JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Article

pubs.acs.org/JACS

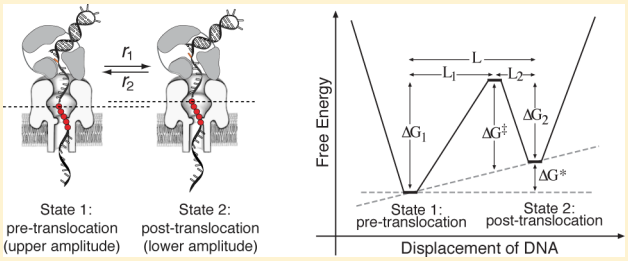
Dynamics of the Translocation Step Measured in Individual DNA Polymerase Complexes

Kate R. Lieberman,* Joseph M. Dahl, Ai H. Mai, Mark Akeson, and Hongyun Wang*

Biomolecular Engineering, University of California, Santa Cruz, California 95064, United States

Supporting Information

ABSTRACT: Complexes formed between the bacteriophage phi29 DNA polymerase (DNAP) and DNA fluctuate between the pre-translocation and post-translocation states on the millisecond time scale. These fluctuations can be directly observed with single-nucleotide precision in real-time ionic current traces when individual complexes are captured atop the α-hemolysin nanopore in an applied electric field. We recently quantified the equilibrium across the translocation step as a function of applied force (voltage), active-site proximal DNA sequences, and the binding of complementary dNTP. To gain insight into the mechanism of this step in the DNAP catalytic cycle, in this study, we have examined the stochastic dynamics of the translocation step. The survival probability of complexes in each of the two states decayed at a single exponential rate, indicating that the observed fluctuations are between two discrete states. We used a robust mathematical formulation based on the autocorrelation function to extract the forward and reverse rates of the transitions between the pre-translocation state and the post-translocation state from ionic current traces of captured phi29 DNAP–DNA binary complexes. We evaluated each transition rate as a function of applied voltage to examine the energy landscape of the phi29 DNAP translocation step. The analysis reveals that active-site proximal DNA sequences influence the depth of the pre-translocation and post-translocation state energy wells and affect the location of the transition state along the direction of the translocation.



INTRODUCTION

DNA polymerases (DNAPs) are molecular motors that catalyze template-directed DNA replication. Movement by replicative DNA polymerases along a DNA substrate during DNA synthesis in steps of one nucleotide is essential for genome integrity. Despite the importance of this translocation step, its kinetics and energetics and their integration in the nucleotide addition cycle during replication are not well understood.

Insight into the structural mechanisms underlying the translocation step has been inferred from crystal structures of complexes formed with primer-template DNA substrates.[1−4] The architecture of the polymerase domain of replicative DNAPs resembles a partially closed right-hand, comprising palm, thumb, and fingers subdomains.[1,3,5,6] In DNAP–DNA complexes containing dNTP complementary to the templating base, elements of the fingers subdomain move relative to their position in complexes lacking dNTP, closing in toward the active site cleft to achieve a tight steric fit with the nascent base pair.

DNAP–DNA binary complexes in the fingers-open conformation and DNAP–DNA–dNTP ternary complexes in the fingers-closed conformation have been considered structural models for the post-translocation and pre-translocation states, respectively.[1,3,4] In this view, the nascent base pair in the closed ternary complex is considered to occupy the site occupied by the terminal base pair of the duplex in a pre-translocation state complex. Crystal structures of A and B family DNAPs reveal steric constraints that, when the polymerase is in the fingers-open conformation, preclude the pre-translocation state binding of DNA. The structures imply that fingers opening accompanies the pre-translocation to post-translocation state transition, and hence fingers closing would accompany the reverse, post-translocation to pre-translocation transition. Recent studies have uncovered intermediates in the fingers-closing transition that are important for nucleotide selection, and that may also be important in the structural mechanism of the translocation step.[2,7−9]

Direct examination of the DNAP translocation step to extract kinetic and energetic parameters is difficult using ensemble methods, and the study of this step in single-molecule experiments requires the challenging combination of high temporal and spatial resolution. We recently reported that the translocation step for the DNAP from bacteriophage phi29 can be directly observed in real time at the single molecule level using the α-hemolysin (α-HL) nanopore, with submillisecond temporal resolution and single-nucleotide spatial precision.[10] The phi29 DNAP is a B family polymerase that catalyzes highly processive DNA replication without the need for accessory proteins such as sliding clamps or helicases.[11] Processive DNA synthesis catalyzed by this enzyme can be monitored with single nucleotide resolution using the nanopore.[12]

Received: September 11, 2012
Published: October 26, 2012

   dx.doi.org/10.1021/ja3090302 | J. Am. Chem. Soc. 2012, 134, 18816−18823

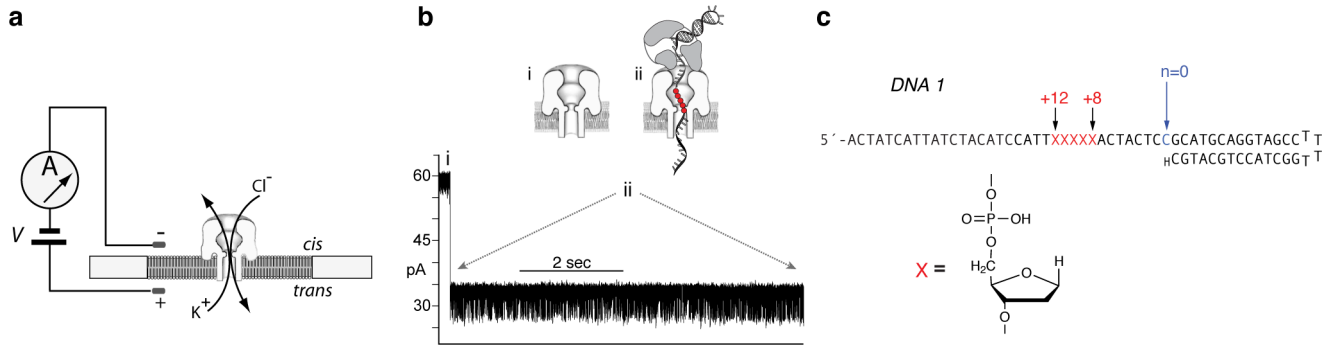


Figure 1. Capture of phi29 DNAP complexes on the α-HL nanopore. In the nanopore device (a), a single α-HL nanopore is inserted in a ~25 μm-diameter lipid bilayer that separates two chambers (cis and trans) containing buffer solution. Current through the nanopore is carried by K^+ and Cl^- ions. A patch clamp amplifier applies voltage and measures ionic current. (b) Representative current trace for a binary complex formed between phi29 DNAP and the DNA1 substrate captured at 180 mV applied potential in buffer containing 10 mM K-Hepes, pH 8.0, 0.3 M KCl, 1 mM EDTA, 1 mM DTT, and 11 mM $MgCl_2$. DNA and phi29 DNAP were added to the nanopore cis chamber to final concentrations of 1 and 0.75 μM, respectively. Cartoons above the current trace illustrate the sequence of events, which is described in the text. (c) Hairpin DNA substrate (DNA1) featuring a 14-base pair duplex region and a single-stranded template region of 35 nucleotides. The primer strand is terminated with a 2′,3′ CMP residue, and the template strand contains a reporter group of five consecutive abasic (1′,2′-H) residues spanning positions +8 to +12 (indicated as red X's in the sequence). The structure of an abasic residue is shown below the DNA sequence. In the cartoons in (a), the abasic residues are shown as red circles.

The equilibrium across the phi29 DNAP translocation step is influenced by applied force and by active-site proximal DNA sequences. Complementary dNTP stabilizes complexes in the post-translocation state but has negligible affinity for complexes in the pre-translocation state. The data support a model in which fluctuations between the pre-translocation state and post-translocation state are driven by Brownian thermal motion, and in which complexes are rectified to the post-translocation state by dNTP binding.[10] In the current study, we have examined the dynamic properties of the translocation state fluctuations in phi29 DNAP–DNA binary complexes. We extracted the forward and reverse translocation rates and quantified the effects of force and active-site proximal DNA sequences on each of the two rates. This analysis has allowed us to evaluate properties of the energy landscape of the phi29 DNAP translocation step, and the manner in which DNA sequences affect this landscape.

METHODS

DNA and Enzyme. DNA oligonucleotides were synthesized at Stanford Protein and Nucleic Acid Facility and purified by denaturing PAGE. DNA hairpins were annealed by heating at 90 °C for 4 min followed by snap cooling in ice water. Wild-type phi29 DNAP was obtained from Enzymatics (Beverly, MA).

Nanopore Methods. Nanopore experiments were conducted as described.[10,12−16] Briefly, a single α-HL nanopore is inserted in a ~25 μm-diameter lipid bilayer that separates two chambers containing buffer solution (10 mM K-Hepes, pH 8.0, 0.3 M KCl, and 1 mM EDTA). DTT and $MgCl_2$ were added to the cis chamber to final concentrations of 1 and 11 mM, respectively. Ionic current was measured with an integrating patch clamp amplifier (Axopatch 200B, Molecular Devices) in voltage clamp mode. Data were sampled using an analog-to-digital converter (Digidata 1440A, Molecular Devices) at 100 kHz in whole-cell configuration and filtered at 5 kHz using a low pass Bessel filter.

State Survival Probability Determinations. Dwell time in each of the two amplitude states was estimated using the single-channel detection function in Clampfit 10 (Molecular Devices). This software uses a half amplitude threshold method to assign transitions between two user-defined amplitude levels.[17] Amplitude levels for each of the two states were determined for the single-channel searches from histograms of all sampled amplitude data points. Five second segments of current trace from each of four independent capture events were analyzed for phi29 DNAP complexes formed with DNA1 (Figure 2b) or DNA4 (Figure S1). Complexes were captured at 180 mV. Each sampled segment contained >1440 transitions between the upper and lower amplitude states.

Error Estimation. The experimental data consist of recorded time traces of ionic current amplitude. Time traces contain numerous independent capture events, each starting with the capture of a random DNAP–DNA complex from the cis chamber and ending with the dissociation or ejection of that complex. Capture events that last ≥5 s are selected. For events that last ≥10 s, we use a segment of 10 s due to a size limit on data extraction in Clampfit. From each event, a sample value of (r_1, r_2) is calculated via the autocorrelation function. In the data measured for each DNA sequence and at each voltage, the number of capture events varies from 12 to 90 with a mean of 29 (Table S1). Thus, for each DNA sequence and at each voltage, we have m samples for r_k (k = 1, 2). Transition rate r_k is reported in the form of (sample mean ± standard error), calculated from the m samples.

For each DNA sequence, parameters (L_1, L_2, $\Delta G^{\ddagger}$, and ΔG^*) are calculated by fitting a straight line to $\log(r_k)$ versus voltage (k = 1, 2). To estimate the error in the values of L_1, L_2, $\Delta G^{\ddagger}$, and ΔG^*, we generate a collection of artificial data sets of r_k versus voltage. We treat the reported value of r_k (sample mean ± standard error) as describing a Gaussian distribution and draw independent samples from the distribution. Each artificial data set yields one set of L_1, L_2, $\Delta G^{\ddagger}$, and ΔG^*. The error is calculated as the sample standard deviation of the collection.

RESULTS AND DISCUSSION

A single α-HL nanopore is inserted into a lipid bilayer that separates two chambers (termed cis and trans) containing buffer solution (Figure 1a). A patch clamp amplifier applies voltage across the bilayer and measures the ionic current that flows through the nanopore, which is carried by K^+ and Cl^- ions in the buffer. Figure 1b shows a typical ionic current trace that results when a complex between phi29 DNAP and a DNA substrate (Figure 1c, DNA1) is captured atop the nanopore at 180 mV applied potential. The ionic current through the open channel (Figure 1b, i) drops rapidly when a complex is captured (Figure 1b, ii). The enzyme is too large to enter the nanopore and thus holds the duplex portion of the DNA substrate atop the pore. The DNA template strand is suspended

dx.doi.org/10.1021/ja3090302 | J. Am. Chem. Soc. 2012, 134, 18816−18823

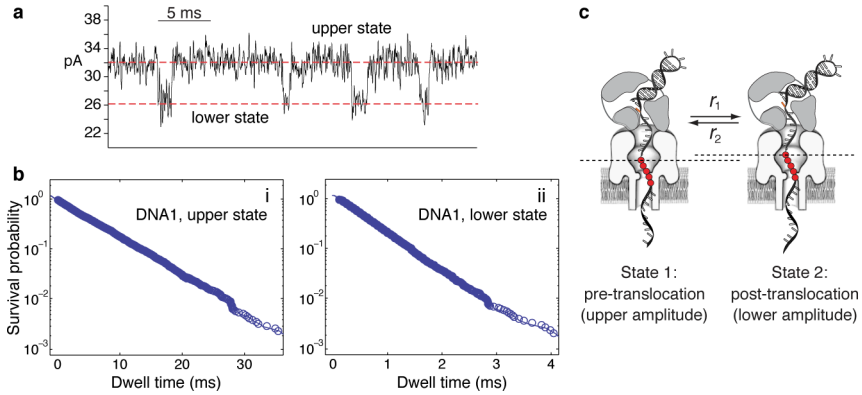


Figure 2. A two-state model for the dynamics of the phi29 DNAP translocation step. (a) Current trace segment for a captured phi29 DNAP–DNA1 complex. The ionic current fluctuates between ~32 and ~26 pA on the millisecond time scale as complexes are held atop the nanopore at 180 mV applied potential. (b) Survival probability of phi29 DNAP–DNA1 complexes in (b, i) the upper amplitude state, and (b, ii) the lower amplitude state. (c) The two-state kinetic model for the phi29 DNAP translocation step. The transitions between the two states are observed as fluctuations in the measured ionic current as the template strand with its embedded abasic reporter displaced in the nanopore lumen.

through the pore lumen, which is just wide enough to accommodate a single strand of DNA.

Phi29 DNAP–DNA complexes reside atop the nanopore for several seconds (Figure 1b, ii). During this period, the measured ionic current fluctuates between two amplitude levels. We have shown that the fluctuations between the two amplitude levels are due to movement of the DNA substrate relative to the enzyme and the nanopore, and that the distance of this movement is ~1 nucleotide.[10] The ability to detect the DNA displacement is achieved by the use of a reporter group in the template strand, comprising five consecutive abasic (1′,2′H) residues (Figure 1c). The fluctuations between the two amplitude levels are detected when the reporter group is displaced in the nanopore lumen.[10] In the upper amplitude state, the primer-template junction of the DNA substrate resides in the polymerase active site, in the pre-translocation state. At 180 mV, the pre-translocation state has an amplitude centered at ~32 pA. Complexes fluctuate on the millisecond time scale between the pre-translocation state and the second, lower amplitude state. In the lower amplitude state, the primer-template junction of the DNA substrate resides in the polymerase active site, in the post-translocation state.[10] The post-translocation state has an amplitude centered at ~26 pA. The fluctuations between the pre-translocation and post-translocation states continue until complexes dissociate or are ejected, after which another complex can be captured.

A Two-State Model for the phi29 DNAP Translocation Step. Close inspection of current traces for captured phi29 DNAP–DNA complexes shows that the measured ionic current fluctuates discretely between the two amplitude levels (Figure 2a), suggesting that along the direction of the DNA translocation, there are two discrete spatial states separated by an energy barrier. To assess whether the transitions between the two spatial states are described by a two-state kinetic model, we examined the dwell time distribution for complexes in each of the two states. We extracted dwell time samples from ionic current traces for phi29 DNAP–DNA1 complexes captured at 180 mV applied potential, using a half amplitude threshold method.[17] Plots of survival probability versus time for complexes in either the upper amplitude, pre-translocation state (Figure 2b, i; Figure S1) or the lower amplitude, post-translocation state (Figure 2b, ii; Figure S1) show that for both states, the dwell times are exponentially distributed. It follows that the transition in each direction is a single kinetic step that can be fully characterized by a transition rate. We therefore use a two-state kinetic model (Figure 2c) to describe the translocation state fluctuations. Complexes fluctuate between two discrete states, with two transition rates: r_1, the forward transition from the upper amplitude, pre-translocation state to the lower amplitude post-translocation state, and r_2, the reverse transition from the post-translocation state to the pre-translocation state (Figure 2c).

Extracting Transition Rates from the Measured Ionic Current Traces. The extraction of dwell time samples using the half amplitude threshold method is subject to errors caused by measurement noise. The errors are not significant when the standard deviation of the measurement noise is small relative to one-half of the distance between two amplitude levels.[17] This is the case when phi29 DNAP–DNA complexes are captured at higher voltages, where the amplitudes of the two states are well separated and thus have a high signal-to-noise ratio. We have applied the half amplitude threshold method to extract dwell time samples from data collected at 180 mV (Figure 2b) and have used these dwell time samples to conclude that the transition in each direction is a single kinetic step. However, to study the transition rates quantitatively across a range of voltages, it is necessary to adopt a more robust method of extracting rates r_1 and r_2 from the measured ionic current traces that can average out the effect of measurement noise. Previously, Bezrukov and Kasianowicz used the autocorrelation function and the associated power spectral density to study protonation kinetics of the α-HL nanopore open channel.[18] We consider the autocorrelation of the time trace. Let I_1 be the ionic current (without measurement noise) of state 1, I_2 is the ionic current (without measurement noise) of state 2, and $S(t)$ is the stochastic state of DNAP–DNA complex at time t.

The ionic current (without measurement noise) at time t is

$$I(t) = \begin{cases} I_1, & S(t) = 1 \\ I_2, & S(t) = 2 \end{cases}$$

The measured ionic current (with measurement noise) is

$$X(t) = I(t) + N(t)$$

We assume that the measurement noise $N(t)$ has zero mean and that $N(t_1)$ is independent of $N(t_2)$.

dx.doi.org/10.1021/ja3090302 | J. Am. Chem. Soc. 2012, 134, 18816–18823

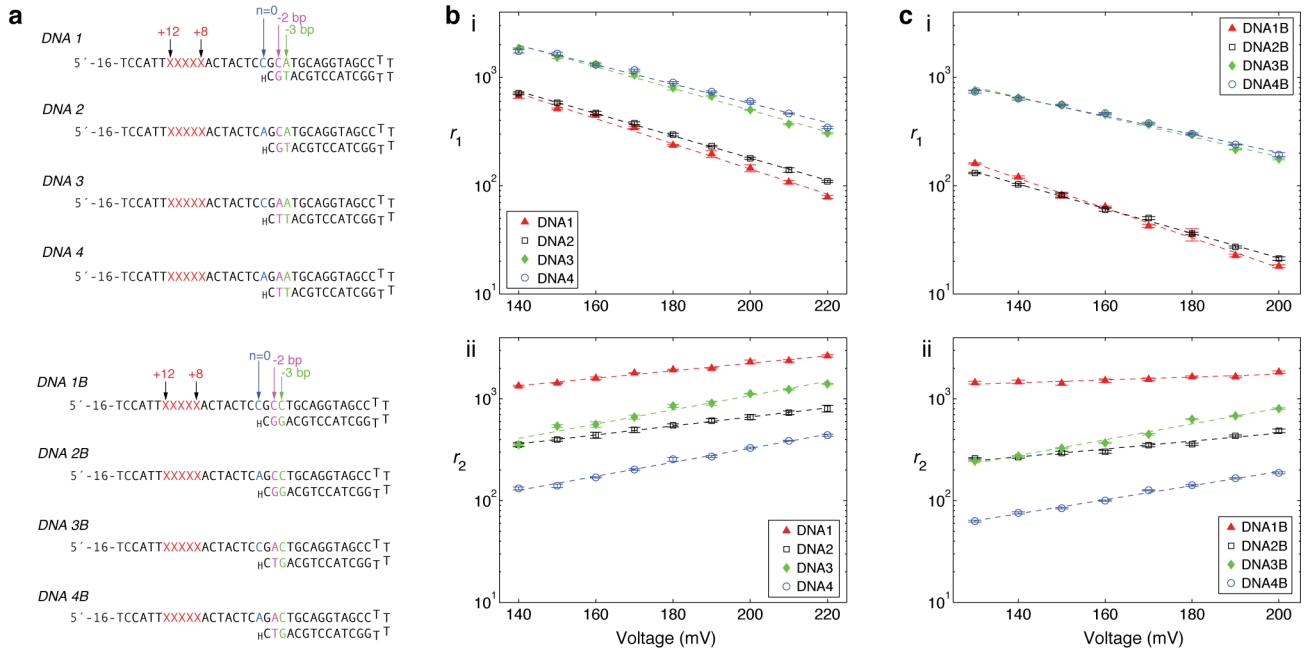


Figure 3. Transition rates of the phi29 DNAP translocation step extracted from ionic current traces using the autocorrelation function. (a) Sequences of the DNA substrates. The sequence determinants that varied in the DNA1−DNA4 series are the residue at template position $n = 0$ (highlighted in blue) and the −2 base-pair of the duplex (highlighted in purple). Sequences in the DNA1B−DNA4B series are the same as in the DNA1−DNA4 series, except that the T−A base-pair (primer-template) at the −3 position of the duplex has been changed to a G−C pair (highlighted in green). The eight DNA substrates are otherwise identical, including the abasic reporter from template positions +8 to +12 (shown as red X's). Sixteen residues at the 5′ end of the template strand are not shown in this panel but can be seen in Figure 1b. Ionic current traces for complexes formed with each of the eight DNAs, captured at 180 mV, are shown in Figure S2. Plots of (b, i and c, i) $\log(r_1)$ versus voltage, and (b, ii and c, ii) $\log(r_2)$ versus voltage, for phi29 DNAP complexes formed with DNAs 1−4 or DNAs 1B−4B. Each plotted point of $\log(r_1)$ or $\log(r_2)$ shows mean ± standard error, which is calculated from a population of 12−90 captured complexes for each of the DNA substrates at each voltage (see Table S1).

Amplitude levels I_1 and I_2 are determined by fitting the distribution of $X(t)$ to a model of two Gaussian modes, as described previously.[10] Once we know the values of I_1 and I_2, we map $X(t)$ to $Y(t)$ and consider the autocorrelation function of Y:

$$Y(t) = \frac{2}{(I_2 - I_1)}\left(X(t) - \frac{I_1 + I_2}{2}\right)$$

$$R(t) = E[Y(t_0)Y(t_0 + t)]$$

The analytical expression of the autocorrelation function in terms of transition rates can be found in ref 19. It follows that $E[Y] = (r_1 - r_2)/(r_1 + r_2)$ and

$$\log(R(t) - (E[Y])^2) = \log(1 - (E[Y])^2) - (r_1 + r_2)t, \quad t > 0$$

Transition rates r_1 and r_2 are expressed as

$$r_1 = \frac{1}{2}(1 + E[Y]) \times (\text{slope of } \log(R(t) - (E[Y])^2))$$

$$r_2 = \frac{1}{2}(1 - E[Y]) \times (\text{slope of } \log(R(t) - (E[Y])^2))$$

where $R(t)$ and $E[Y]$ are estimated from the data $\{(t_i, Y_i), i = 1, 2, .., N\}$ as

$$E[Y] \approx \frac{1}{N}\sum_{i=1}^{N} Y_i,\ R(t_j) \approx \frac{1}{N-j}\sum_{i=1}^{N-j} Y_i \cdot Y_{i+j},$$

$$j = 1, 2, \ldots$$

Fitting an exponential function to data points of $R(t)$, and identifying and avoiding the effect of filtering on $R(t)$, is discussed in Figures S2 and S3.

The Forward and Reverse Translocation Rates. Force is exerted on phi29 DNAP−DNA complexes when they reside atop the pore, due to the voltage pulling on the DNA template strand. The direction of the applied force opposes the forward translocation, and thus voltage shifts the translocation equilibrium toward the pre-translocation state.[10] The effects of the applied force on the equilibrium across the translocation step could be exerted either by impeding the rate of forward translocation, by increasing the rate of reverse translocation, or by a combined effect on both rates. Plots of $\log(r_1)$ versus voltage and $\log(r_2)$ versus voltage for phi29 DNAP complexes formed with the DNA1 substrate both fit to straight lines (Figure 3b,i and b,ii, red triangles). At 180 mV, the transition rates for complexes formed with DNA1 are r_1 = 235.71 ± 4.4 s^{-1} and r_2 = 1932.9 ± 25.06 s^{-1}. Both rates are affected by the applied force; across the series of applied voltages, r_1 decreases from 672.94 ± 33.68 s^{-1} at 140 mV, to 78.68 ± 2.42 s^{-1} at 220 mV; r_2 increases from 1337.9 ± 24.34 s^{-1} at 140 mV, to 2643.2 ± 69.06 s^{-1} at 220 mV.

Sequences proximal to the polymerase active site in both the duplex and the single-stranded template of DNA substrates influence the equilibrium across the translocation step.[10] We

dx.doi.org/10.1021/ja3090302 | J. Am. Chem. Soc. 2012, 134, 18816−18823

evaluated the effects on the transition rates caused by changing active-site proximal DNA sequence determinants, using eight DNA substrates that differ from each other at the −2 and −3 base-pairs of the duplex, or at n = 0 of the template strand (Figure 3a). Sample ionic current traces for complexes formed with each of the eight DNA substrates, captured at 180 mV, are shown in Figure S4. Plots of $\log(r_1)$ versus voltage and $\log(r_2)$ versus voltage for complexes formed with each of the eight DNA substrates fit to straight lines across the voltage range examined (Figure 3b, i and ii and c, i and ii). When complexes formed with different DNA substrates are compared, both the vertical intercepts and the slopes of log(rate) versus voltage vary, indicating that active-site proximal DNA sequences affect both the rate at a given voltage and the dependence of the rate on voltage. To evaluate the effects of active-site proximal DNA sequences on the phi29 DNAP translocation step in a biophysical framework, we studied the free energy landscape of the two translocation states and the transitions between them.

The Free Energy Landscape of the Translocation Step. We consider each transition rate as the escape rate from a potential well, and adopt a schematic representation of the energy landscape (Figure 4). Our approach is to use a simple model so that the model parameters can be determined from the data. The schematic energy landscape is described by four parameters: L_1, distance between the pre-translocation state and the transition state; L_2, distance between the post-translocation state and the transition state; ΔG_1, free energy difference between the pre-translocation state and the transition state; and ΔG_2, free energy difference between the post-translocation state and the transition state.

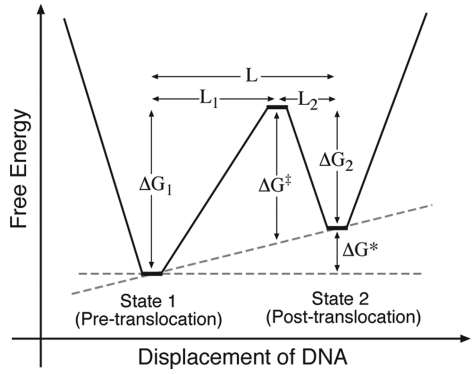


Figure 4. Schematic representation of the free energy landscape of the two translocation states and the transitions between them. The energy landscape is characterized by four parameters (L_1, L_2, ΔG_1, ΔG_2), or equivalently (L_1, L_2, $\Delta G^{\ddagger}$, ΔG^*). The advantage of using $\Delta G^{\ddagger}$ is that it is independent of voltage and provides a measure of energy barrier height (see text for details).

$L = L_1 + L_2$ is the distance between the pre-translocation and post-translocation states.

The transition rates, as given by Kramers approximation, have the expressions:

$$r_k \approx \gamma D \exp\left(\frac{-\Delta G_k}{k_B T}\right), \quad k = 1, 2$$

where D is the diffusivity of the DNA with respect to the pore, and γ is a constant reflecting the geometric detail of energy landscape near the three extremes (the bottoms of the two potential wells and the top of the energy barrier). In our simple schematic representation of energy landscape, the geometric detail near the three extremes is not specified, and there is not enough information from the data to determine the geometric detail. To proceed, we absorb the effect of the geometric detail near the three extremes into the free energy differences ΔG_1 and ΔG_2. Consequently, constant γ becomes independent of voltage and DNA sequence. In this framework, ΔG_k (k = 1, 2) contains the effects of both the depth of the potential well and the geometric detail near extremes.

The free energy differences vary with the applied voltage:

$$\Delta G_1 = \Delta G_1|_{V=140} + \alpha(V - 140)L_1$$

$$\Delta G_2 = \Delta G_2|_{V=140} - \alpha(V - 140)L_2$$

where V is the applied voltage, αV is the force induced by the voltage, and α is the constant coefficient relating the applied voltage and the force. Log(rate) versus voltage has the expression below:

$$\log(r_1) \approx -(V - 140)\frac{\alpha L_1}{k_B T} - \frac{\Delta G_1|_{V=140}}{k_B T} + \log(\gamma D)$$

$$\log(r_2) \approx (V - 140)\frac{\alpha L_2}{k_B T} - \frac{\Delta G_2|_{V=140}}{k_B T} + \log(\gamma D)$$

For k = 1, 2, respectively, $(\Delta G_k|_{V=140})/(k_B T)$ and $(\alpha L_k)/(k_B T)$ are calculated by fitting a straight line to $\log(r_k)$ versus V. In the calculation, we set $\log(\gamma D)$ = 18, which is consistent with a value of 10^{-8}–10^{-7} cm^2/s for the translational diffusion of

Table 1. Parameters of the Free Energy Landscape of the Translocation Step[a]

DNA substrate[b]	L[c]	L_1[d]	L_2[e]	$\Delta G^{\ddagger}$[f]	ΔG^* (140 mV)[g]
DNA1	1 ± 0.025	0.760 ± 0.023	0.241 ± 0.010	10.949 ± 0.021	0.637 ± 0.044
DNA2	0.960 ± 0.024	0.671 ± 0.014	0.289 ± 0.020	11.894 ± 0.015	−0.723 ± 0.039
DNA3	1.113 ± 0.044	0.655 ± 0.021	0.458 ± 0.039	11.335 ± 0.019	−1.582 ± 0.085
DNA4	1.030 ± 0.038	0.581 ± 0.036	0.449 ± 0.014	11.964 ± 0.022	−2.744 ± 0.067
DNA1B	0.988 ± 0.029	0.896 ± 0.021	0.092 ± 0.019	10.963 ± 0.061	2.513 ± 0.029
DNA2B	0.995 ± 0.029	0.741 ± 0.014	0.254 ± 0.025	12.649 ± 0.037	0.958 ± 0.033
DNA3B	1.106 ± 0.034	0.600 ± 0.027	0.506 ± 0.020	11.979 ± 0.015	−0.861 ± 0.042
DNA4B	1.001 ± 0.027	0.550 ± 0.024	0.451 ± 0.012	12.716 ± 0.017	−2.165 ± 0.035

[a]All distances are reported in units L(DNA1); see text for details. Error bars indicate the standard errors; see Methods for a description of the error estimations. [b]DNA sequences are shown in Figure 3. [c]L = distance of the translocation step. [d]L_1 = distance from the pre-translocation state to the transition state of the translocation step. [e]L_2 = distance from the transition state of the translocation step to the post-translocation state. [f]$\Delta G^{\ddagger}$ = height of the transition state energy barrier ($k_B T$). [g]ΔG^* = free energy difference ($k_B T$) between the pre-translocation and post-translocation states at 140 mV applied potential.

dx.doi.org/10.1021/ja3090302 | J. Am. Chem. Soc. 2012, 134, 18816−18823

DNA.[20] Because the value of α is unknown, we normalize all distances by L(DNA1). That is, all values of distances are reported in units of L(DNA1).

Both ΔG_1 and ΔG_2 are related to the energy barrier height, and both vary with the voltage. To separate the dependence on voltage and the dependence on energy barrier height, instead of using ΔG_1 and ΔG_2, we introduce $\Delta G^{\ddagger} = (L_2/L)\Delta G_1 + (L_1/L)\Delta G_2$, which is independent of voltage and measures the energy barrier height. We also introduce $\Delta G^* = \Delta G_1 - \Delta G_2$, which is the free energy difference between the two states and is not affected by the energy barrier height. We thus use four parameters to specify the schematic energy landscape: L_1, L_2, $\Delta G^{\ddagger}$, and $\Delta G^*|_{V=140}$. We determined the values for these four parameters from the plots of the extracted rates, for complexes formed between phi29 DNAP and each of the eight DNA substrates (Table 1).

The Distance of the Translocation Step (L). The values of L, the distance between the pre-translocation and the post-translocation states (reported in units of L(DNA1)), for complexes formed with each of the eight DNA substrates are given in Table 1. In prior mapping experiments, using DNA substrates with the same active-site proximal sequences as DNA1, in which the position of the abasic block in the template strand was varied in single nucleotide increments, we established that the translocation distance between the two amplitude states is ~1 nucleotide.[10] Thus, L(DNA1), the unit used for reporting all distances in Table 1, is ~1 nucleotide.

When the values of L for complexes formed with the eight DNA substrates are compared, small differences in the distance of the displacement are revealed (Table 1). These differences may reflect small effects of sequence on the binding of DNA in the two translocation states of DNAP−DNA complexes. The differences may be exacerbated by the absence in these experiments of the pyrophosphate and dNTP ligands that stabilize the pre-translocation and post-translocation states, respectively. Crystal structures of complexes in the post-translocation state indicate that DNA recognition by phi29 DNAP is mediated almost entirely through interactions that are sequence-independent, as expected for an enzyme that must rapidly replicate DNA without significant sequence bias.[1,21] Nonetheless, there may be small sequence-dependent differences in the length or angles of noncovalent bonds between the enzyme and DNA in the active site that are revealed when the fluctuations across the translocation step are examined over a range of force.

The Location of the Transition State (L_1 and L_2). While the effects of DNA substrate sequences on the distance between the pre-translocation and post-translocation states are modest, their effects on the location of the transition state along the translocation direction are more pronounced. The location of the transition state is specified by L_1 (distance from the pre-translocation state to the transition state) and L_2 (distance from transition state to the post-translocation state). The distances range from a transition state that is located ~90% along the pathway to the post-translocation state for complexes formed with DNA1B, to a transition state located ~55% along the pathway to the post-tranlocation state for complexes formed with DNA4B (Table 1). During each forward or reverse transition, a set of noncovalent bonds between the enzyme and the DNA substrate is broken and a different set is formed. The DNAP motor has a step size along the direction of translocation that is only a few angstroms. It is likely that the translocation step is accompanied by a conformational change perpendicular to the direction of translocation (partial or complete fingers domain opening and closing) that cannot be directly observed in our experiments but which contributes to the energy landscape. Sequence-dependent differences in the strength and location of the sets of noncovalent bonds between the enzyme and DNA as the pathway is traversed may contribute to the changes in the location of the transition state caused by varying the active-site proximal DNA sequences.

The two series of substrates, DNAs 1−4 and DNAs 1B−4B, differ from one another in the identity of the base-pair at the −3 position in the duplex (Figure 3a). Within both of the substrate series, the identities of the determinants at the −2 base-pair and at the template n = 0 position individually affect the location of the transition state; a T−A pair at the −2 position shifts the transition state closer to the pre-translocation state relative to a G−C pair at this position. Likewise, a dAMP residue at n = 0 shifts the transition state closer to the pre-translocation state, relative to a dCMP residue at this position. Within both DNA substrate series, the combination of the T−A −2 base-pair and the dAMP residue at n = 0 moves the location of the transition state toward the pre-translocation state more than either individual determinant does, although the effects are not linearly additive (Table1).

The Height of the Transition State Barrier ($\Delta G^{\ddagger}$). The height of the energy barrier that separates the pre-translocation and post-translocation states is measured by the parameter $\Delta G^{\ddagger}$, defined in terms of ΔG_1 and ΔG_2. The meaning of $\Delta G^{\ddagger}$ can be understood by connecting the two translocation state wells with a straight line in the free energy plot (Figure 4). $\Delta G^{\ddagger}$ is the height of the energy barrier above that line. $\Delta G^{\ddagger}$ is independent of voltage even though both ΔG_1 and ΔG_2 (the depths of the two potential wells) vary with the applied voltage.

As described above, ΔG_1 and ΔG_2 (and consequently $\Delta G^{\ddagger}$) are calculated using $\log(\gamma D) = 18$. While $\log(\gamma D) = 18$ is a reasonable estimate, it is probably not exact. If the value of $\log(\gamma D)$ is changed, the estimated values of ΔG_1, ΔG_2, and $\Delta G^{\ddagger}$ will all change by the same amount. However, it is important to note that differences between the values of $\Delta G^{\ddagger}$ for complexes formed with different DNA substrates are independent of $\log(\gamma D)$; the reported values of $\Delta G^{\ddagger}$ may differ from the true values by a constant, but the effect of DNA substrates on $\Delta G^{\ddagger}$ is correctly reflected in Table 1. Thus, to compare the free energy barrier that separates the pre-translocation and post-translocation states among complexes formed with DNA substrates that vary in sequence, we compare the values of $\Delta G^{\ddagger}$. There are differences of 1−2 k_BT in the values for this parameter as a function of DNA substrate sequence (Table 1), suggesting that the height of the barrier to the translocation step is not dramatically affected by substrate sequences as they pass through the active site during DNA synthesis.

The Free Energy Difference between the Pre-translocation and the Post-translocation States (ΔG^*). The parameter $\Delta G^* = \Delta G_1 - \Delta G_2$ is unaffected by the value of $\log(\gamma D)$ used in the calculation of ΔG_1 and ΔG_2. In contrast to the modest differences in $\Delta G^{\ddagger}$ caused by differences in DNA substrate sequences, the differences in ΔG^* are substantial. The base-pair at the −3 position of the duplex in DNAs 1−4 is a T−A pair, while in DNAs 1B−4B it is a G−C pair. For each of the four members of both DNA series, the G−C pair at the −3 position shifts the translocational equilibrium toward the pre-translocation state relative to the otherwise identical substrate bearing a T-A pair at the −3 position. Within both of the

substrate series, the determinants at the −2 base-pair and at the template $n = 0$ position separately affect the equilibrium across the translocation step. A T-A pair at the −2 position shifts the equilibrium toward the post-translocation state, relative to a G−C pair at this position. Likewise, a dAMP residue at $n = 0$ shifts the equilibrium toward the post-translocation state, relative to a dCMP residue at this position. Reminiscent of the effects of these two sequence determinants on the distance to the transition state, within both of the DNA substrate series the combination of the T−A −2 base-pair and the dAMP residue at $n = 0$ shifts the equilibrium state further toward the post-translocation state than either individual determinant does, although these effects are not linearly additive (Table1).

The Influence of Active-Site Proximal DNA Substrate Sequences on Translocation Step Dynamics. Qualitative inspection of the rate plots and energy landscape parameters shows that some trends, such as the effects of the base-pair at −2 position and the residue at $n = 0$ on the location of the transition state and on the equilibrium parameter ΔG^*, can be attributed to the interactions of phi29 DNAP with individual DNA sequence determinants. Notably, the identity of the residue at $n = 0$ has a more profound effect on the reverse transition (r_2) than it does on the forward transition (r_1); its effects are manifested predominantly in the vertical displacement of the fitted lines for the response of r_2 to voltage (Figure 3b, i, ii, and c, i, ii), suggesting that it affects the depth of the post-translocation state well (ΔG_2) but not the pre-translocation state well (ΔG_1). This may arise from a contribution of the nucleotide to the energy barrier for the reverse transition (r_2) that relates to movement of the kink in the template strand backbone between the $n = 0$ and $n = +1$ residues observed for phi29 DNAP (and other DNAPs) in post-translocation state crystal structures.[1]

When the combined effects of active-site proximal DNA substrate sequences on the phi29 DNAP translocation step are evaluated quantitatively, for the most part individual sequence determinants do not display simple linear (additive or subtractive) effects on the parameters of the free energy landscape. Instead, the influences of each of the sequence determinants are context dependent. These results suggest that phi29 DNAP interacts with active-site proximal DNA sequences (including, but not limited to, the subset that we have examined) along the translocation pathway as a composite array.

The ability to analyze the dynamics of the phi29 DNAP translocation step in detail relies on the unique combination of spatial and temporal resolution afforded by the nanopore assay. In this study, the fluctuation rates across the translocation step were measured in the absence of ligands, and thus they apply to the behavior inherent to binary complexes in response to thermal fluctuations. This behavior obtains during the point in the catalytic cycle that follows the release of pyrophosphate and precedes incoming dNTP binding. We are extending our analysis to the behavior of complexes in the presence of complementary dNTP, which binds to and rectifies complexes in the post-translocation state, driving the motor forward. In a subsequent study, we will simultaneously extract the forward and reverse translocation rates and the dNTP association and dissociation rates. This will further broaden our understanding of the mechanism of this essential DNA polymerase motor step and its integration into the catalytic cycle.

During DNA replication in vivo, phi29 DNAP displaces the downstream DNA strand complementary to the template by actively promoting unwinding of the strands ahead of the advancing enzyme.[11,22] Recent single-molecule optical trap experiments have revealed that downstream duplex sequences can influence the rate of synthesis, with G−C-rich tracts slowing the rate.[22] It is plausible that these effects on the synthesis rate, imposed by the necessity for the polymerase to unwind downstream duplexes of differing stability, affect the forward translocation rate (rather than the rate of other steps in the nucleotide addition cycle). The differences in the synthesis rates were observed in the presence of dNTPs, which would strongly impede the reverse translocation. Thus, the more stable downstream duplex segments may in effect exert a stronger opposing force to the forward translocation. The downstream DNA strand complementary to the template is not present in the DNA substrates used in the nanopore experiments, and thus it remains to be seen whether such downstream sequence effects could obscure (or amplify) the effects of active-site proximal sequences on the energy landscape of the translocation step.

■ ASSOCIATED CONTENT

Supporting Information

Figure S1 (survival probability plots for dwell times in the upper and lower amplitude states for phi29 DNAP−DNA4 complexes); Figure S2 (autocorrelation for a complex formed between phi29 DNAP and a DNA substrate with a template strand consisting of poly dCMP); Figure S3 (fitting an exponential function to autocorrelation for complexes formed with DNA1, captured at 220 and 180 mV); Figure S4 (sample ionic current traces for complexes formed between phi29 DNAP and DNAs 1−4 and 1B−4B); and Table S1 (number of phi29 DNAP−DNA captured complexes used to determine the translocation rates). This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

krlieberman@gmail.com; hongwang@soe.ucsc.edu

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We are grateful to Jacob Schrieber for software, Robin Abu-Shumays for discussions, and Ashley Cox, William Dunbar, Jungsuk Kim, Raj Maitra, and Chris O'Donnell for technical assistance. This work was supported by NIH Grant 1R01GM087484-02 from NIGMS (to K.R.L. and M.A.) and by NSF Grant DMS-0719361 (to H.W.).

■ REFERENCES

(1) Berman, A. J.; Kamtekar, S.; Goodman, J. L.; Lázaro, J. M.; de Vega, M.; Blanco, L.; Salas, M.; Steitz, T. A. *EMBO J.* **2007**, *26*, 3494−3505.

(2) Golosov, A. A.; Warren, J. J.; Beese, L. S.; Karplus, M. *Structure* **2010**, *18*, 83−93.

(3) Johnson, S. J.; Taylor, J. S.; Beese, L. S. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 3895−3900.

(4) Yin, Y. W.; Steitz, T. A. *Cell* **2004**, *116*, 393−404.

(5) Doublié, S.; Tabor, S.; Long, A. M.; Richardson, C. C.; Ellenberger, T. *Nature* **1998**, *391*, 251−258.

(6) Franklin, M. C.; Wang, J.; Steitz, T. A. *Cell* **2001**, *105*, 657−667.

(7) Berezhna, S. Y.; Gill, J. P.; Lamichhane, R.; Millar, D. P. *J. Am. Chem. Soc.* **2012**, *134*, 11261−11268.

(8) Wang, W.; Wu, E. Y.; Hellinga, H. W.; Beese, L. S. *J. Biol. Chem.* **2012**, *287*, 38215−38226.
(9) Wu, E. Y.; Beese, L. S. *J. Biol. Chem.* **2011**, *286*, 19758−19767.
(10) Dahl, J. M.; Mai, A. H.; Cherf, G. M.; Jetha, N. N.; Garalde, D. R.; Marziali, A.; Akeson, M.; Wang, H.; Lieberman, K. R. *J. Biol. Chem.* **2012**, *287*, 13407−13421.
(11) Salas, M.; Blanco, L.; Lázaro, J. M.; de Vega, M. *IUBMB Life* **2008**, *60*, 82−85.
(12) Lieberman, K. R.; Cherf, G. M.; Doody, M. J.; Olasagasti, F.; Kolodji, Y.; Akeson, M. *J. Am. Chem. Soc.* **2010**, *132*, 17961−17972.
(13) Akeson, M.; Branton, D.; Kasianowicz, J. J.; Brandin, E.; Deamer, D. W. *Biophys. J.* **1999**, *77*, 3227−3233.
(14) Benner, S.; Chen, R. J.; Wilson, N. A.; Abu-Shumays, R.; Hurt, N.; Lieberman, K. R.; Deamer, D. W.; Dunbar, W. B.; Akeson, M. *Nat. Nanotechnol.* **2007**, *2*, 718−724.
(15) Garalde, D. R.; Simon, C. A.; Dahl, J. M.; Wang, H.; Akeson, M.; Lieberman, K. R. *J. Biol. Chem.* **2011**, *286*, 14480−14492.
(16) Hurt, N.; Wang, H.; Akeson, M.; Lieberman, K. R. *J. Am. Chem. Soc.* **2009**, *131*, 3772−3778.
(17) Colquhoun, D.; Sigworth, F. J. In *Single-Channel Recording*; Sakmann, B., Neher, E., Eds.; Plenum Press: New York, 1995; pp 483−587.
(18) Bezrukov, S. M.; Kasianowicz, J. J. *Phys. Rev. Lett.* **1993**, *70*, 2352−2355.
(19) Kogan, Sh. *Electronic Noise and Fluctuations in Solids*; Cambridge University Press: New York, 1996.
(20) Lukacs, G. L.; Haggie, P.; Seksek, O.; Lechardeur, D.; Freedman, N.; Verkman, A. S. *J. Biol. Chem.* **2000**, *275*, 1625−1629.
(21) Wang, M.; Xia, S.; Blaha, G.; Steitz, T. A.; Konigsberg, W. H.; Wang, J. *Biochemistry* **2011**, *50*, 581−590.
(22) Morin, J. A.; Cao, F. J.; Lazaro, J. M.; Arias-Gonzalez, J. R.; Valpuesta, J. M.; Carrascosa, J. L.; Salas, M.; Ibarra, B. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, 8115−8120.