Supporting Information

Single-Molecule Titration in a Protein Nanoreactor Reveals the Protonation/Deprotonation Mechanism of a C:C Mismatch in DNA

Hang Ren, Cameron G. Cheyne, Aaron M. Fleming, Cynthia J. Burrows*, and Henry S. White*

Department of Chemistry, University of Utah, 315 S 1400 E, Salt Lake City, UT 84112

Table of Contents

S1. Experimental section

Chemicals and Materials. All chemicals were purchased from Sigma-Aldrich (St. Louis, MO) and used as received unless otherwise specified. All solutions were prepared using deionized water (resistivity > 18 MΩ cm). Buffer solutions (0.25 M KCl, 10 mM phosphate, prepared at pH 6.0-7.5) were prepared and filtered using a sterile 0.22-μm Millipore filter before use. Wild-type α-hemolysin (α-HL) was purchased as a lyophilized monomer powder from List Biological Laboratories (Campbell, CA) and diluted with deionized water to 1 mg/mL and stored at -80 °C. The phospholipid 1,2-diphytanoyl-*sn*-glycero-3-phosphocholine (DPhPC) was obtained from Avanti Polar Lipids (Alabaster, AL) and diluted with decane to a concentration of 10 mg/mL. Glass nanopore membranes were fabricated using bench top techniques as previously described and were chemically modified with 2% (v/v) 3-cyanopropyldimetyhlchlorosilane in dry acetonitrile.[1]

The phosphate buffers for kinetic isotope effect were prepared using D_2O instead of deionized water. The pD was measured using a pH meter calibrated with standard buffers in H_2O by subtracting the reading with 0.4.[2]

DNA Synthesis and Purification Procedure. The DNA was synthesized by the DNA Core Facility at the University of Utah utilizing phosphoramidites (Glen Research, Sterling, VA). After synthesis, each oligodeoxynucleotide was cleaved from the synthetic column and deprotected according to the manufacturer's protocols. Purification of the synthesized DNA was performed utilizing an ion-exchange HPLC procedure running a linear gradient of B from 25% to 100% over 30 minutes while monitoring UV absorbance at 260 nm for collection (A = 20mM Tris, 1 M NaCl at pH 7 in 10% CH_3CN/90% H_2O, B =

10% CH_3CN/90% H_2O. Flow rate: 3 mL/ min). The purification salts were removed by dialysis for 36 h. All DNA oligomers were annealed by mixing the 23-mer and 47-mer in a 1:1 ratio at 90 °C for 5 minutes in a water bath and cooled to room temperature over 3 hours. The annealed DNA was stored at -20 °C prior to use.

Ion Channel Recordings. The current-time (*i-t*) measurements were performed using a low noise Nanopatch system provided by Electronic BioSciences (San Diego, CA). The current was sampled at 10 KHz and filtered using a 4-pole Bessel low-pass filter at 1 KHz. A pressure of 40 to 80 mmHg was applied to the inside of the GNM using a gas-tight syringe once the lipid bilayer was formed across the glass nanopore membrane, which aids in the reconstitution of the protein channel. Protein reconstitution was indicated by a single jump in the current of approximately 0.25 pA mV^{-1} at 20 °C in 0.25 M KCl and 10 mM phosphate buffer (pH from 6.0 to 7.5). Duplex DNA was added to the grounded *cis* side of the protein channel prior at a final concentration of ~0.8 μM. A 70 mV bias (*trans* vs *cis*) was applied across the two Ag/AgCl electrodes on either side of the protein channel (trans vs cis) to hold the DNA molecule inside the channel unless otherwise specified. The temperature was controlled by a custom PID thermoelectric cooler (CUI Inc., CP20151) and a thermocouple residing in the external reservoir. All experiments were run at 20.0 ± 0.1 °C unless otherwise specified.

Data Collection and Analysis. The *i-t* blockades lasting longer than 1 ms were identified as duplex DNA, and those less than 1 ms were identified as single-stranded DNA translocations. Events were extracted using QuB software (https://milesculabs.biology.missouri.edu/QuB_Downloads.html).[3] QuB was used to assign one of three states (1: open channel, 2: intra-helical state, 3: extra-helical state) to

continuous *i-t* data using the software's 'idealization' SKM (segmental K-means) method to extract changes in states. Histograms of the duration of each state (intra-helical and extra-helical) were plotted in the log-binned form to determine lifetime constants as shown in S3.

S2. Effect of voltage and temperature on base flipping kinetics

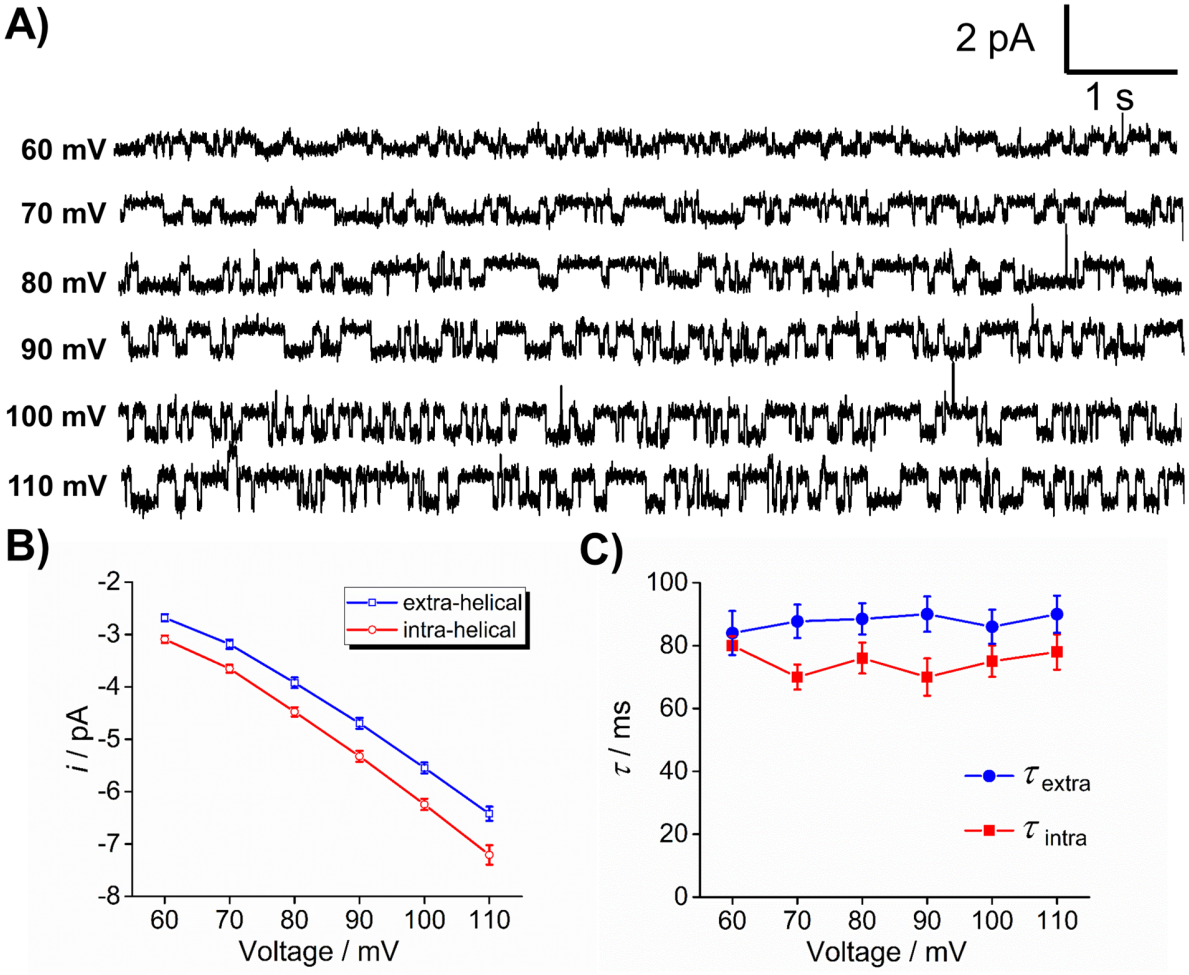

Figure S1. A) Representative *i-t* traces for C:C mismatch at different voltages (*trans* vs *cis*). B) Current levels for intra-helical (red) and extra-helical state (blue). C) Lifetimes of the extra-helical state (τ_{extra}) and intra-helical state (τ_{intra}) as a function of applied voltages across the α-HL at 20 °C (corresponding log-binned dwell time histograms are shown in Figure S2). Solution contained 250 mM KCl and 10 mM PB (pH 6.2). Error bars indicate 95% confidence intervals from fitting.

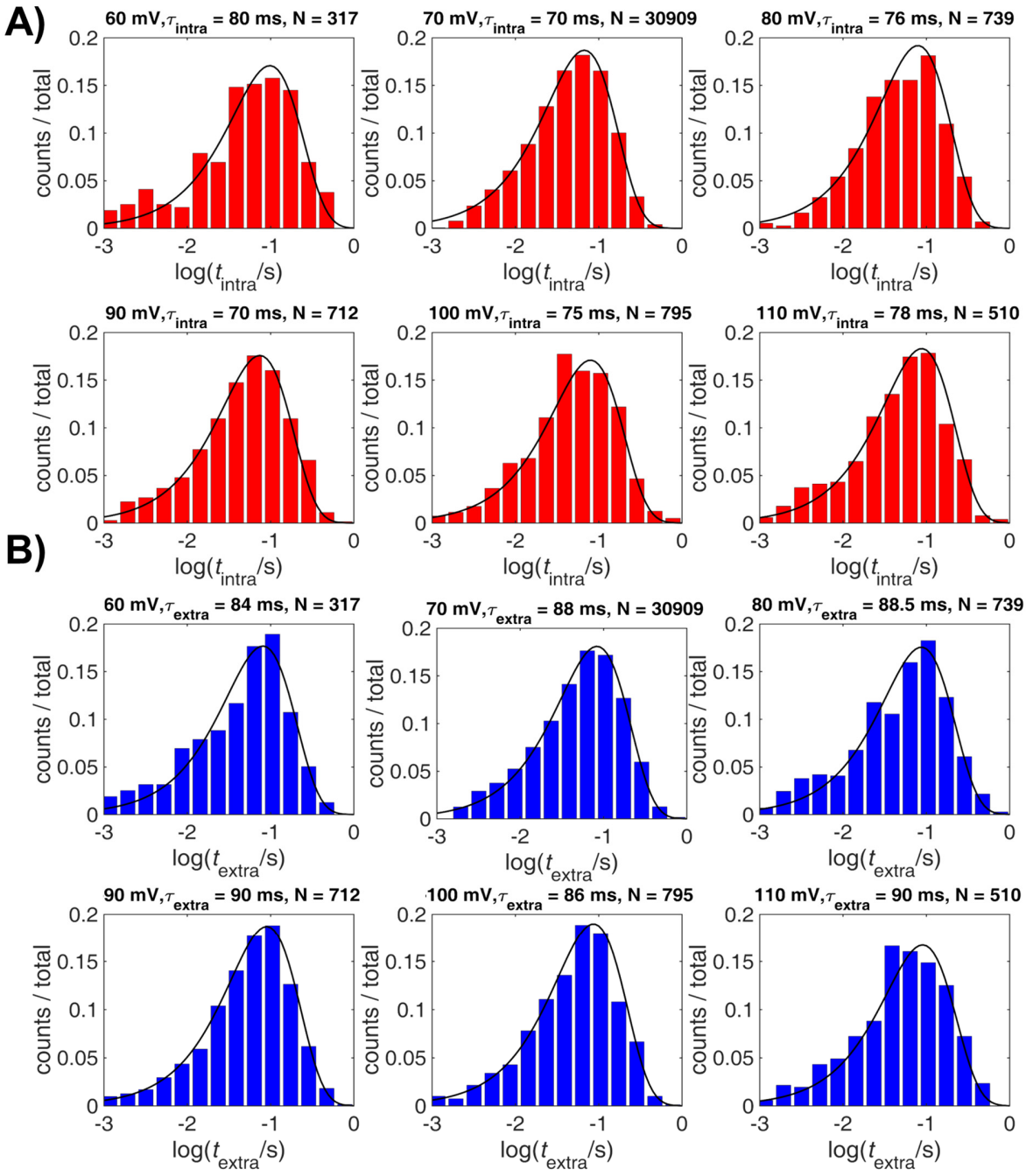

Figure S2. Log-binned dwell time histograms for A) intra-helical and B) extra-helical states of a C:C mismatch at different applied voltages at 20 °C. Solution contained 250 mM KCl and 10 mM PB (pH 6.2).

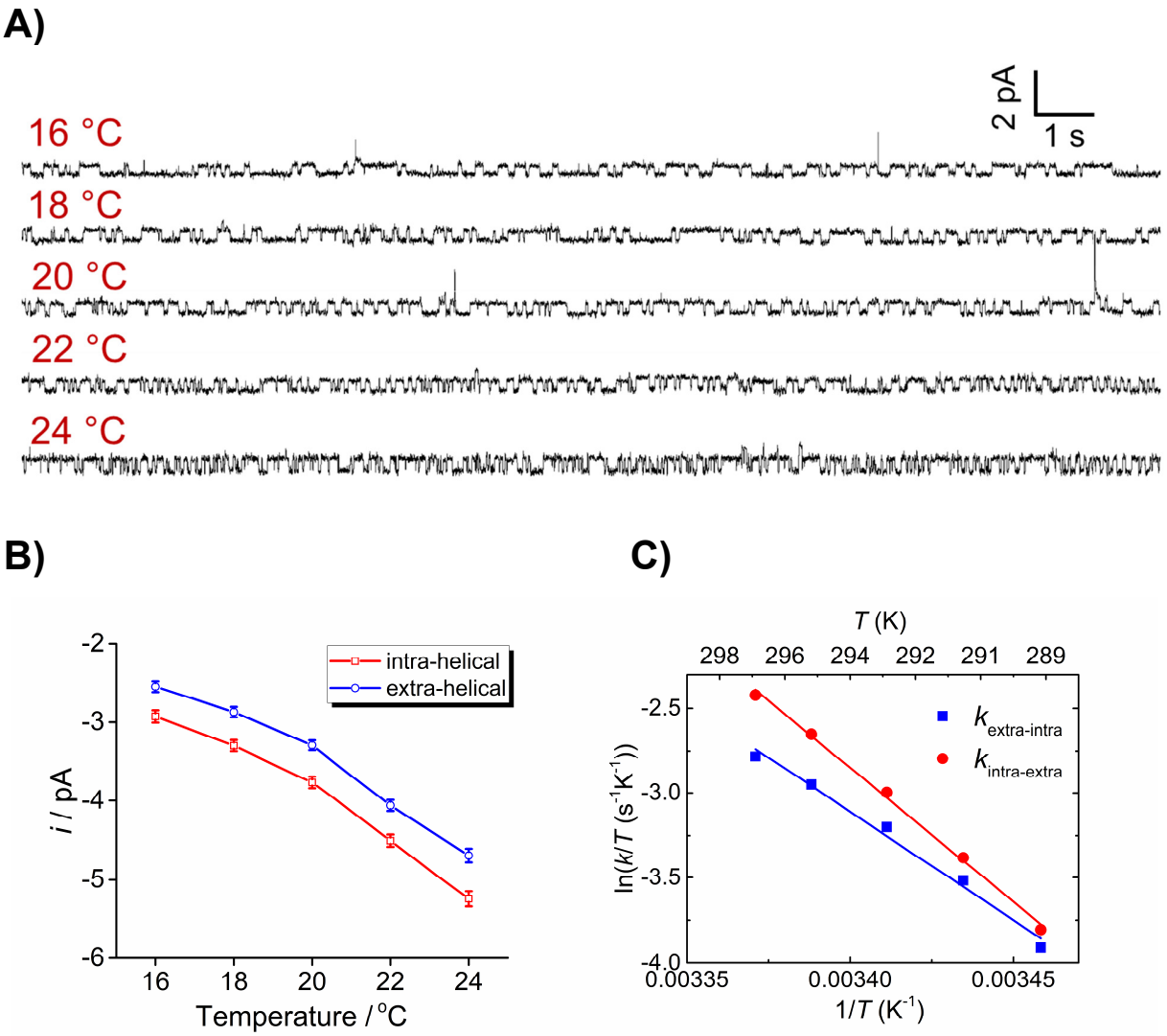

Figure S3. A) *i-t* traces for a dsDNA with a C:C mismatch captured in the α-HL nanopore at different temperatures. B) Current levels for intra- and extra-helical states at different temperatures. C) Eyring plot for the single-molecule kinetics of base flipping-in ($k_{extra\text{-}intra}$) and flipping-out ($k_{intra\text{-}extra}$). The solution contained 250 mM KCl and 10 mM phosphate buffer (pH 6.2). 70 mV was applied across the α-HL channel (*trans* vs *cis*).

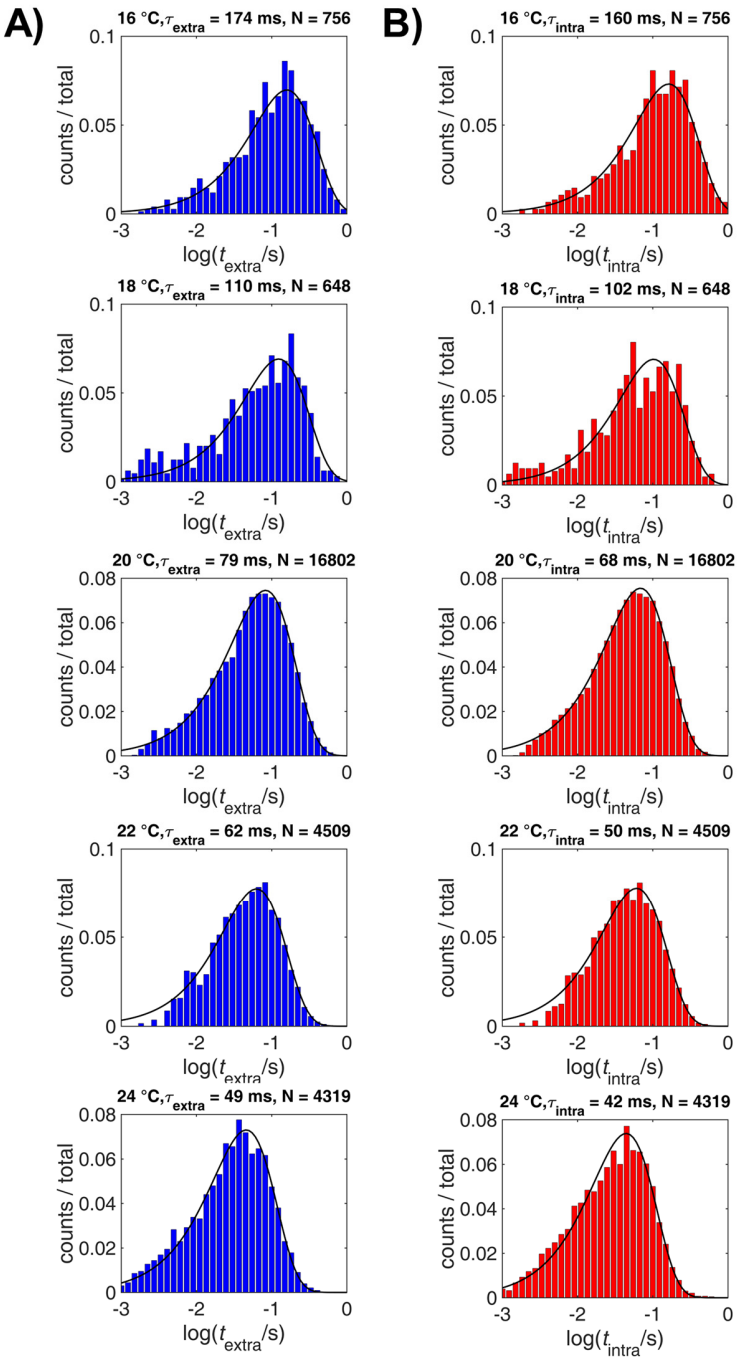

Figure S4. A) Log-binned dwell time histograms for A) intra-helical and B) extra-helical states of a C:C mismatch at different at different temperature. The solution contained 250 mM KCl and 10 mM phosphate (pH 6.2). 70 mV was applied across the α-HL channel (*trans* vs *cis*).

Table S1. $\Delta H^{\ddagger}$, $\Delta S^{\ddagger}$ and $\Delta G^{\ddagger}$ obtained from Eyring plot for base flipping kinetics at pH 6.2.

	extra→intra	intra→extra
$\Delta H^{\ddagger}$(kcal/mol)	25	31
$\Delta S^{\ddagger}$(e.u.)	33	54
$\Delta G^{\ddagger}$(kcal/mol)	16	15

S3. Optimizing the residence time of ds-DNA inside protein channel.

We optimized the voltage applied across the α-HL channel to maximize the time that a single duplex molecule resides in the channel. It can be seen from Figure S5 that by lowering the voltage from 120 mV to 70 mV (*trans* vs *cis*), the unzipping time increases by ~3 orders of magnitude (~ 1 – 10^3 s).

The increased dwell time is attributed to the decreased force on the molecule and therefore decreased probability of unzipping. However, voltages lower than 70 mV result in a decrease of DNA's dwell time because the electric force is not strong enough to prevent the duplex molecule from diffusing out of the α-HL channel.[4] At 70 mV, the capture rate is less than 1 molecule per minute while the average dwell time of the molecule is 3000 s, thus giving confidence that the same molecule is present throughout the length of the experiment when performing pH titrations.

In addition, lifetimes of intra-helical and extra-helical states are not a function of voltage from 70 to 120 mV, as shown in Figure S1. This suggests that the kinetics of base flipping is unaffected by our applied voltage.

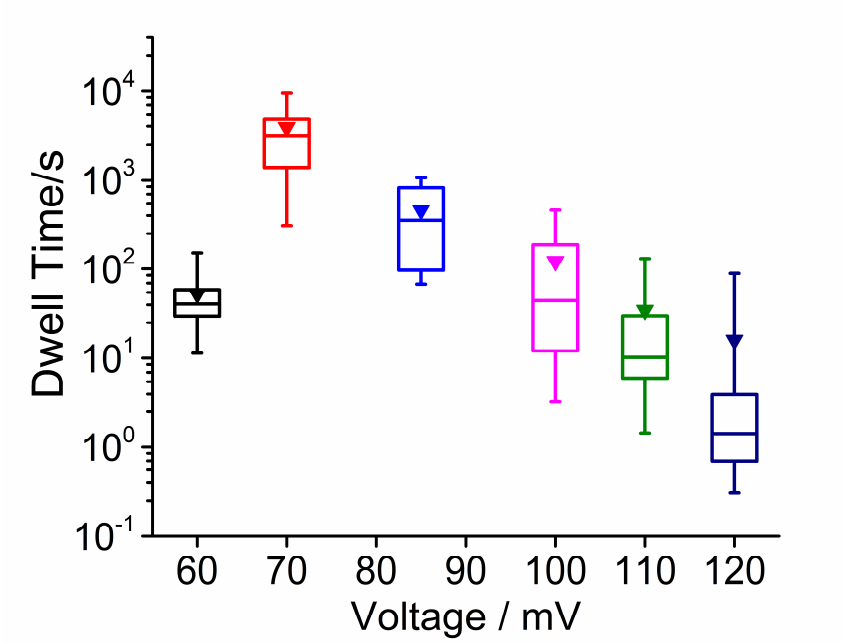

Figure S5. Dwell time of ds-DNA inside the α-HL at different voltages (*trans* vs *cis*) at 20 °C. The box edges are 25% and 75% percentiles. The bars are 5% and 95% percentiles. Horizontal lines inside the box are the medians and triangles are the means. Sample size at each voltage is >25. Solution contains 250 mM KCl and 10 mM PB (pH 6.2). The DNA sequence is shown in Figure 1 of the main text.

S4. Log-binned histogram for determining lifetimes.

If the same current level corresponds to multiple sub-state, the dwell time histograms will exhibit a mixture of multiple exponentials in the form of:

$$f(t) = a_1{\tau_1}^{-1}e^{-t/\tau_1} + a_2{\tau_2}^{-1}e^{-t/\tau_2} + \cdots$$

$$= \sum_i a_i{\tau_i}^{-1}e^{-t/\tau_i} \quad \text{(eq S1)}$$

where $\sum_i a_i = 1$.

Such a distribution can be fitted to obtain characteristic lifetimes, τ_i. However, if the lifetimes are similar, or the weight of one component is small, it can be difficult to determine if the distribution is mono- or multi-exponential, which leads to large fitting errors. In this situation, log-binned histograms have been shown to be advantageous in analyzing and representing sums of multiple exponential distributions.[5] Taking the log transform of t as x = ln(t), the probability distribution of x has been shown to be:

$$f(x) = \sum_i a_i{\tau_i}^{-1}e^{\left(x-\frac{e^x}{\tau_i}\right)} \quad \text{(eq S2)}$$

The shape of f(x) is no longer exponential. Instead, it is a negatively skewed bell-shaped curve with peak positions at $t = \tau_i$.

The advantage of the log-binned histogram is shown in an example. For a distribution of $f(t) = a_1{\tau_1}^{-1}e^{-t/\tau_1} + a_2{\tau_2}^{-1}e^{-t/\tau_2}$, where a_1= 0.9, τ_1= 10 ms, and a_2= 0.1, τ_2= 100 ms, it is very hard to distinguish the two exponential components, as shown in Figure S6A. However, the log-transform plot clearly indicates two peaks located at $\tau_1$and τ_2, as shown in Figure S6B. The log-binned histogram compared to the original histogram results in

less errors in the fitting parameters. Therefore, all the dwell time histogram in this paper will be log-binned.

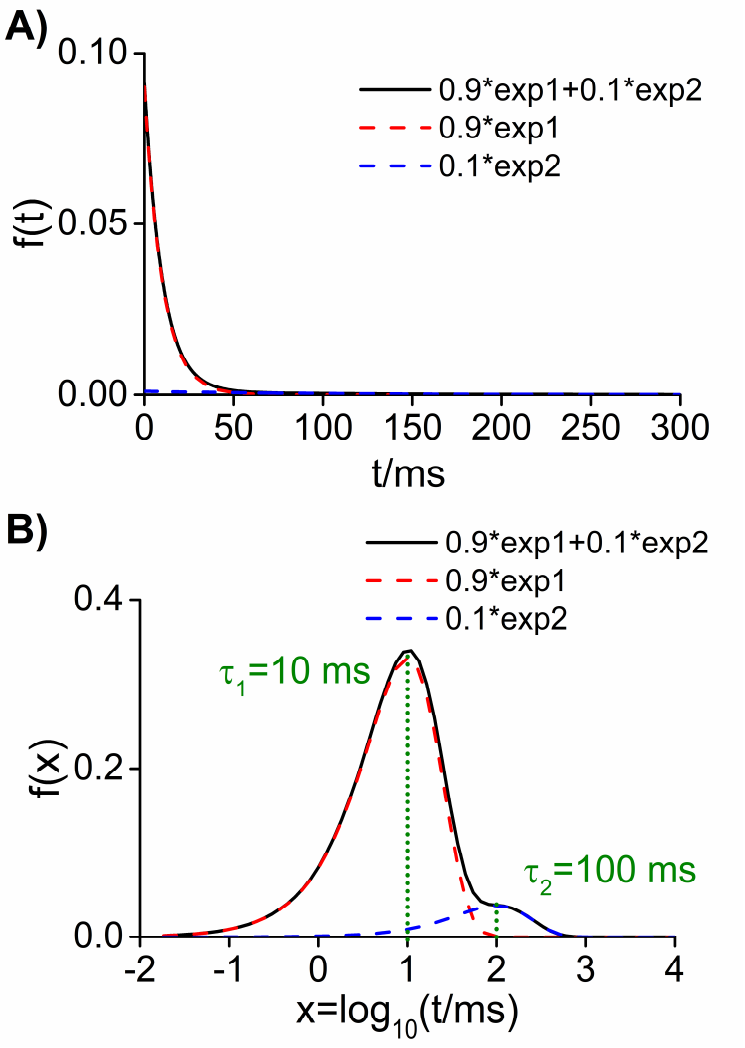

Figure S6. Histograms of two exponential distributions, exp1 (red line), exp2 (blue line), and their sum (black line) with A) conventional binning and B) log binning. Exp1 has the time constant (τ_1) of 10 ms and exp2 has the time constant (τ_2) of 100 ms.

S5. Prolonged measurement of base flipping kinetics inside the protein channel

At different pH, the durations of both intra-helical and extra-helical states can be well fitted with mono-exponential distributions, as shown in Figure S7. This indicates that transition between the intra-helical states and extra-helical states (i.e., base flipping) follows an apparent first-order process with rate constants of $k_{\mathrm{intra-extra}} = 1/\tau_{\mathrm{intral}}$ and $k_{\mathrm{extra-intra}} = 1/\tau_{\mathrm{extra}}$. As the pH decrease from 7.5 to 6.6, τ_{intra} becomes longer and τ_{extra} remains unchanged.

The stability of the base flipping kinetics over extended periods (> 1 h) was evaluated by a moving average of a cluster of dwell times for each state. As shown in Figure S8, the average lifetimes for extra-helical states (τ_{extra}) and intra-helical states (τ_{intra}) of base flipping are stable at all pH over hours within the expected confidence interval (based on the Erlang distribution, which is the distribution of the sum of exponential variables). This indicates that the kinetics of base flipping is stable throughout experimental measurement (Figure S8).

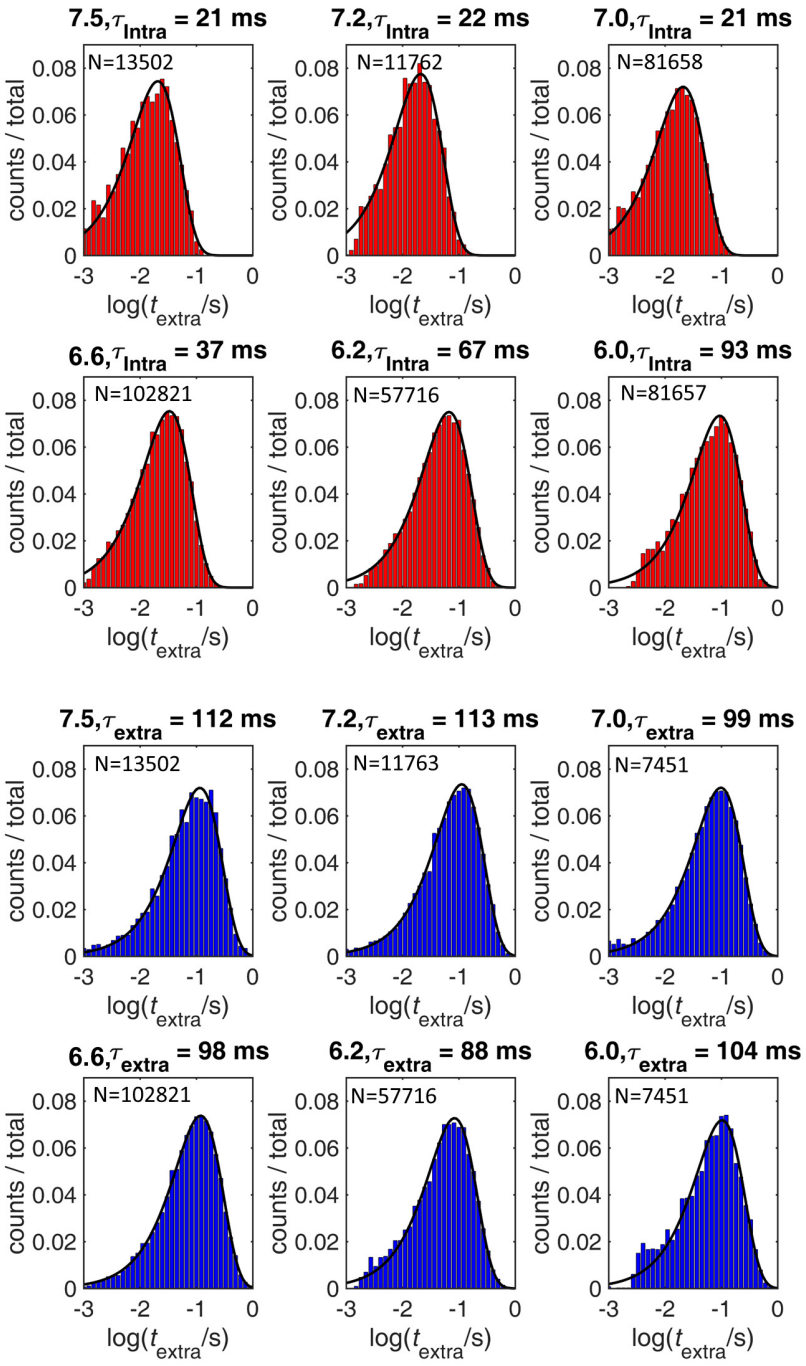

Figure S7. Log-binned histograms of durations of A) intra-helical states and B) extra-helical states at different pH. Black lines are best fits using mono-exponential distributions (apparent first order kinetics).

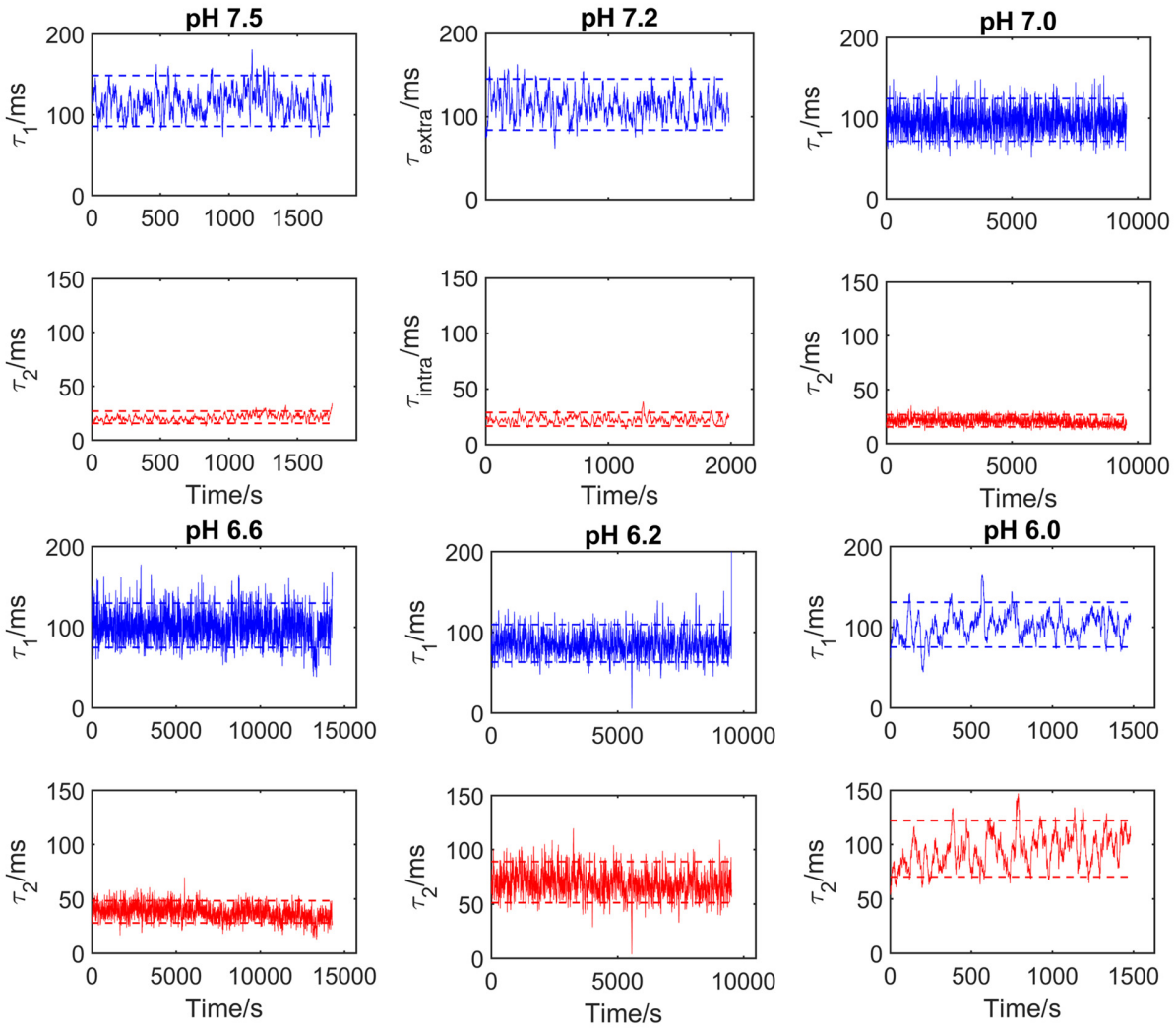

Figure S8. Stability of extra-helical lifetimes (τ_{extra}, blue) and intra-helical lifetimes (τ_{intra}, red) of a single molecule of DNA with C:C mismatch in the protein channel at different pH. Each point is an average of the duration of 100 consecutive events (~15 s). Dashed lines are 99% confidence intervals.

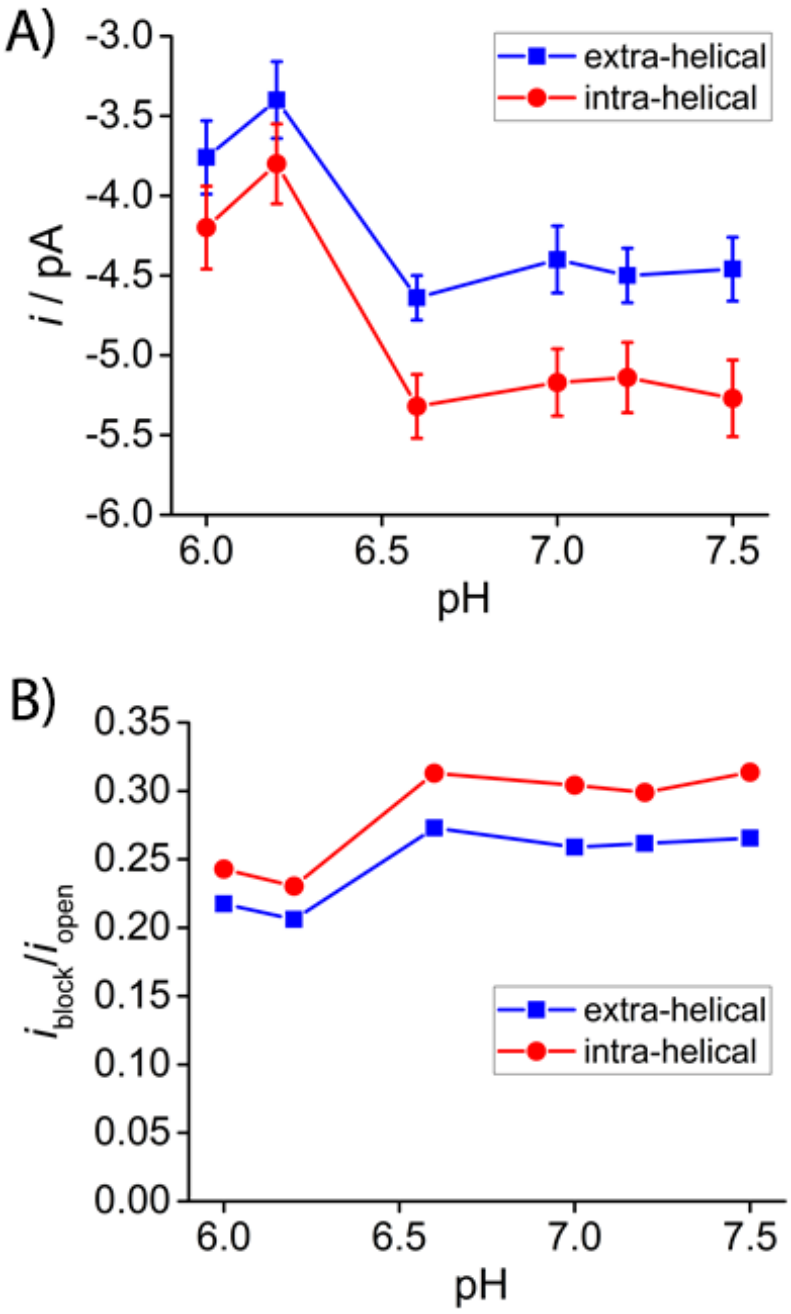

Figure S9. Effect of pH on the current levels of intra-helical and extra-helical states. A) Current amplitudes of intra-helical and extra-helical states. B) Relative blockade (i_{block}/i_{open}) of intra-helical and extra-helical states.

S6. Titration of a single DNA molecule inside the protein channel

We performed single-molecule titration inside the protein channel. The raw *i-t* traces are shown in Figure S10. With the addition of acid (HCl), the duration of intra-helical states (lower states in the *i-t* trace) becomes longer. Such a change can be reversed by the addition of base (KOH). The analysis of the lifetime of each state is shown in Figure 4 of the main text. The time delay between adding acid (or base) and the change in the *i-t* response is due to slow solution mixing in the cell.

Note that during the addition of the acid or base, the opening of Faraday cage and the addition of the acid or base result in ~ 20 s of *i-t* traces showing large noise (indicated by the red arrows in Figures S10A and B), which prevents the observation of if base flip of the DNA molecule. Immediately after the ~20 s noise, the signal from base flipping is highly likely to be from the same molecule for two reasons. First, the probability for the molecule to unzip/escape in 20 s is small given a mean dwell time of ~3000 s at 70 mV. In addition, even if the molecule is lost, capture of a second DNA molecule after the lost in this 20 s noise is unlikely given a capture rate of < 1 molecule per minute at 70 mV. Together, these almost guarantee that after 20 s if *i-t* traces indicate the presence of a DNA molecule inside the protein channel, it is very likely to be the same molecule.

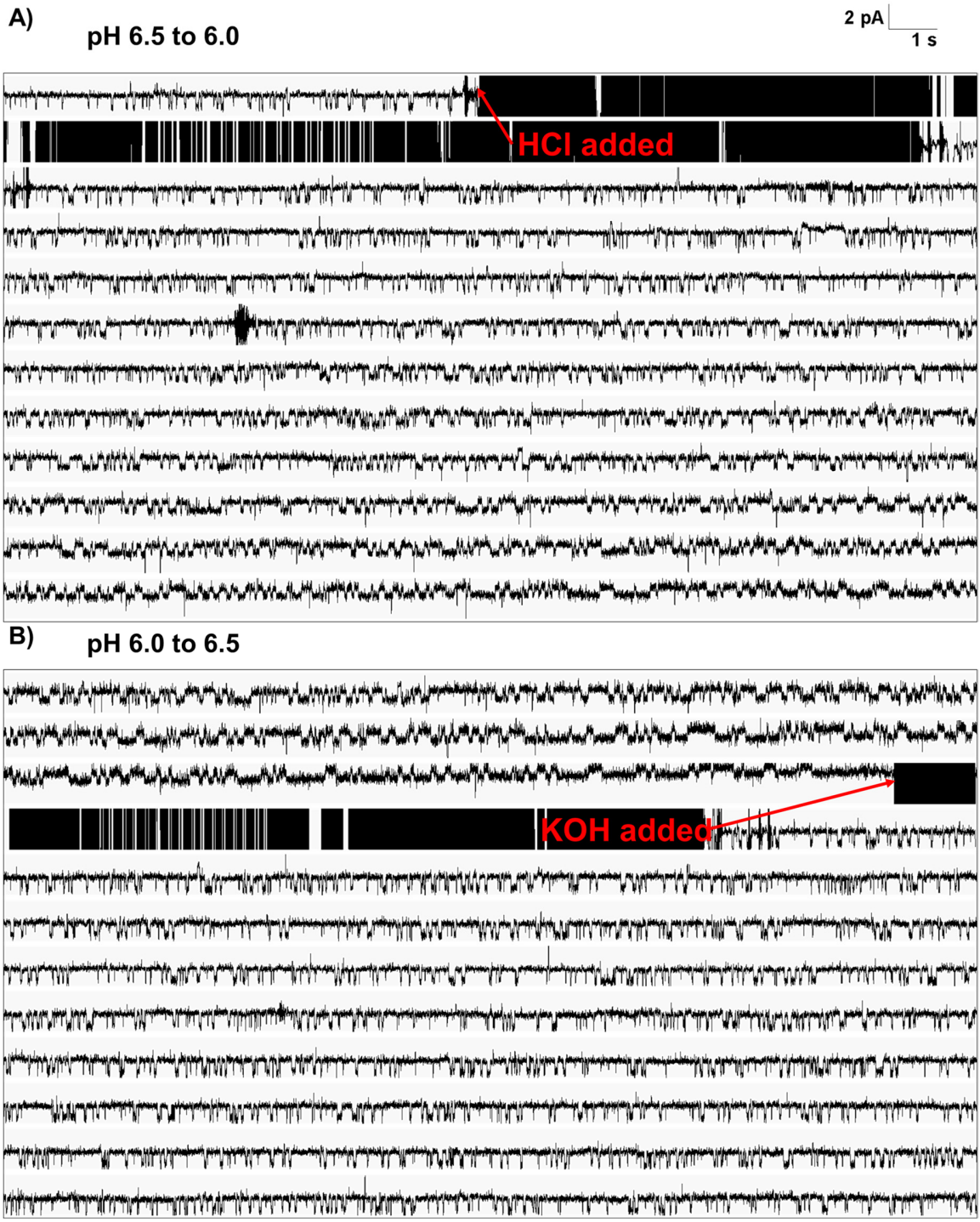

Figure S10. Raw *i-t* trace during the titration of a single captured dsDNA in the α-hemolysin nanopore. Titration A) from pH 6.5 to 6.0 and B) from pH 6.0 to 6.5.

S7. Simulated *i-t* traces using a hidden Markov model

A Monte Carlo method was used to simulate the *i-t* traces based on a proposed hidden Markov kinetic model (Figure S11). Markov models have been widely used to model the kinetics of single-molecule reactions.[6] "Markov" means that the future states only depend on the *current* state, but not on the history (e.g., how long it has stayed in the current state, what paths were taken to arrive at the current state). The 4 states in our model include intra-helical protonated state ($S_{intra\text{-}H+}$), intra-helical deprotonated state (S_{intra}), extra-helical protonated state ($S_{extra\text{-}H+}$), and extra-helical deprotonated state (S_{extra}). In our experiment, when a DNA molecule is captured in the α-HL, only two major current levels are identified instead of four states that are proposed in the model. This means that some states are "hidden" in the measurement, and it is likely that some states are sharing the same current level. In our model, we assume that the two extra-helical states ($S_{extra\text{-}H+}$ and S_{extra}) display one current level i_1, while the two intra-helical states ($S_{intra\text{-}H+}$ and S_{intra}) share another current level i_2, as shown in Figure S11.

In a Monte Carlo simulation, time is discretized into time intervals, δ*t*, and transition between states are simulated using a transition probability defined by our model in Figure S11. When δ*t* is sufficiently small, the probability that a molecule transits from state *i* to state *j* in δ*t* (w_{ij}) is equal to the product of transition rate from *i* to *j* (k_{ij}) and the length of the interval (δ*t*). (i.e., $w_{ij} = k_{ij}\,\delta t$ as $\delta t \rightarrow 0$). Multiple transitions (e.g., *i-j-i-j*) in a short δ*t* are safely ignored. The probability of staying at the same state *i*, denoted as w_{ii}, is equal to one minus the sum of probabilities of transitions from *i* to other states i.e., $w_{ii} = 1 - \sum_{j \neq i} k_{ij}\,\delta t$. Using this method, the probabilities of all possible transitions in δ*t* can be

calculated based on the kinetic parameters in our model (Figure S11), which are summarized in a transition probability matrix in Table S2. The entry in the i^{th} row and j^{th} column denotes the probability of transition from state *i* to state *j*.

Table S2. Transition probability matrix for transitions between the four states (row to column) during δ*t*.

start \ *end*	**S H+ intra**	**S H+ extra**	**S intra**	**S extra**
S H+ intra	1 - $k^{\mathrm{H^+}}_{\mathrm{intra\rightarrow extra}}\delta t$ -$k^{\mathrm{intra}}_{\mathrm{d}}\delta t$	$k^{\mathrm{H^+}}_{\mathrm{intra\rightarrow intra}}\delta t$	$k^{\mathrm{intra}}_{\mathrm{d}}\delta t$	0
S H+ extra	$k^{\mathrm{H^+}}_{\mathrm{extra\rightarrow intra}}\delta t$	1 - $k^{\mathrm{H^+}}_{\mathrm{extra\rightarrow intra}}\delta t$ - $k^{\mathrm{extra}}_{\mathrm{d}}\delta t$	0	$k^{\mathrm{extra}}_{\mathrm{d}}\delta t$
S intra	$k^{\mathrm{intra}}_{\mathrm{p}}[\mathrm{H^+}]\delta t$	0	1 - $k^{\mathrm{intra}}_{\mathrm{p}}[\mathrm{H^+}]\delta t$ - $k_{\mathrm{intra\rightarrow extra}}\delta t$	$k_{\mathrm{intra\rightarrow extra}}\delta t$
S extra	0	$k^{\mathrm{extra}}_{\mathrm{p}}[\mathrm{H^+}]\delta t$	$k_{\mathrm{extra\rightarrow intra}}\delta t$	1 - $k^{\mathrm{extra}}_{\mathrm{p}}[\mathrm{H^+}]\delta t$ - $k_{\mathrm{extra\rightarrow intra}}\delta t$

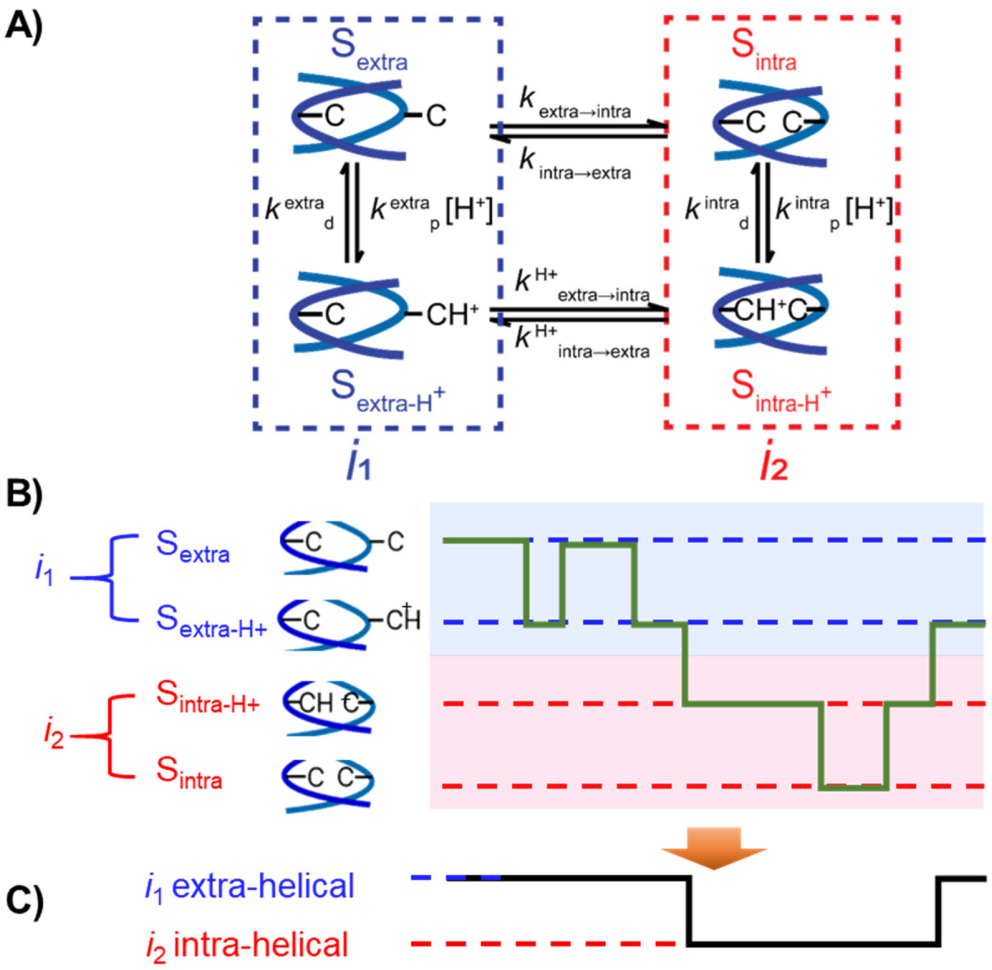

Figure S11. A) A 4-state hidden Markov kinetic model for protonation/deprotonation and base flipping of C:C mismatch in a DNA molecule. The states in blue and red boxes are intra-helical and extra-helical states, respectively. B) Hidden states vs time generated from the kinetic model in A) using Monte Carlo method. C) Observable states vs time (simulated *i-t* traces) derived from the hidden states vs time in B).

Our simulation starts randomly from one of the four states in Figure S11A. The time interval in the simulation (i.e., δt) is 0.1 ms, which is comparable to the limit of the

experimental measurement (10 kHz sampling frequency) and is 2 orders of magnitude smaller than the smallest lifetimes experimentally observed (τ_{intra} = 16 ms). During each δt, a random number uniformly distributed between 0 and 1 is generated and compared with the transition probability (as shown in Table S2) to decide which transition should occur. This process is repeated in the following δt intervals, which generate the sequence of states as a function of time as shown in Figure S11B. Note that from the simulation, we can monitor exactly the transitions between all four states (the hidden states). Since some states share the same experimental signal, e.g., both S_{extra} and $S_{extra\text{-}H+}$ have the same current level, we can covert the hidden states vs time to observable states (current) vs time by grouping the hidden-states according to their current level, as illustrated in Figure S11C. The simulation continues until the total simulation time is reached (1200 s for our simulation). The simulations are conducted at different pH values. Note that in our simulation, we also reflect the missing events caused by the filter (low pass filtered at 1 kHz for data analysis) by concatenating events separated by one short event (<0.5 ms).

To obtain the best kinetic parameters based on the experimental results, we perform a global optimization by minimizing the mean squared error (MSE) of the dwell time histograms of both intra-helical and extra-helical states between simulations and experiments at different pH (7.5, 7.2, 7.0, 6.6, 6.2, and 6.0) using a particle swarm optimization package in Matlab.

The fastest protonation rate is the transport-limited rate, which can be calculated from the ionic current (i_{total}, ~5 pA) and transference number of proton in the bulk solution (t_{H+}). Briefly, the current carried by proton (i_{H+}) can be estimated from $i_{H+} = i_{total} \times t_{H+}$. The rate of H^+ translocation ($k_{mt\ H+}$) is obtained from $k_{mt\ H+} = i_{H+}/q_{H+}$, where q_{H+} is the charge of one

proton (equals 1.6×10^{-19} C). $k_{mt\ H+}$ at different pH using this calculation is summarized in Table S3. Note that in this calculation we assume perm-selectivity of cation, which has been shown to be true under our electrolyte condition (ionic strength < 1 M) when DNA resides in the protein channel.[7] At pH 7.0, the mass transport limited rate for protonation is 56 s^{-1}, which corresponds to an apparent protonation rate of 5.6×10^{8} M^{-1} s^{-1}.

Table S3. Rate of H^+ translocation ($k_{mt\ H+}$) and the corresponding transport-limited protonation rate at different pH.

pH	$k_{mt\ H+}$ (s^{-1})	Transport-limited protonation rate (M^{-1} s^{-1})
7.5	18	5.6×10^{8}
7.0	56	5.6×10^{8}
6.5	180	5.6×10^{8}
6.0	560	5.6×10^{8}

The optimized kinetic parameters are listed in Table 1 in the main text. The simulated *i-t* traces with these optimized parameters are shown in Figure S12. The comparison of the log dwell time histograms of intra-helical and extra-helical states and the corresondping lifetimes from experiment and simulation are shown in Figures S13 and S14.

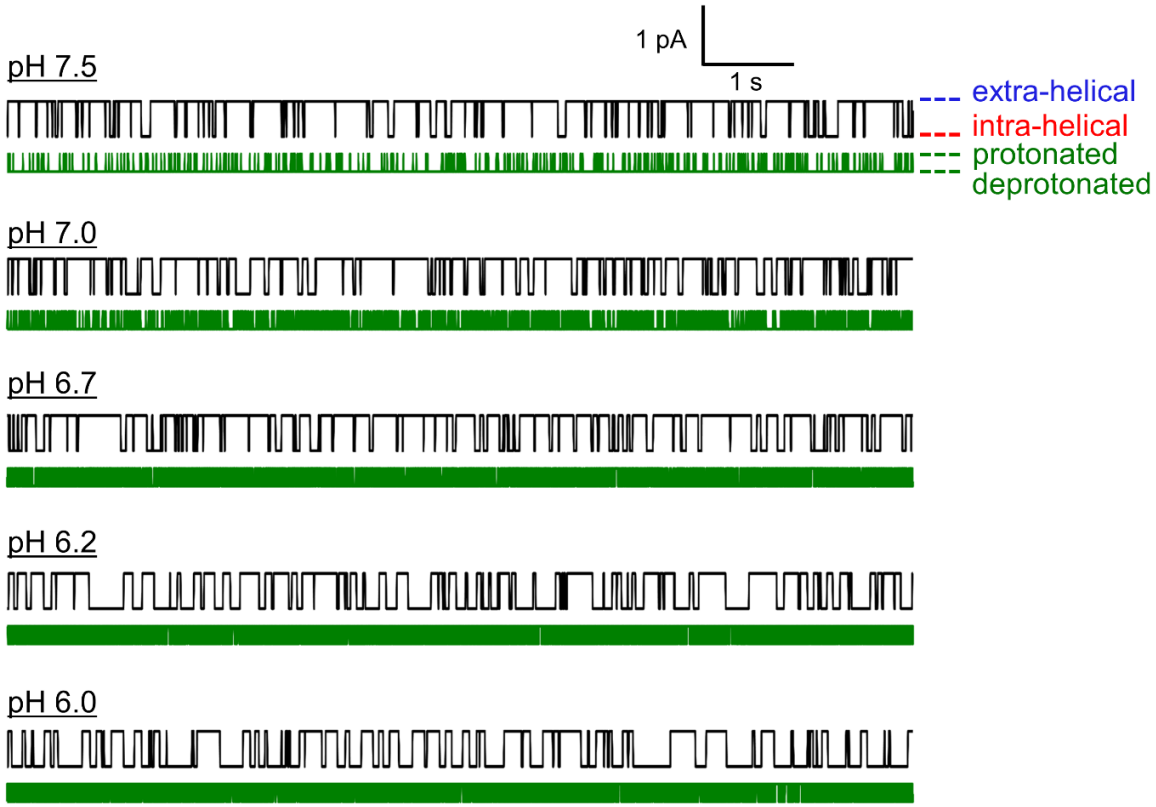

Figure S12. Simulated current states (black) and protonation states (green) vs time for the base flipping of C:C mismatch at different pH using a 4-state kinetic model in Figure S11A and optimized parameters in Table 1 of the main text. At each pH, 20 s of the simulated traces are displayed. The corresponding histograms of the durations of extra-helical and intra-helical states are shown in Figure S13.

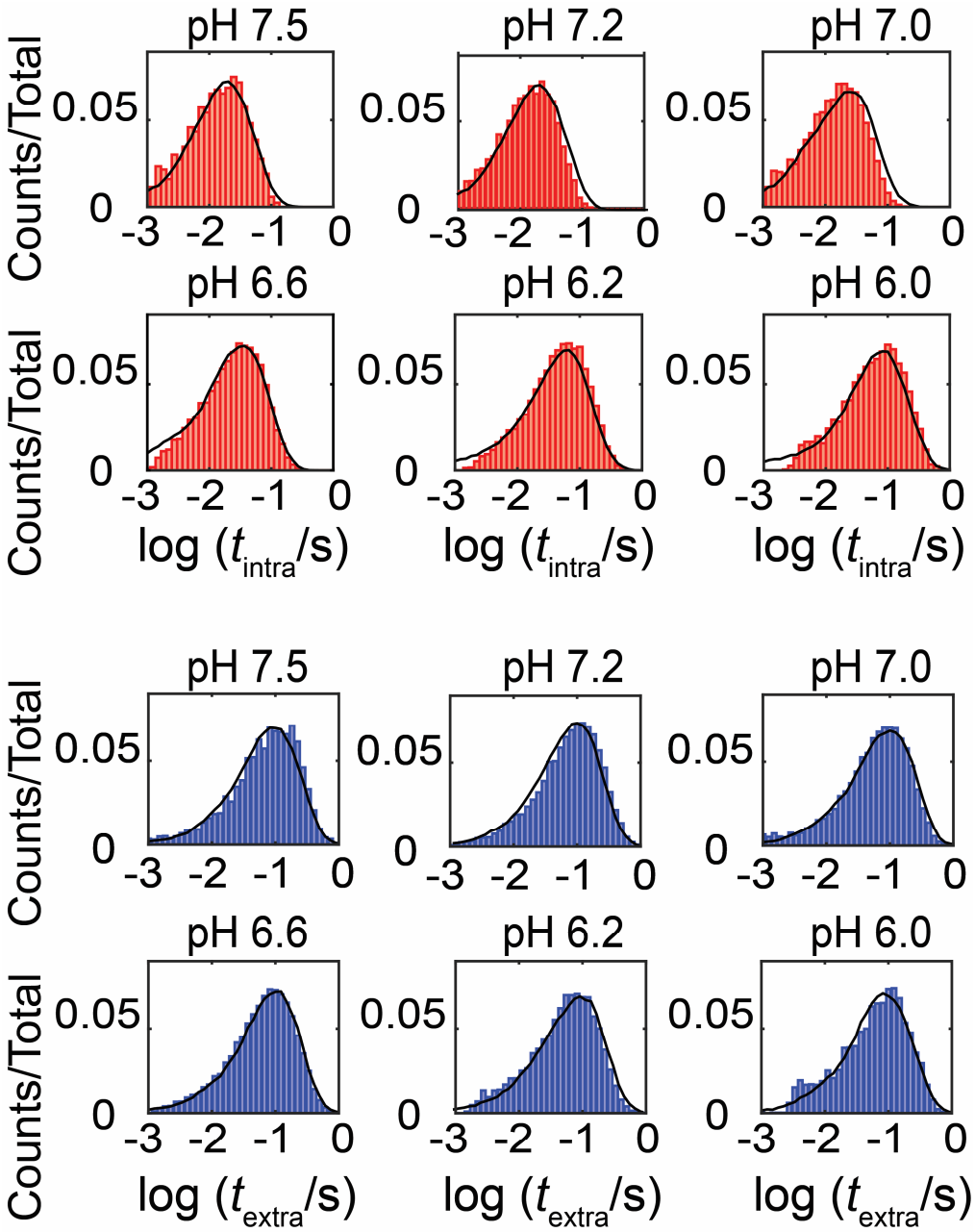

Figure S13. Comparison of dwell time distributions between experiment (bars) and simulation (black lines) for intra-helical (top) and extra-helical (bottom) states at different pH. Simulations contain 1200 s of data. The comparison of the lifetimes between experiment and simulation is shown in Figure S14.

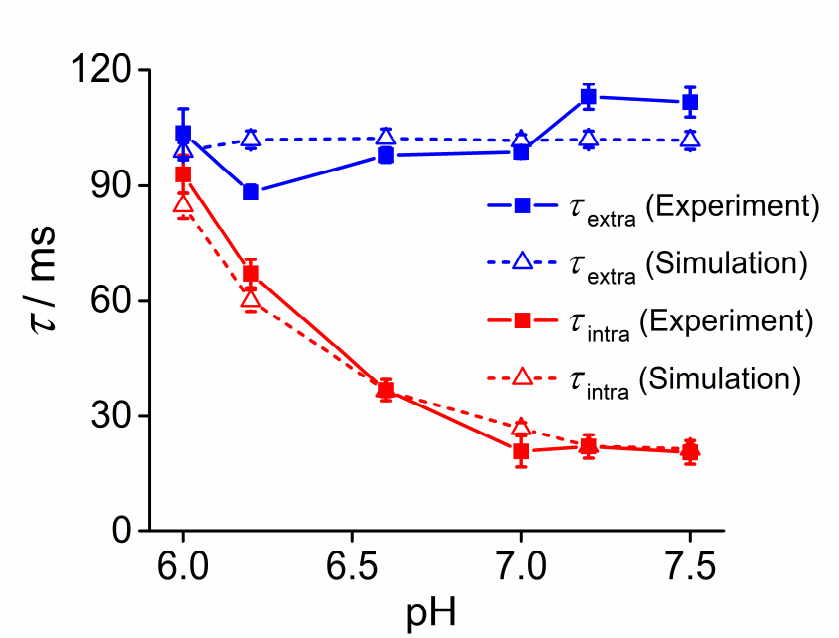

Figure S14. Comparison of Intra-helical (red) and extra-helical (blue) lifetimes of C:C mismatch as a function of pH obtained between experiment (solid) and simulation (dashed). The solution contained 0.25 M KCl and 10 mM phosphate.

S8. Simulation using a three-state model

As discussed in the main text (also seen Table 1 in the main text), because base flipping out from the protonated states is slow ($k_{\mathrm{intra}\rightarrow\mathrm{extra}}^{\mathrm{H}^+}$< 0.1 s^{-1}) and the pK_a for the extra-helical C:C is < 5.6, the extra-helical protonated state rarely occurs in the pH range we studied (pH from 6.0 to 7.5). Therefore, the 4-state model (Figure S11A) can be effectively reduced to a 3-state model as shown in Figure S15. In this way, the fitting parameters are greatly reduced to only 4. Note that the deprotonation rate at intra-helical state ($k_{\mathrm{d}}^{\mathrm{intra}}$) is determined by $\mathrm{p}K_{\mathrm{a}}^{\mathrm{intra}}$ and protonation rate ($k_{\mathrm{p}}^{\mathrm{intra}}$). The optimized parameters for the 3-state model based on our experimental results are shown in Table S4. The optimized parameters in the 3-state model agrees with those in the 4-state model.

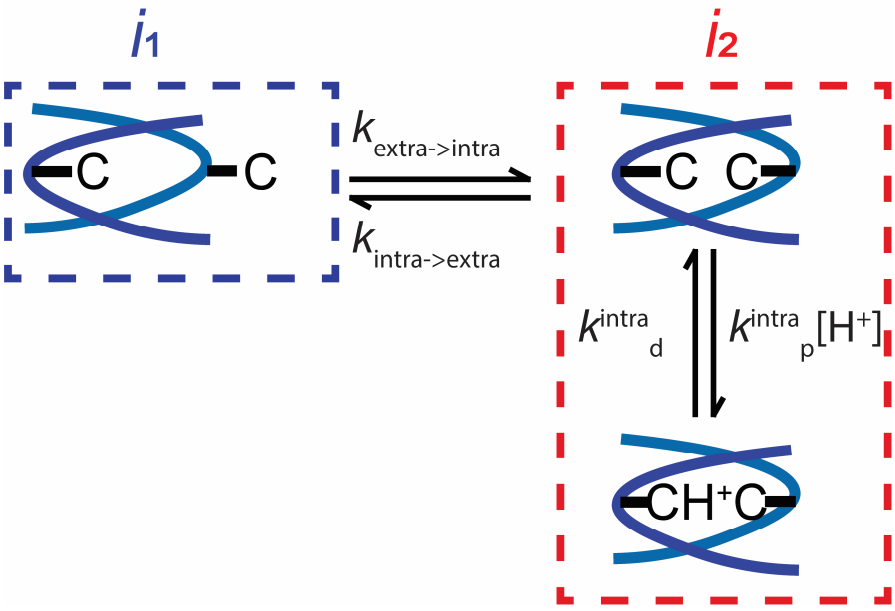

Figure S15. A 3-state hidden Markov model for the pH-dependent base flipping.

Table S4. Comparison of the optimized parameters for pH-dependent base flipping between the 3-state model in Figure S15 and the 4-state model in Figure S10A.

Parameters	3-state model	4-state model
$k_{\mathrm{intra}\rightarrow\mathrm{extra}}$	65 s^{-1}	63 s^{-1}
$k_{\mathrm{extra}\rightarrow\mathrm{intra}}$	10 s^{-1}	10 s^{-1}
$\mathrm{p}K_{\mathrm{a}}^{\mathrm{intra}}$	6.7	6.7
$k_{\mathrm{p}}^{\mathrm{intra}}$	> 2 ×10^{8} M^{-1} s^{-1}	> 2 ×10^{8} M^{-1} s^{-1}

S9. Single-molecule kinetic isotope effect on base flipping

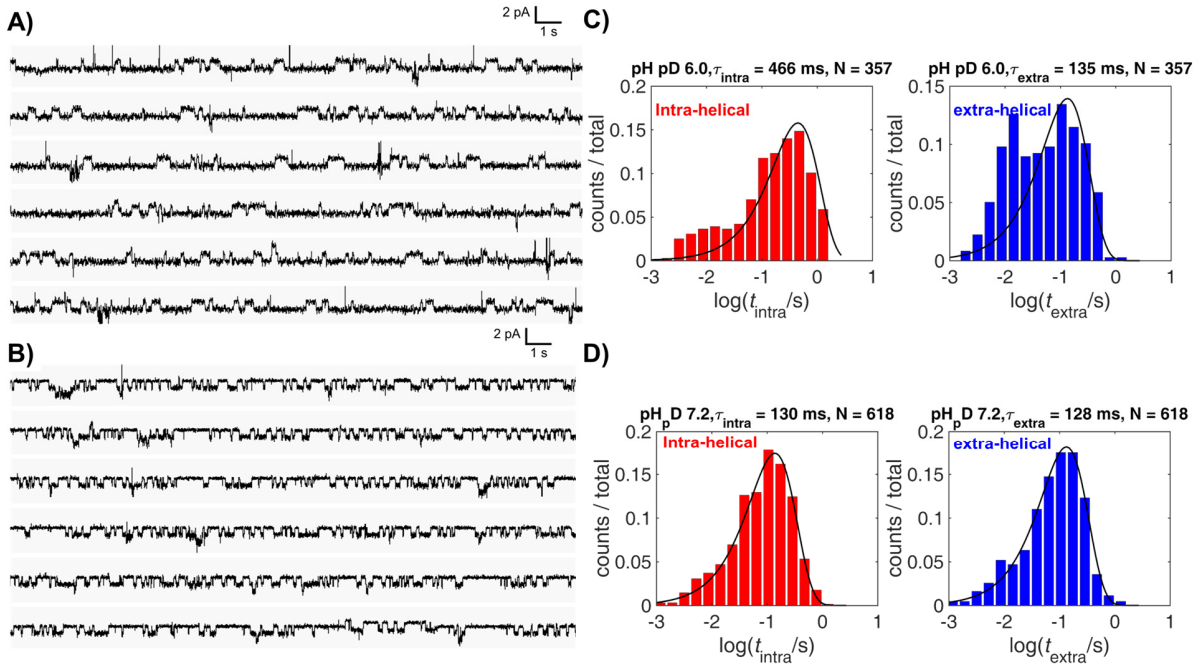

Figure S16. *i-t* traces for single molecules of DNA with C:C mismatch in α-HL at different pH in D_2O with A) pD 7.2 and B) pD 6.0. The corresponding dwell time histograms for intra-helical (red) and extra-helical (blue) states are shown in C) and D). The solutions contain 250 mM KCl and 10 mM phosphate buffer. Voltage of at 70 mV (*trans* vs *cis*) was applied.

Table S5. Kinetic isotope effect of the observed rate for base flipping out (k_{obs}^{out}) and flipping in (k_{obs}^{in}).*

	pH or pD 6.0		pH or pD 7.2	
	k_{obs}^{out} (s^{-1})	k_{obs}^{in} (s^{-1})	k_{obs}^{out} (s^{-1})	k_{obs}^{in} (s^{-1})
H_2O	11.8±0.5	10.1±0.3	47±2	9.9±0.2
D_2O	2.1±0.3	7.7±0.5	9±1	7.8±0.8
KIE	5.2±0.8	1.3±0.2	5.0±0.6	1.3±0.1

*Solution contains 0.25 M KCl, 10 mM phosphate buffer in H_2O or D_2O.

S10. Effect of protonation rate on intra-helical duration distributions

To obtain a qualitative understanding of the effect of protonation rate (k_p) on the observed *i-t* traces and the distribution of durations of each state, we simulated the *i-t* traces at different k_p (keeping the other parameters unchanged as in Table 1 in the main text). The simulated results at different pH are shown in Figures S17-S19.

When the k_p is ~50% or less than the transport-limited rate (Table S3), we observe a deviation from the perfect mono-exponential distribution for the intra-helical states (especially at lower pH, see Figures S17-S19). Indeed, a second peak in the log-binned distribution of intra-helical durations appears at lower k_p, indicating the intra-helical states can be described apparently by two first order processes with different time constants (sections C in Figures S17-S19). Therefore, k_p is at least 2×10^8 M^{-1} s^{-1} for the protonation at intra-helical state. Note that the variation in k_p shows no effect on the durations extra-helical states (sections B in Figures S17-S19).

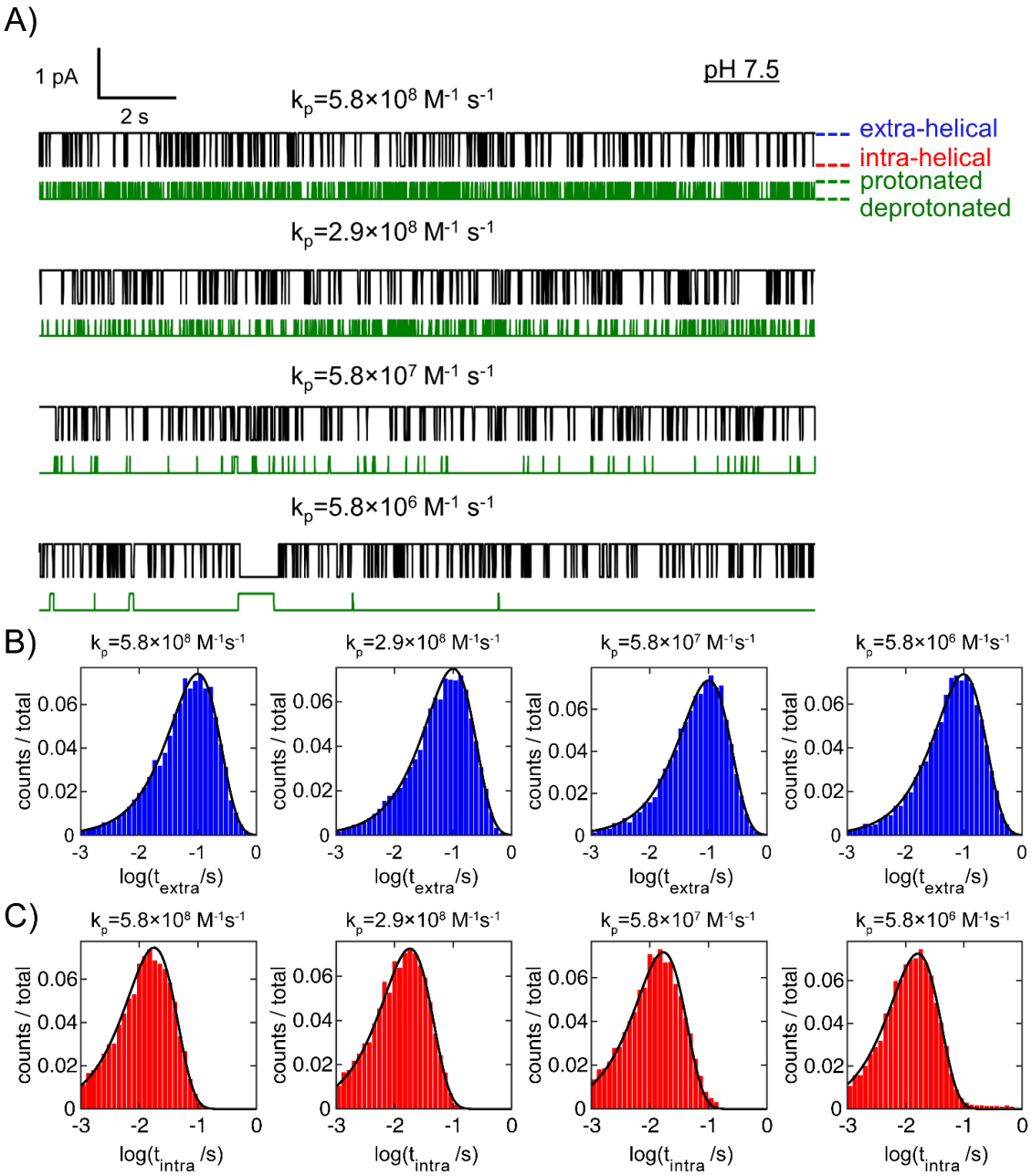

Figure S17. Effect of protonation rate (k_p) on the *i-t* traces and dwell time histogram at pH 7.5. A) Simulated current (black) and protonation states (green) vs time (10 s displayed) at different k_p. Log dwell time histogram of B) extra-helical and C) intra-helical states at different k_p. Black lines are the best fits of mono-exponential distribution.

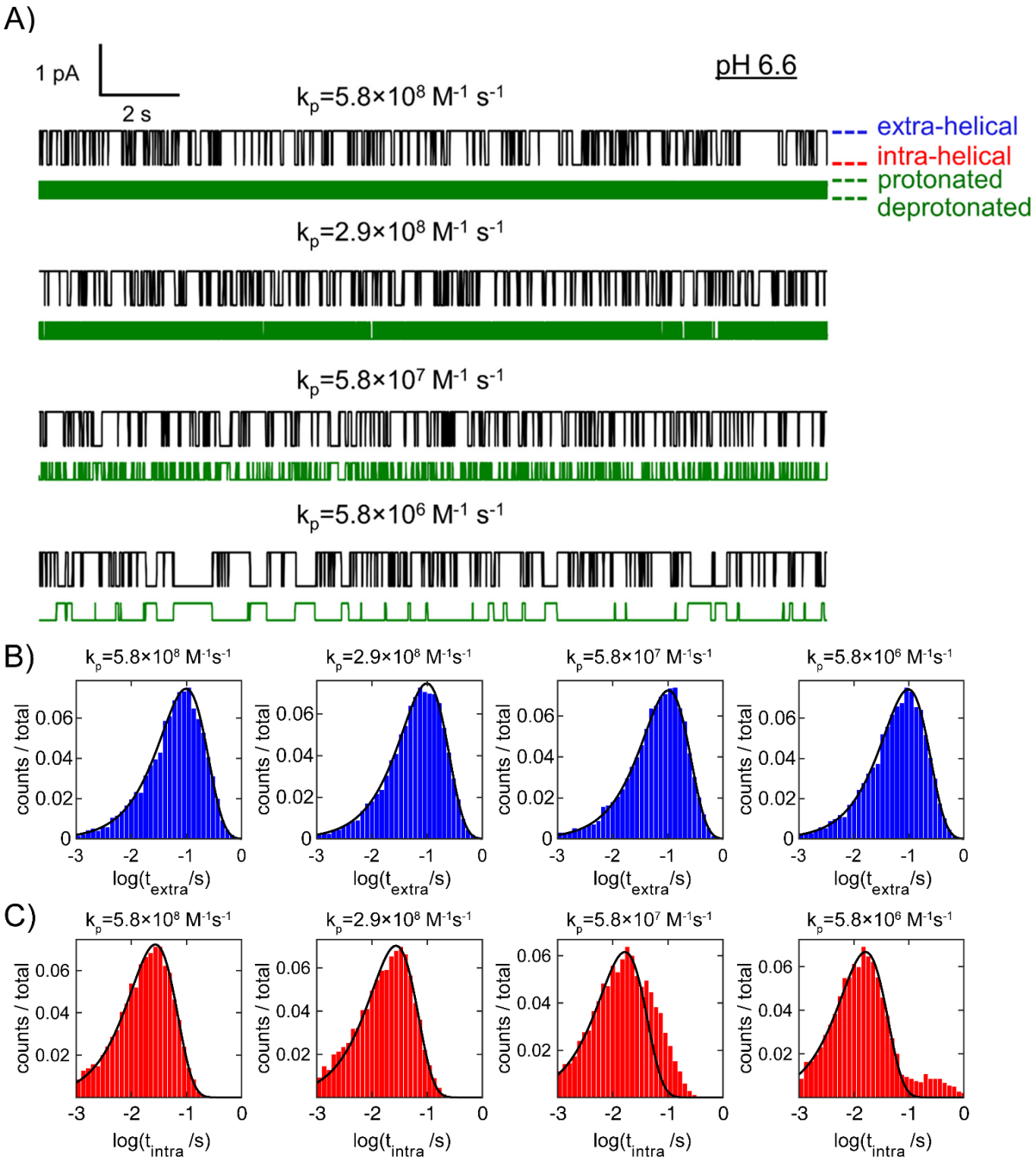

Figure S18. Effect of protonation rate (k_p) on the *i-t* traces and dwell time histogram at pH 6.6. A) Simulated current (black) and protonation states (green) vs time (10 s displayed) at different k_p. Log dwell time histogram of B) extra-helical and C) intra-helical states at different k_p. Black lines are the best fits of mono-exponential distribution.

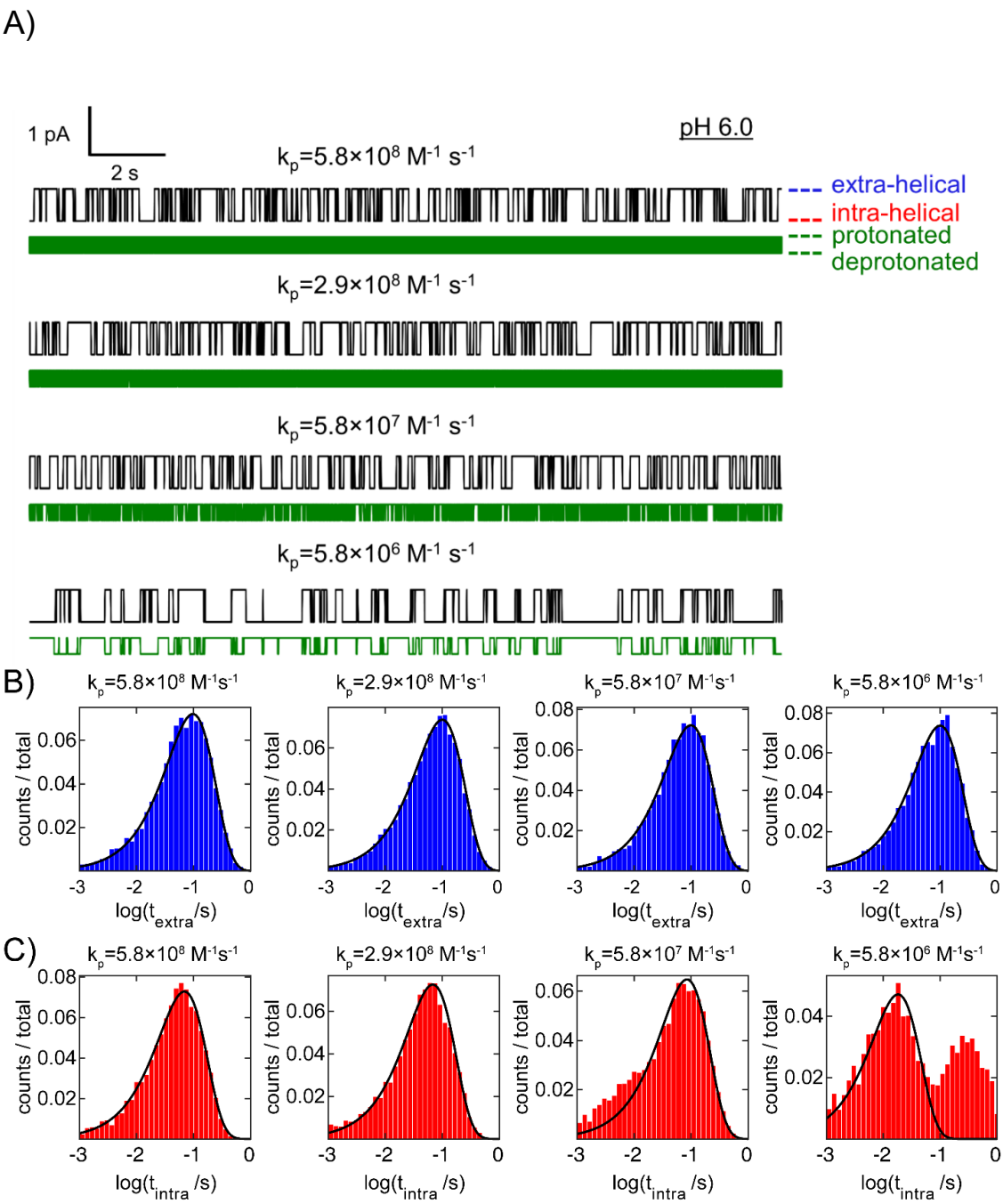

Figure S19. Effect of protonation rate (k_p) on the *i-t* traces and dwell time histogram at pH 6.0. A) Simulated current (black) and protonation states (green) vs time (10 s displayed) at different k_p. Log dwell time histogram of B) extra-helical and C) intra-helical states at different k_p. Black lines are the best fits of mono-exponential distribution.

S11. Simulation disallowing acid-base reaction at intra-helical states

We further simulated the scenario where no protonation/deprotonation is allowed at intra-helical state. The other parameters are the same as those in Table 1 in the main text. Again, we observed the appearance of another peak in the dwell time histogram, indicating the appearance of a process with another time constant. In the *i-t* traces, excessively long intra-helical states are observed, which correspond to the protonated intra-helical states (indicated by stars in Figure S20B). Such results disagree with our experimental data, where only one peak in the dwell time histogram is observed (i.e., one characteristic lifetime is obtained, see Figure S7) and no excessively long intra-helical state in the *i-t* traces is observed (Figure 3A in the main text).

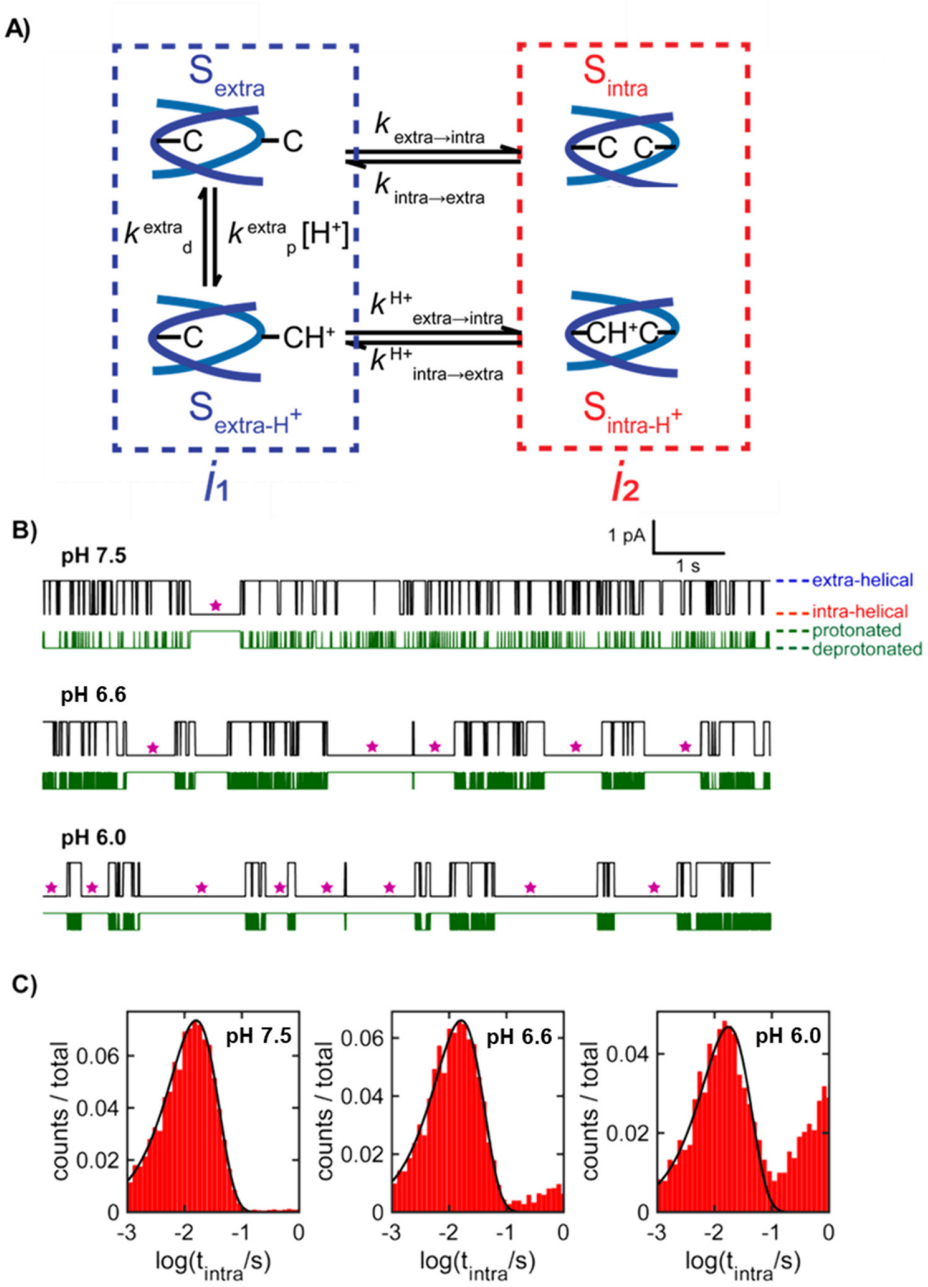

Figure S20. Simulation with no protonation/deprotonation allowed at intra-helical states. A) The corresponding 4-state model. B) Simulated current (black) and protonation state (green) vs time at different pH. The intra-helical protonated states are labeled with stars. C) Dwell time histogram of intra-helical states at different pH.

S12. References

S1. Zhang, B.; Galusha, J.; Shiozawa, P. G.; Wang, G.; Bergren, A. J.; Jones, R. M.; White, R. J.; Ervin, E. N.; Cauley, C. C.; White, H. S., *Anal. Chem.* **2007,** *79* (13), 4778-4787.

S2. Krężel, A.; Bal, W., *J. Inorg. Biochem.* **2004,** *98* (1), 161-166.

S3. Milescu, L. S.; Akk, G.; Sachs, F., *Biophys. J.* **2005,** *88* (4), 2494-2515.

S4. Ervin, E. N.; Kawano, R.; White, R. J.; White, H. S., *Anal. Chem.* **2008,** *80* (6), 2069-2076.

S5. Colquhoun, D.; Sigworth, F., Fitting and Statistical Analysis of Single-Channel Records. In *Single-Channel Recording*; Sakmann, B., Neher, E., Eds.; Springer: New York, 1995; pp 483-587.

S6. McKinney, S. A.; Joo, C.; Ha, T., *Biophys. J.* **2006,** *91* (5), 1941-1951.

S7. Bonthuis, D.; Zhang, J.; Hornblower, B.; Mathé, J.; Shklovskii, B. I.; Meller, A., *Phys. Rev. Lett.* **2006,** *97* (12), 128104.