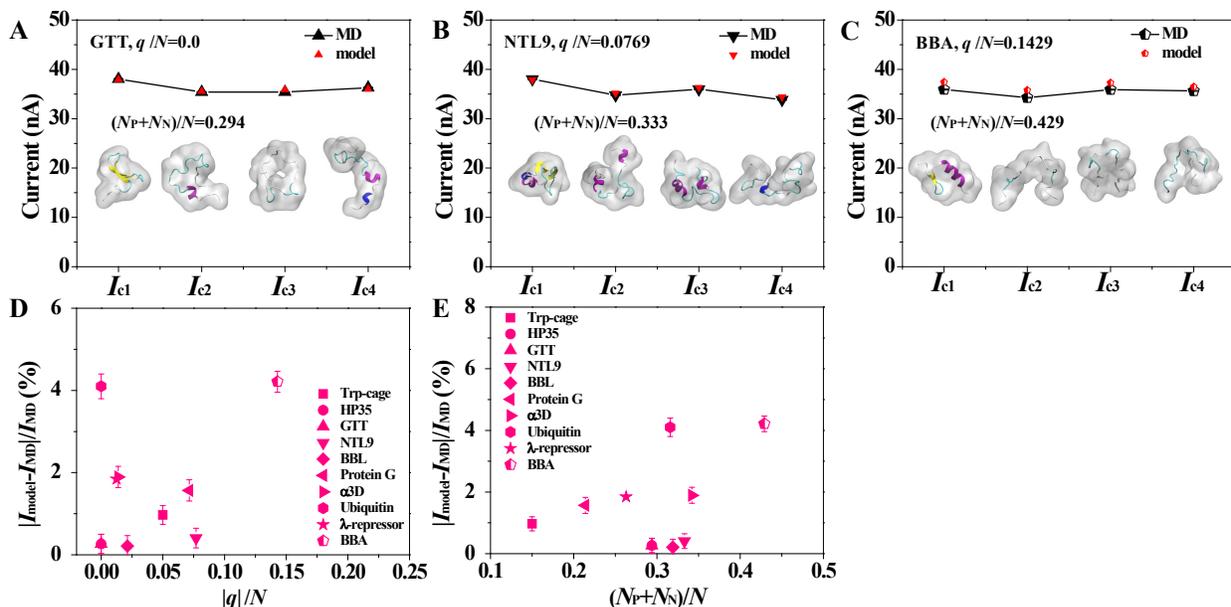


Figure S5: The effect of protein charge on agreement between the theoretical model and all-atom MD simulation. (A–C) Ionic currents computed using the theoretical model (red) and MD simulations (black) for four conformations of GTT (panel A), NTL9 (panel B) and BBA (panel C). Images in each panel illustrate the four conformations of the proteins. The ratio of the protein’s total charge q or the number of charged residues ($N_P + N_N$) to the total number of residues N is specified in each panel. The ionic current were computed using the same setup as in Figure 1A and Figure S1. (D–E) Percent difference error between predictions of the model I_{model} and results of all-atom MD simulation I_{MD} versus $|q|/N$ (panel D) and $(N_P + N_N)/N$ (panel E) for ten folded proteins, nine from Figure 1A and Figure S1 and BBA.

Supplementary Methods 3: Calculations of rotation-averaged ionic currents.

To determine the effect of protein orientation on the ionic current, the folded and unfolded conformations of the nine weakly charged proteins (Figure 1A and Figure S1) were placed at the center of the nanochannel in different orientations. 1728 different protein orientations were generated by rotating the protein about the z , x , and y axes in 30° increments from 0° to 360° . The center of mass of the protein was located within 0.3 \AA from the nanopore in all conformations considered. For each protein orientation, the 3D distance map was generated from the atomic coordinates of the protein and the nanochannel as described in SI Methods 2. The ionic current for each conformation was computed using our theoretical model (main text Figure 3A), producing 1728 independent ionic current values for the folded and unfolded states of each protein at the same bias and ion concentration conditions as in Figure 1A and Figure S1. Figure S6 shows the variation of the nanochannel ionic current with protein orientation. Similar rotation average values were obtained when protein orientations were randomly chosen, Figure S7A–C. Figure S7D–F compares the rotation-averaged currents to the currents obtained by averaging the ensembles of folded and unfolded conformations observed in MD simulations of protein folding, see main text Figure 5 and SI Methods 4.

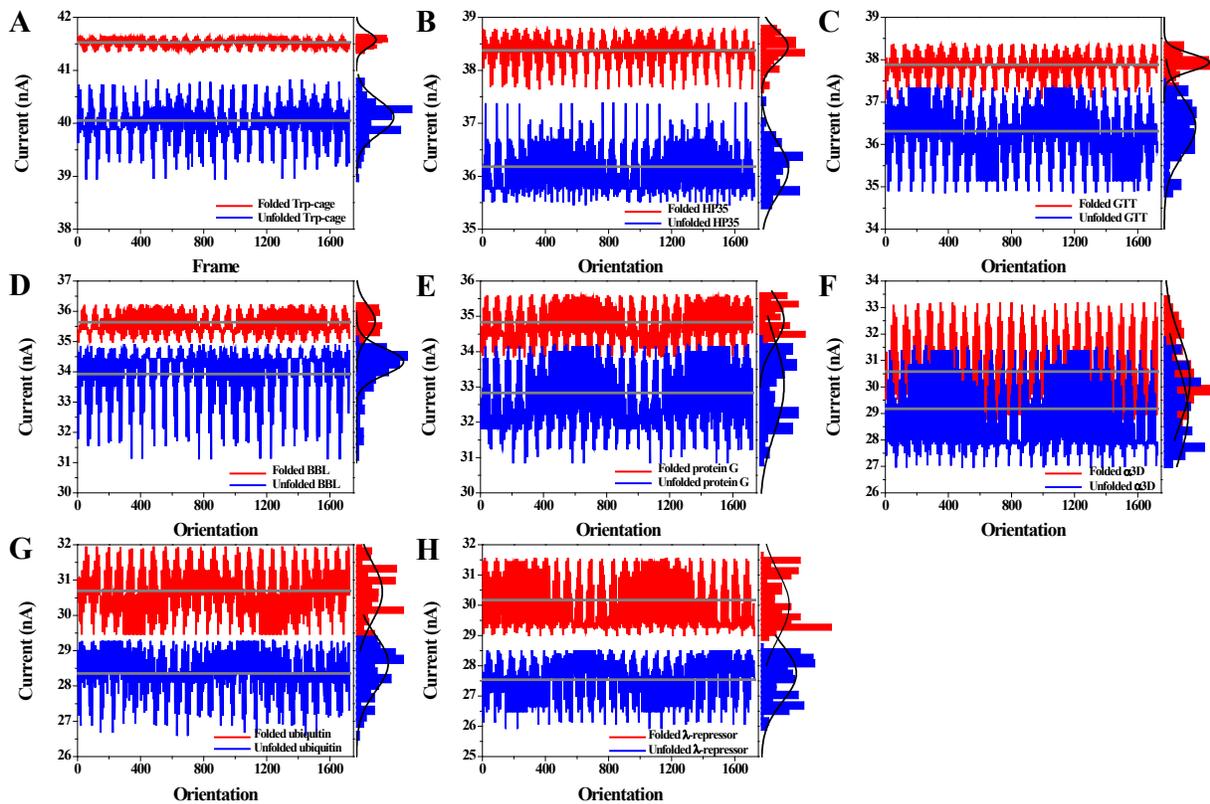


Figure S6: The effect of protein orientation on the blockade current. Shown is the blockade current, I , for 1728 unique orientations of Trp-cage (panel A), HP35 (panel B), GTT (panel C), BBL (panel D), protein G (panel E), α 3D (panel F), ubiquitin (panel G) and λ -repressor (panel H) in the folded (red) and unfolded (blue) states. Horizontal lines indicate the rotation-averaged blockade currents. The blockade currents were computed using a theoretical model featured in the main text Figure 3. For each protein state, the plot also shows an all-point histogram of the currents (off the right axis) and its fit to a Gaussian distribution (black line).

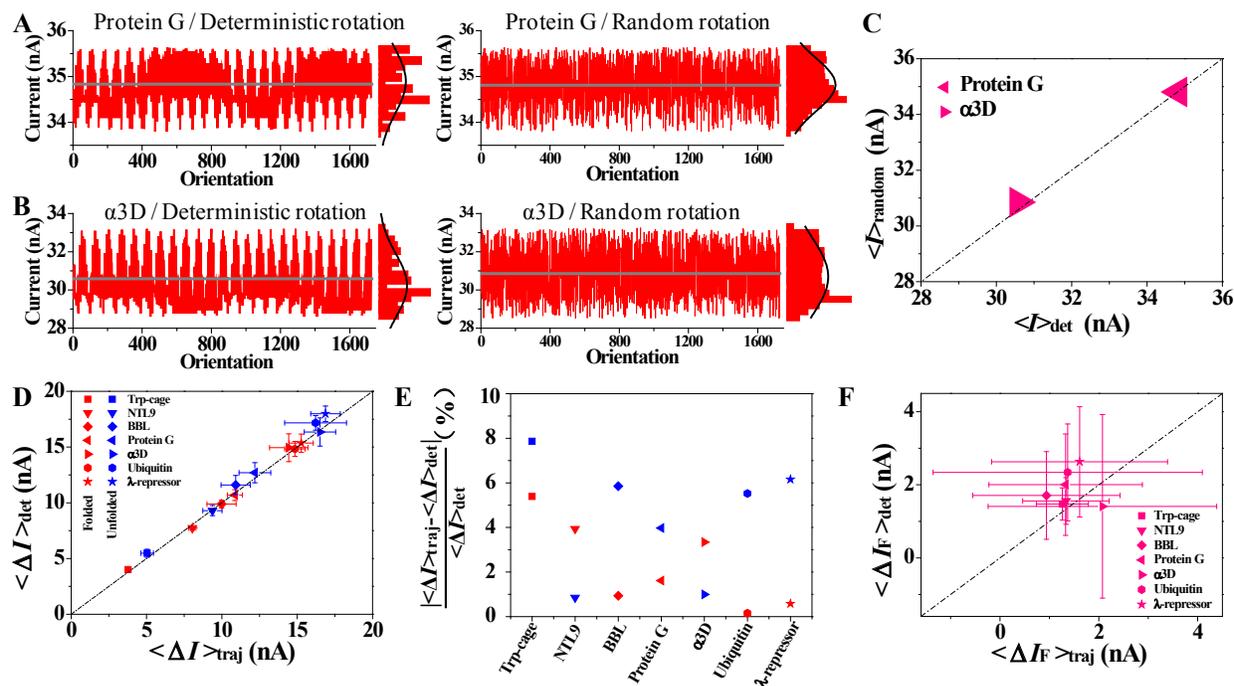


Figure S7: Validation of the protocol for determination of the rotation-averaged ionic currents. (A–B) Blockade ionic currents computed for 1728 orientations of protein G (panel A) and α 3D (panel B) produced by rotating the protein about axis z , x , and y in 30° increments (left column) or by randomly selecting the rotation angles (right column). Horizontal lines indicate the rotation-averaged blockade currents. For each protein state, the plot also shows an all-point histogram of the currents (off the right axis) and its fit to a Gaussian distribution (black line). (C) Comparison of the rotation-averaged blockade currents obtained using the deterministic (horizontal axis) and random (vertical axis) methods of protein orientation selection. Dashed line indicates perfect agreement between the two methods. (D) Comparison of the rotation-averaged blockade current amplitudes, $\langle \Delta I \rangle = I_0 - \langle I \rangle$, obtained using the deterministic rotation-averaging protocol (vertical axis) and by averaging over ensembles of conformations observed in MD simulations of protein folding (horizontal axis). In the analysis of protein folding trajectories, protein conformations of Q values greater than 0.9 or smaller than 0.1 were considered to represent folded and unfolded states, respectively, except for BBL and λ -repressor, see caption to Figure S17 for detailed description. Dashed line indicates perfect agreement between the two rotation-averaging methods. Error bars indicate standard deviations of the current values associated with conformational sampling. (E) Percent difference error between the blockade current values obtained using different averaging methods. The greater percent difference error for the unfolded state can be attributed to the conformational heterogeneity of the state. (F) Same as in panel D but for the ΔI_F values. In all panels, the blockade current for each orientation was computed using a theoretical model featured in the main text Figure 3.

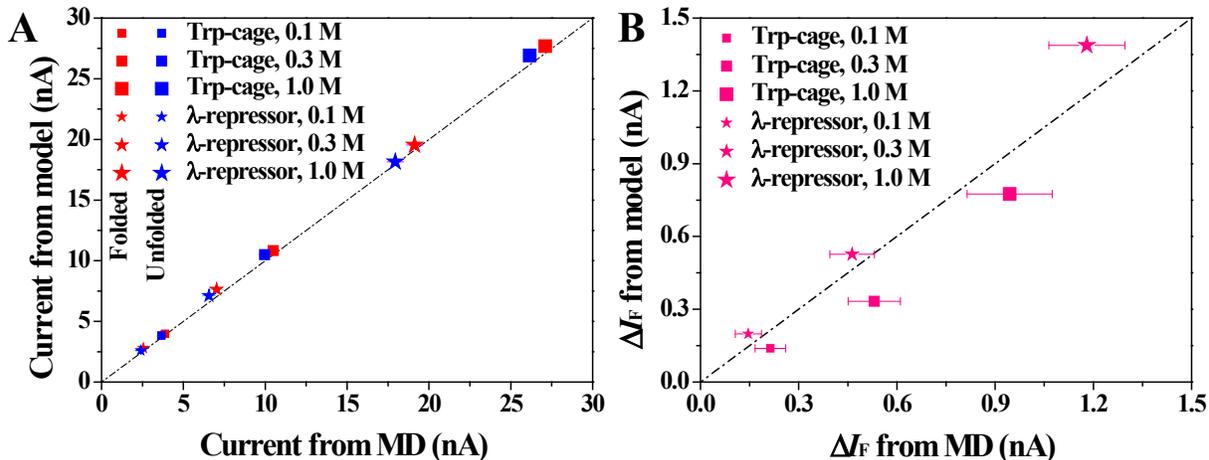


Figure S8: Theory *versus* simulation for ionic current blockades at variable electrolyte conditions. (A) The blockade current predicted by the theoretical model *versus* the blockade current measured in MD simulations for the folded and unfolded Trp-cage and λ -repressor at 0.1, 0.3 and 1.0 M KCl electrolyte. Except electrolyte concentration, the simulation conditions were identical to those featured in main text Figure 1. Dashed line indicates perfect agreement between model and simulation. Theoretical calculation of the blockade currents at an arbitrary electrolyte concentration was done using the same functional dependence of $\sigma(r)/\sigma_0$ as at 2 M KCl, main text Figure 3A; the bulk electrolyte conductivity value σ_0 was set to the value measured in all-atom MD simulations of a nanochannel filled with electrolyte of the target concentration and in the absence of protein. Error bars represent standard errors. (B) Same as in panel A but for ΔI_F .

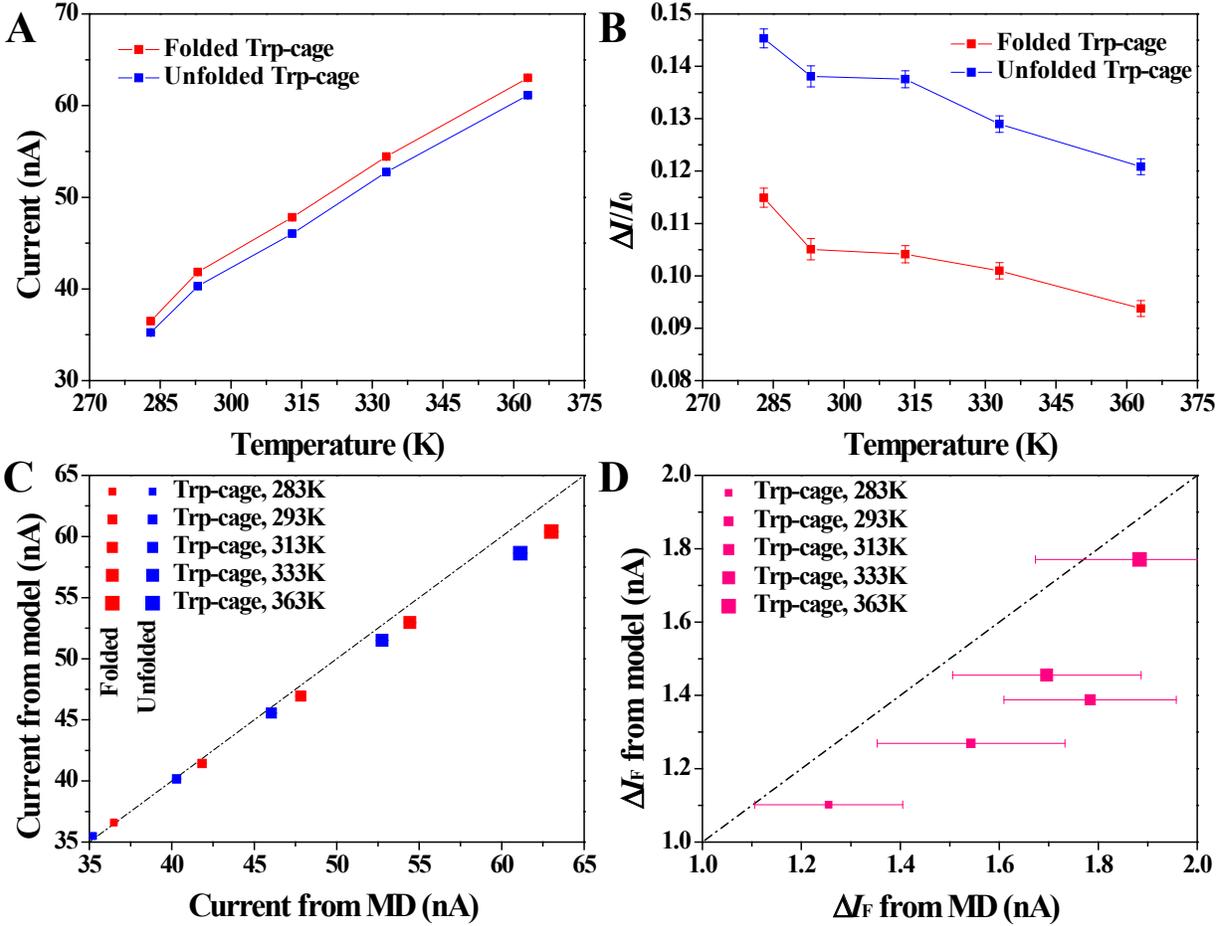


Figure S9: The effect of temperature on the ionic current difference produced by protein folding. (A) Blockade currents measured in all-atom MD simulations of a 6-nm-diameter nanochannel containing a Trp-cage protein in either folded or unfolded states as a function of temperature. Except temperature, the simulation conditions were identical to those featured in main text Figure 1. (B) Temperature dependence of the relative current blockades $\Delta I/I_0$ measured in all-atom MD simulations. Here, I_0 denotes the open pore current and $\Delta I = I_0 - I_f$ (or $I_0 - I_{unf}$) denotes the current blockade amplitude produced by the folded (or unfolded) Trp-cage protein. (C) The blockade current predicted by the theoretical model *versus* the blockade current measured in MD simulations for the folded and unfolded Trp-cage at different temperatures. Theoretical calculation of the blockade currents at arbitrary temperature was done using the same functional dependence of $\sigma(r)/\sigma_0$ as at 2 M KCl, main text Figure 3A; the bulk electrolyte conductivity value σ_0 was set to the value measured in all-atom MD simulations of a nanochannel system filled with 2 M KCl, at the target temperature and in the absence of protein. Dashed line indicates perfect agreement between model and simulations. (D) Same as in panel C but for ΔI_F . Error bars represent standard errors.

Supplementary Methods 4: Real-time detection of folding-unfolding transitions by ionic current measurement.

Our calculations of ionic current signatures produced by folding or unfolding of a protein in a nanopore were based on the millisecond all-atom, explicit solvent MD trajectories reported by the David E. Shaw group.^{1,3} To make our analysis computationally tractable, the all-atom simulation trajectories were pruned with a time interval ranging from 0.1 μ s to 1.0 μ s, making sure that at least a thousand of protein conformations were preserved in each data set. Interestingly, our sampling of the MD trajectories mirrors that of advanced ionic current recording systems of 1 to 5 MHz bandwidth.⁶

To compute an ionic current trace, each protein conformation was first translated to have its center of mass at the center of the nanochannel. Following that, the ionic current I was computed using the theoretical model described in main text Figure 3 and SI Methods 2 at the same bias and ion concentration conditions as in Figure 1A and Figure S1. For clarity, all resulting ionic current traces were smoothed with a low-pass filter.

To characterize the blockade currents of folding intermediates (main text Figure 5D–F, Figure S19 and Figure S20), we analyzed millisecond all-atom MD trajectories^{1,3} and randomly selected 1000 protein conformations for each of the twenty Q value intervals (from 0 to 1, in 0.05 increments) from replica exchange simulations² for analysis. The theoretical model was used to compute the ionic current for each of the conformations; the resulting set of ionic current values was averaged to determine the mean value and the standard error of mean in each of the Q value intervals. Similar analysis was performed to determine the ionic current of folding intermediates using R_g and RMSD as the folding coordinate.

We computed the fraction of native contacts, Q , using Eq. 13 of Ref. 7 with a cutoff of 8 Å excluding $i - j \leq 4$ where i and j are residue indices.⁷ For completeness, we reproduce Eq. 13 below:

$$Q = \frac{\sum_{i < j-1} \theta(r_c - r_{ij}^N) \exp \left[-\frac{(r_{ij} - r_{ij}^N)^2}{2\sigma_{ij}^2} \right]}{\sum_{i < j-1} \theta(r_c - r_{ij}^N)}. \quad (3)$$

In the above expression, r_{ij}^N and r_{ij} denote distances between C α atoms of residues i and j in the native (N) and current structures, respectively; θ is the Heaviside step function; $r_c = 8 \text{ \AA}$; $\sigma_{ij} = |i - j|^{0.15} \text{ \AA}$.

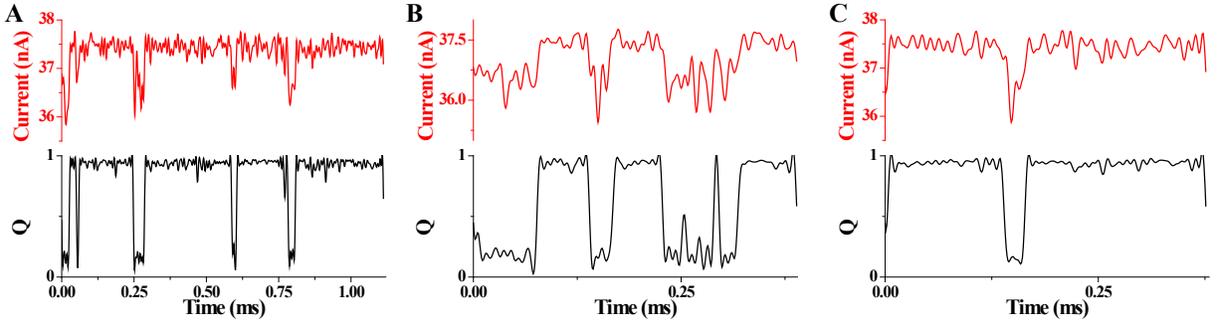


Figure S10: Folding-unfolding transitions of NTL9 observed by ionic current through the nanopore. Panels A, B and C correspond to three independent all-atom MD simulations of NTL9. The ionic current recording from a single protein undergoing folding-unfolding transitions are shown at the top panels. The corresponding changes in the folding parameter Q are shown at the bottom panels. The traces shown were sampled at 1 MHz and smoothed with a low-pass filter at 200 kHz.

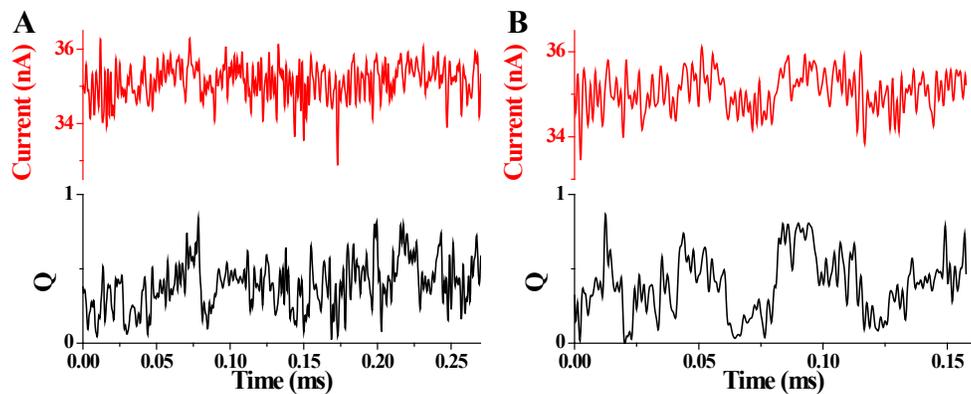


Figure S11: Folding-unfolding transitions of BBL observed by ionic current through the nanopore. Panels A and B correspond to two independent all-atom MD simulations of BBL. The ionic current recording from a single protein undergoing folding-unfolding transitions are shown at the top panels. The corresponding changes in the folding parameter Q are shown at the bottom panels. The traces shown were sampled at 5 MHz and smoothed with a low-pass filter at 1 MHz.

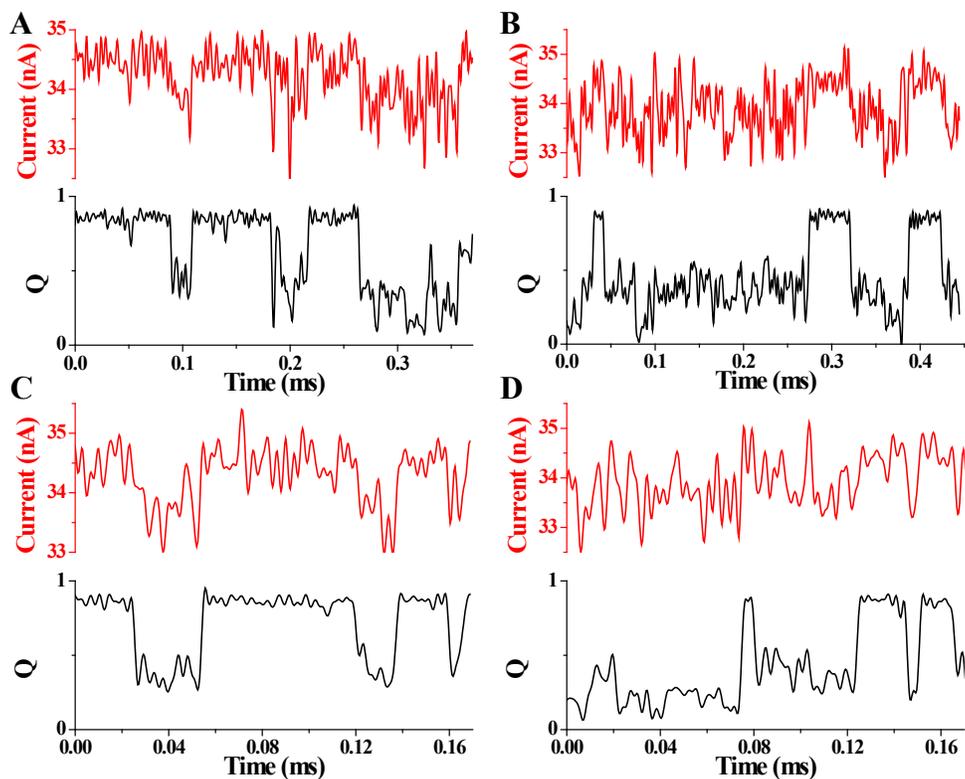


Figure S12: Folding-unfolding transitions of protein G observed by ionic current through the nanopore. Panels A–D correspond to four independent all-atom MD simulations of protein G. The ionic current recording from a single protein undergoing folding-unfolding transitions are shown at the top panels. The corresponding changes in the folding parameter Q are shown at the bottom panels. The traces shown were sampled at 2.5 MHz and smoothed with a low-pass filter at 500 kHz.

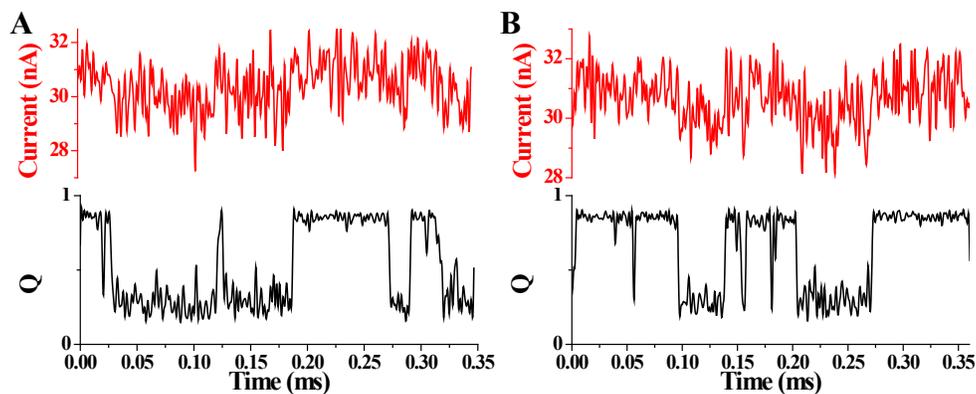


Figure S13: Folding-unfolding transitions of $\alpha 3D$ observed by ionic current through the nanopore. Panels A and B correspond to two independent all-atom MD simulations of $\alpha 3D$. The ionic current recording from a single protein undergoing folding-unfolding transitions are shown at the top panels. The corresponding changes in the folding parameter Q are shown at the bottom panels. The traces shown were sampled at 3.3 MHz and smoothed with a low-pass filter at 670 kHz.

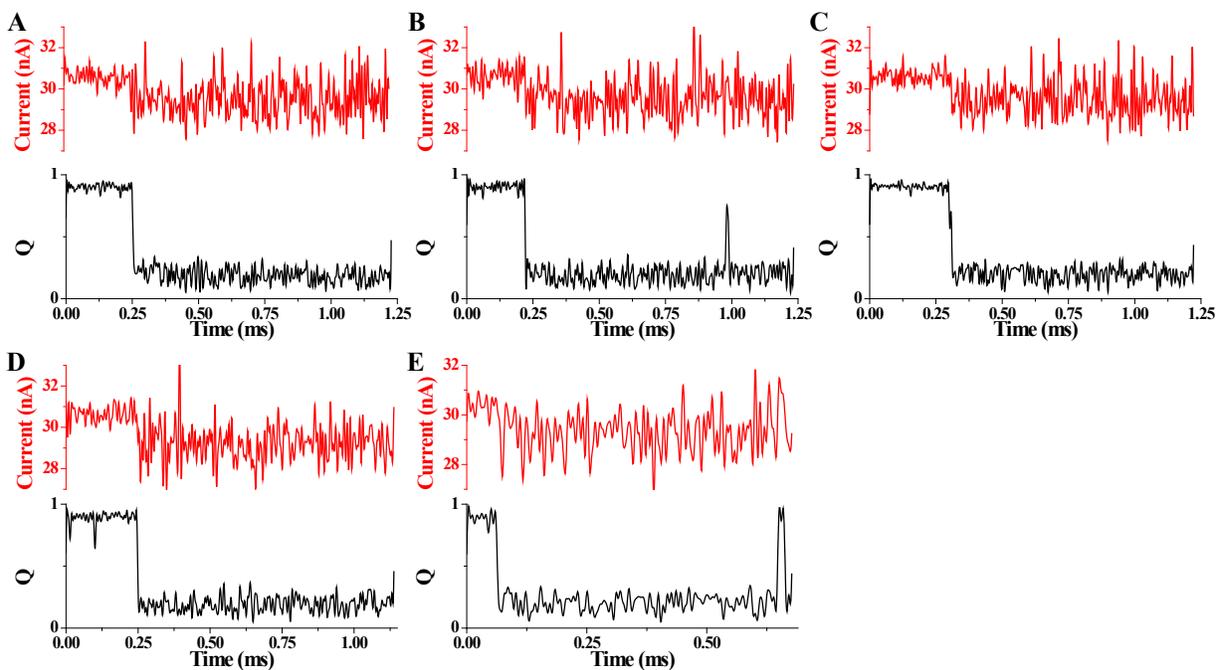


Figure S14: Folding-unfolding transitions of ubiquitin observed by ionic current through the nanopore. Panels A–E correspond to five independent all-atom MD simulations of ubiquitin. The ionic current recording from a single protein undergoing folding-unfolding transitions are shown at the top panels. The corresponding changes in the folding parameter Q are shown at the bottom panels. The traces shown were sampled at 1 MHz and smoothed with a low-pass filter at 200 kHz.

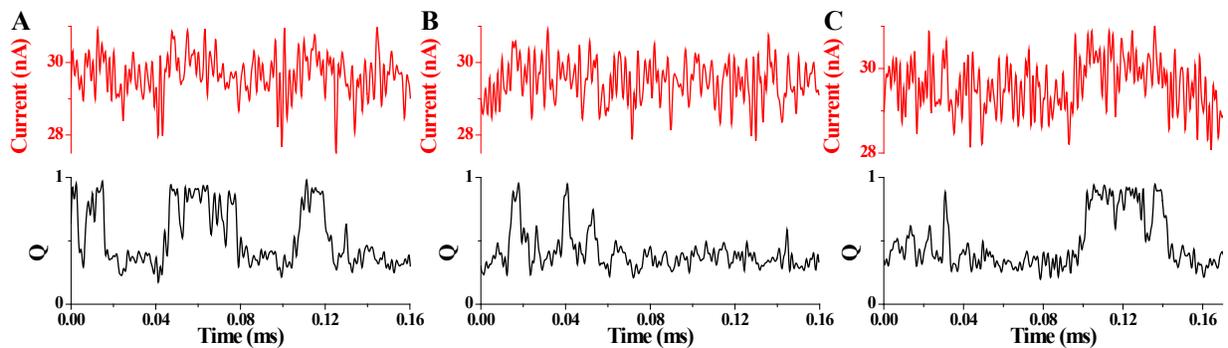


Figure S15: Folding-unfolding transitions of λ -repressor observed by ionic current through the nanopore. Panels A, B and C correspond to three independent all-atom MD simulations of λ -repressor. The ionic current recording from a single protein undergoing folding-unfolding transitions are shown at the top panels. The corresponding changes in the folding parameter Q are shown at the bottom panels. The traces shown were sampled at 5 MHz and smoothed with a low-pass filter at 1 MHz.

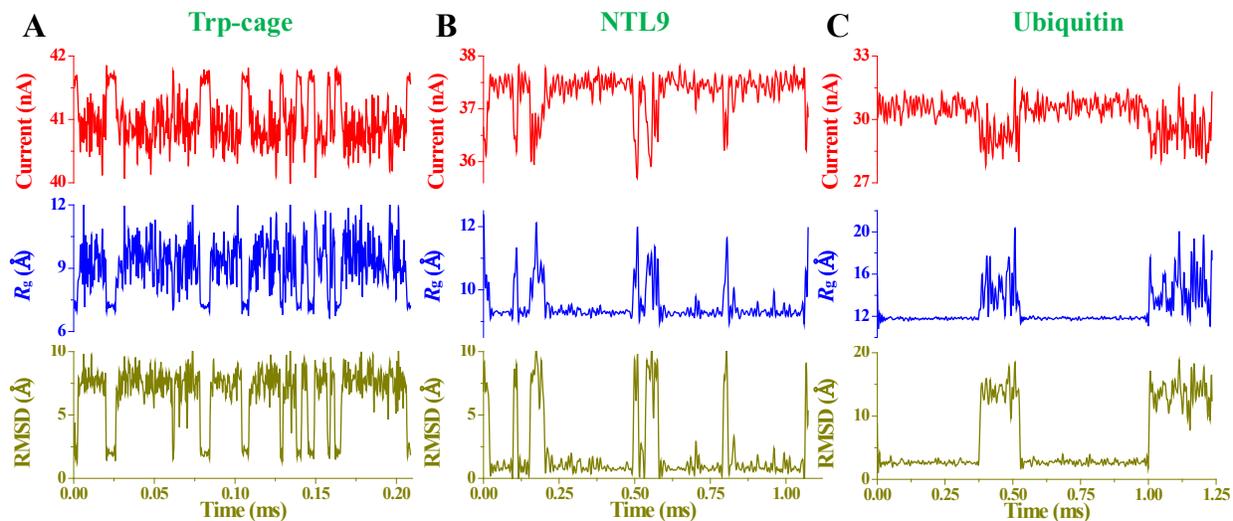


Figure S16: Real-time detection of protein folding-unfolding transitions. (A–C) Simulated ionic current recording from a single protein undergoing folding-unfolding transitions (top panels) for Trp-cage (A), NTL9 (B) and ubiquitin (C), respectively. The corresponding changes in the radius of gyration (R_g) and RMSD are shown at the middle panels and bottom panels, respectively. The traces shown for Trp-cage were sampled at 10 MHz and smoothed with a low-pass filter at 2 MHz. And the traces shown for NTL9 and ubiquitin were sampled at 1 MHz and smoothed with a low-pass filter at 200 kHz.

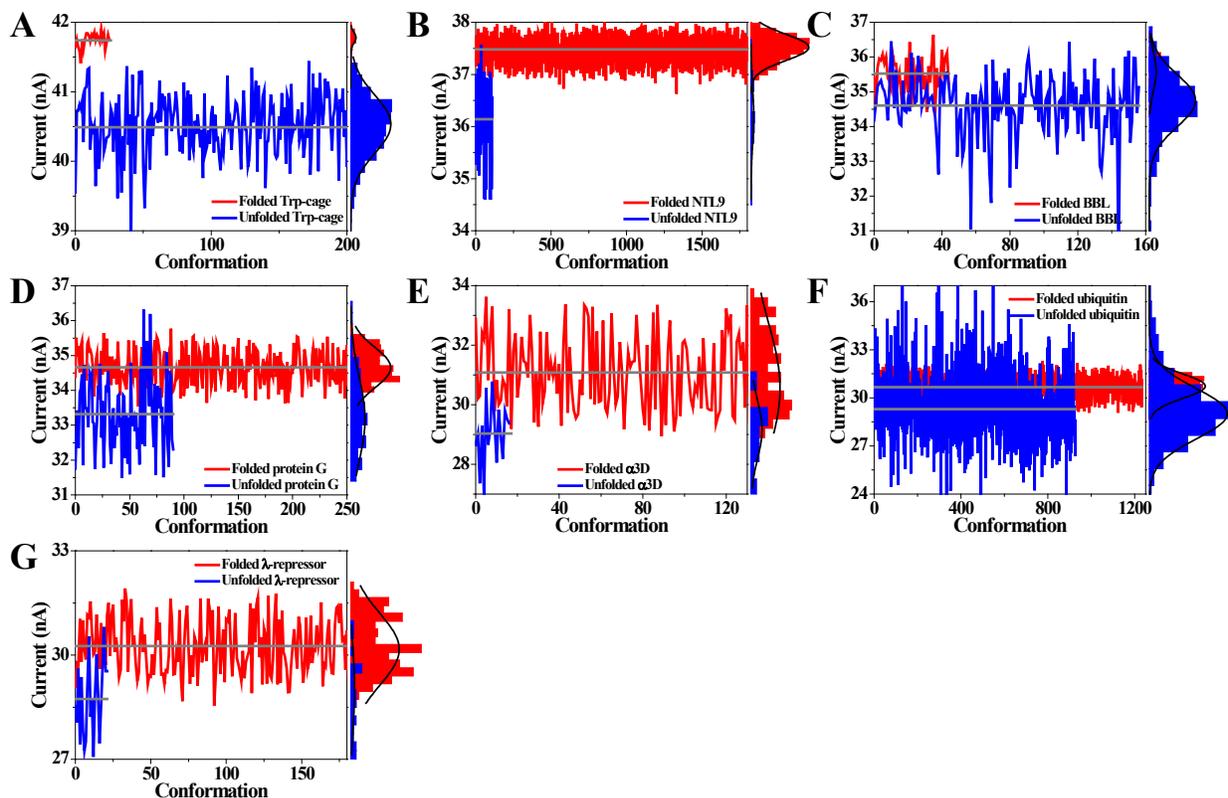


Figure S17: The effect of protein conformation on ionic current blockade in folded and unfolded states. Shown is the blockade current, I , for ensembles of folded ($Q > 0.9$, red) and unfolded ($Q < 0.1$, blue) conformations of Trp-cage (A), NTL9 (B), BBL (C), protein G (D), α 3D (E), ubiquitin (F) and λ -repressor (G). The conformations were selected from all-atom MD trajectories of protein folding-unfolding transitions reported by the D.E. Shaw group.^{1,3} For each conformation, the blockade current was computed using our theoretical model. Because of limited sampling, we used a more relaxed criterion to select folded conformations of BBL ($Q > 0.8$) and unfolded conformations of λ -repressor ($Q < 0.2$). Horizontal lines indicate the averaged blockade currents. For each protein state, the plot shows an all-point histogram of the currents (off the right axis) and its fit to a Gaussian distribution (black line).

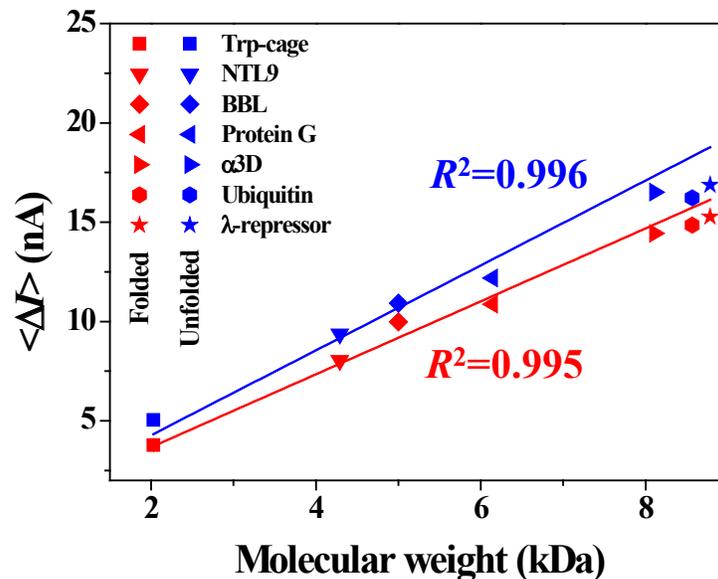


Figure S18: Blockade current amplitude $\langle \Delta I \rangle = I_0 - \langle I \rangle$ versus molecular weight of the proteins averaged over ensembles of folded ($Q > 0.9$) and unfolded ($Q < 0.1$) conformations observed in all-atom MD simulations of folding-unfolding transitions (see SI Figure S17). Because of limited sampling, more relaxed selection criteria was applied to folded conformations of BBL ($Q > 0.8$) and unfolded conformations of λ -repressor ($Q < 0.2$). Lines represent linear fit to the data. The corresponding plot of the ΔI_F value is shown in main text Figure 4D.

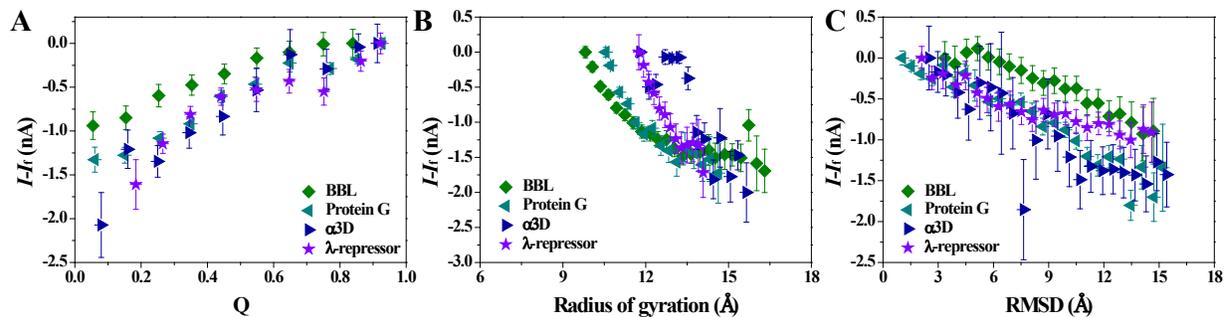


Figure S19: Protein folding intermediates characterization by ionic current. (A–C) The average difference between the current blockades produced by a folding intermediate of that protein (I) and a fully folded protein (I_f). In panels A,B and C, the protein folding intermediates are categorized according to their Q values, radius of gyration and RMSD from the native folded states. Error bars represent standard errors. Protein conformations used for the analysis derived from the millisecond folding trajectories of BBL, protein G, α 3D and λ -repressor reported by the D.E. Shaw group.¹

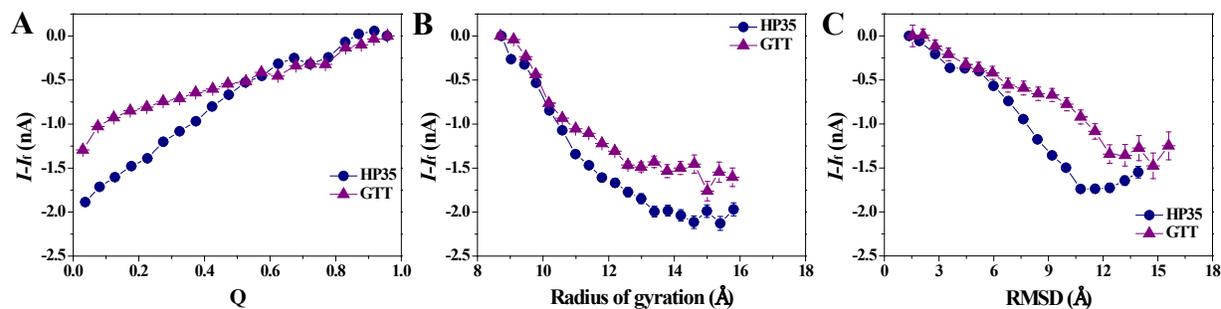


Figure S20: Protein folding intermediates of HP35 and GTT observed by ionic current. (A–C) The average difference between the current blockades produced by a folding intermediate of that protein (I) and a fully folded protein (I_f). In panels A, B and C, the protein folding intermediates are categorized according to their Q values, radius of gyration and RMSD from the native folded states. Protein conformations used for the analysis were obtained from replica-exchange simulations of villin head piece (HP35) and WW domain (GTT) folding,² Error bars represent standard errors.

References

1. Lindorff-Larsen, K.; Piana, S.; Dror, R. O.; Shaw, D. E. How Fast-Folding Proteins Fold. *Science* **2011**, *334*, 517–520.
2. Yoo, J.; Aksimentiev, A. Refined Parameterization of Nonbonded Interactions Improves Conformational Sampling and Kinetics of Protein Folding Simulations. *J. Phys. Chem. Lett.* **2016**, *7*, 3812–3818.
3. Piana, S.; Lindorff-Larsen, K.; Shaw, D. E. Atomic-Level Description of Ubiquitin Folding. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 5915–20.
4. Restrepo-Pérez, L.; John, S.; Aksimentiev, A.; Joo, C.; Dekker, C. *Nanoscale*, Published online, doi: 10.1039/C7NR02450A.
5. Belkin, M.; Aksimentiev, A. Molecular Dynamics Simulation of DNA Capture and Transport in Heated Nanopores. *ACS Appl. Mater. Interfaces* **2016**, *8*, 12599–12608.

6. Shekar, S.; Niedzwiecki, D. J.; Chien, C.-C.; Ong, P.; Fleischer, D. A.; Lin, J.; Rosenstein, J. K.; Drndic, M.; Shepard, K. L. Measurement of DNA Translocation Dynamics in a Solid-State Nanopore at 100 ns Temporal Resolution. *Nano Lett.* **2016**, *16*, 4483–4489.
7. Eastwood, M.; Hardin, C.; Luthey-Schulten, Z.; Wolynes, P. Evaluating Protein Structure-Prediction Scheme Using Energy Landscape Theory. *IBM J. Res., Dev.* **2001**, *45*, 475–497.