Conformational dynamics during high-fidelity DNA replication and translocation defined using a DNA polymerase with a fluorescent artificial amino acid

Received for publication, October 29, 2020, and in revised form, December 2, 2020. Published, Papers in Press, December 3, 2020, https://doi.org/10.1074/jbc.RA120.016617

Tyler L. Dangerfield* and Kenneth A. Johnson**

From the Institute for Cellular and Molecular Biology, Department of Molecular Biosciences, University of Texas, Austin, Texas, USA

Edited by F. Peter Guengerich

We address the role of enzyme conformational dynamics in specificity for a high-fidelity DNA polymerase responsible for genome replication. We present the complete characterization of the conformational dynamics during the correct nucleotide incorporation forward and reverse reactions using stopped-flow and rapid-quench methods with a T7 DNA polymerase variant containing a fluorescent unnatural amino acid, (7-hydroxy-4-coumarin-yl) ethylglycine, which provides a signal for enzyme conformational changes. We show that the forward conformational change (7–6000 s−1) is much faster than chemistry (300 s−1) while the enzyme opening to allow release of bound nucleotide (1.7 s−1) is much slower than chemistry. These parameters show that the conformational change selects a correct nucleotide for incorporation through an induced-fit mechanism. We also measured conformational changes occurring after chemistry and during pyrophosphorolysis, providing new insights into processive polymerization. Pyrophosphorolysis occurs via a conformational selection mechanism as the pyrophosphate binds to a rare pretranslocation state of the enzyme–DNA complex. Global data fitting was achieved by including experiments in the forward and reverse directions to correlate conformational changes with chemical reaction steps. This analysis provided well-constrained values for nine rate constants to establish a complete free-energy profile including the rates of DNA translocation during processive synthesis. Translocation does not follow Brownian ratchet or power stroke models invoking nucleotide binding as the driving force. Rather, translocation is rapid and thermodynamically favorable after enzyme opening and pyrophosphate release, and it appears to limit the rate of processive synthesis at 4 °C.

DNA polymerases are ideal systems to study enzyme specificity because fidelity is physiologically important and alternative substrates (noncognate base pairs) are well known. DNA polymerases alter their substrate specificity during each cycle of template-dependent nucleotide incorporation, allowing efficient selection of the correct nucleotide over the structurally similar mismatched nucleotides (1). Selection of the correct nucleotide is assisted by base pairing of the incoming nucleotide with the templating base. However, the small free energy differences between base pairing for correct versus incorrect bases cannot account for the extremely high specificity of these enzymes (2), although hydrogen bonds do enforce base-pair geometry that contributes to satisfying steric requirements (3). Proofreading mechanisms further improve overall accuracy, but exonuclease specificity is dictated by the binding of the primer/template at the polymerase site to modulate transfer to the exonuclease site (4). DNA polymerases vary widely from low-fidelity repair enzymes (5–9) to high-fidelity enzymes necessary for genome replication (10–12). It is surprising that high-fidelity enzymes catalyze replication at faster rates than low-fidelity enzymes, raising additional questions regarding the molecular mechanisms responsible for enzyme specificity.

DNA polymerase from bacteriophage T7 provides a simple enzyme capable of fast and accurate DNA replication. Crystal structures of the T7 DNA polymerase (Fig. 1) show large structural rearrangements, particularly in the “fingers” domain, upon nucleotide binding (13). In a minimal model (Equation 1), the enzyme–DNA complex with a primer strand n-nucleotides in length (EDn) binds nucleotide (N) in the open state. A transition to the closed state (EDnN → FDnN) then facilitates the chemical reaction followed by enzyme opening and release of pyrophosphate (PPi).

\[
ED_n + N \rightleftharpoons k_1 \quad ED_n N \quad \rightleftharpoons k_2 \quad FD_n N \quad \rightleftharpoons k_3 \quad FD_{n+1} PP_i \quad \rightleftharpoons k_4 \quad ED_{n+1} + PP_i
\]

(1)

Similar conformational changes have been observed for both high- and low-fidelity polymerases, but the contribution of these prechemistry steps to enzyme specificity has been controversial, leaving many questions unanswered. What role does this conformational change play in selecting the correct nucleotide over the three incorrect but structurally similar substrates? Is the rate of the conformational change faster or slower than the chemistry and how does this affect the fidelity? How do the kinetics of the conformational change differ when comparing low- and high-fidelity DNA polymerases?

* For correspondence: Kenneth A. Johnson, kajohnson@mail.utexas.edu.
There is little question that the change in structure from an open to a closed state leads to alignment of catalytic residues, and that certainly influences the rate of the chemical reaction. However, when the conformational change is faster than the chemistry, the question as to whether the structural change directly influences specificity must be based on complete kinetic analysis of the pathway. Because of the disparate results obtained for low- and moderate-fidelity polymerases, it is crucial that the role of the conformational change be defined for a high-fidelity enzyme. Previous attempts to examine the kinetics of the conformational change using T7 DNA polymerase were flawed because they were based on constructing a cys-lite mutant to afford site-specific labeling with a fluorescent probe (1). Unfortunately, the cys-lite variant containing nine mutations reduced the fidelity and stability of the enzyme. Here we take advantage of new methods to site specifically label T7 DNA polymerase with a fluorescent unnatural amino acid, (7-hydroxy-4-coumarin-yl)ethylglycine, under conditions where the fluorescent variant retains the fidelity and reaction kinetics of the wildtype enzyme (21). Unless noted otherwise, all experiments were performed with the labeled enzyme.

Prior measurements of translocation using single molecule methods (22, 23) produced data that suggest a rapid equilibrium constant for translocation with an equilibrium constant near unity, so that dNTP binding will trap the DNA in the translocated state. However, it is important to note that single molecule measurements at equilibrium do not necessarily reflect the population of intermediates formed during processive synthesis. Here we provide a unique solution to the problem by determining the rate constants for translocation during synthesis as part of global data fitting.

We use the coumarin fluorescence signal to measure the rates of the forward and reverse conformational change steps for the enzyme in concert with measurements of the rates of polymerization. Corresponding measurements for the reverse reaction (pyrophosphorolysis) and global data fitting of all experiments provides well-constrained rate constants to define a complete free-energy profile for the reaction pathway, including DNA translocation.

Results

Kinetics and equilibrium of nucleotide binding

We first measured the kinetics and equilibrium for nucleotide binding steps \((K_1, k_{2}, \text{and } k_{-2} \text{ in Equation } 2)\) using a 2',3'-dideoxy-nucleotide-terminated primer to allow dATP binding but prevent the chemical reaction.

\[
ED_{dd} + dATP \rightleftharpoons k_1 \cdot ED_{dd}dATP \rightleftharpoons \frac{k_1}{k_{-2}} FD_{dd}dATP
\]

A 27-nt (nucleotide) 2',3'-dideoxy-terminated primer strand was annealed to a 45-nt template strand (27\(d_d\)/45, denoted as \(D_{dd}\)) and then incubated with the labeled enzyme to form the \(ED_{dd}\) complex. Initial experiments performed at 20 °C indicated that the rate of the conformational change \((k_2)\) was too
Conformational dynamics for T7 DNA polymerase

Figure 2. Kinetics and equilibrium of nucleotide binding steps preceding chemistry. A, stopped flow dATP binding rate. T7 DNA polymerase E514Cou (750 nM), thioredoxin (15 μM), and DNAad (1 μM) were mixed with dATP (5–100 μM) and Mg²⁺ (12.5 mM) to start the reaction. Colored dots are stopped flow traces at various dATP concentrations, and black lines are single exponential fits to the data. B, decay rate versus [dATP]. Rates are derived from the single exponential fits to the data in (A). The dashed gray line through the data points is the fit to a linear function with a y-intercept of 0 and a slope of 14.5 ± 0.3 μM⁻¹ s⁻¹. The solid black line through the data is the fit to a hyperbola \(K_{\text{obs}}: 12,200 \pm 6850 \text{s}^{-1}, k_{\text{dissoc}}: 805 \pm 470 \text{μM} \). Only data up to 75 μM are shown because the reaction was complete in the dead time of the stopped flow instrument at higher concentrations. C, 2D FitSpace calculation. The relationship between ground-state binding \(1/K_i\) and the conformational change step \(k_f\) is shown in terms of \(χ^2\) from the 2D FitSpace calculation in KinTek Explorer. The red area of the graph shows well-constrained rate constant combinations, whereas areas in green and blue are not well constrained. This experiment sets a lower limit of 6000 s⁻¹ for the rate of the conformational change at a 0.99 χ² confidence threshold and shows the linear relationship between \(k_f\) and \(1/K_i\). D, fluorescence measurement of dATP dissociation rate. T7 DNA polymerase E514Cou (500 nM), thioredoxin (10 μM), DNAad (600 nM), and dATP (500 nM) were mixed with wildtype T7 DNA polymerase (10 μM), thioredoxin (50 μM), and 27/45 DNA (with a 3'-OH group, 10 μM), which serves as a trap for nucleotide released from the ternary T7 DNA polymerase E514Cou-DNAad-dATP complex. The black line through the data is the best fit to a single exponential function, with an observed rate of 1.39 ± 0.002 s⁻¹. E, quench flow measurement of dATP dissociation rate. T7 DNA polymerase E514Cou (500 nM), thioredoxin (5 μM), DNAad (500 nM), BSA (0.1 mg/ml), and \([α-32P]-\text{dATP}\) (500 nM) were mixed with wildtype T7 DNA polymerase (5 μM), thioro- redoxin (20 μM), and 27/45 DNA (5 μM) to serve as a free nucleotide trap. The black line through the data is the best fit to a double exponential function, where the small initial fast phase corresponds to unbound \([α-32P]-\text{dATP}\) incorporated rapidly into the trap DNA and the second slower phase corresponding to the nucleotide dissociation rate has a rate of 3.0 ± 0.4 s⁻¹. F, equilibrium dATP binding titration. A solution of T7 DNA polymerase E514Cou (400 nM), thioredoxin (8 μM), DNAad (600 nM), and Mg²⁺ (12.5 mM) was titrated with dATP from a Hamilton syringe while monitoring fluorescence over the course of the 5-min titration with constant stirring. The signal was corrected for the small dilution and fit to a quadratic equation (black line) giving \(K_{\text{dissoc}}: 107 ± 1 \text{nM}\) and an enzyme concentration of 400 ± 1 nM.

Fast to measure in the stopped flow except at very low concentrations of nucleotide. To better estimate the rate of the conformational change, as well as other steps in the pathway that are too fast to measure accurately at 20 °C, we opted to perform all experiments in this paper at 4 °C.

**dATP binding kinetics**

The nucleotide binding rate was measured by mixing a preformed labeled E-DNAad complex with various concentrations of dATP (5–100 μM) and Mg²⁺ in the stopped flow (Fig. 2A). Although a good signal could be obtained by direct excitation of the coumarin label, a larger signal with the same kinetics was obtained when fluorescence emission from the coumarin was excited by FRET from tryptophan residues in the enzyme. An increase in fluorescence was observed and fit well to a single exponential function (Equation 3) at concentrations giving a measurable signal (up to 100 μM).

The exponential function was:

\[
y = A_0 + A_1 \cdot (1 - e^{-λ_1 t})
\]

where \(A_0\) is the starting fluorescence, \(A_1\) is the amplitude, \(λ_1\) is the observed decay rate, and \(t\) is time.

The plot of the observed rate at each dATP concentration (Fig. 2B) can be fit using a straight line (dashed gray line in Fig. 2B) or a hyperbola (Equation 4, solid black line in Fig. 2B). The model in Equation 2 predicts a hyperbolic concentration dependence of the observed rate.

\[
rate = \frac{k_{\text{max}} \cdot [\text{dATP}]}{K_{d+[\text{dATP}]} + [\text{dATP}]} = \frac{K_1 k_2 [\text{dATP}]}{1 + K_1 [\text{dATP}]} + K_2
\]

The slope of the linear fit was 14.5 ± 0.3 μM⁻¹ s⁻¹, which defines the second-order rate constant for nucleotide binding \((K_1k_2)\) leading to the \(FD_{dN}\) state (Equation 2). A slightly
better fit to the data was obtained assuming a rapid-equilibrium ground-state dATP binding ($K_i$) followed by a conformational change step ($k_2$). Conventional analysis of this data set using the two-step binding model reveals the large errors on the extrapolated maximum rate ($k_2 + k_{-2} = 12,200 \pm 6850 \text{ s}^{-1}$) and the $K_d = 1/K_f$ (805 ± 470 μM). The initial slope of the concentration dependence provides an accurate estimate for $K_f k_2$ for the two-step model, even though neither $K_f$ nor $k_2$ is well defined.

Fitting the primary experimental data globally using KinTek Explorer including confidence contour analysis (24) provided a well-defined lower limit on $K_2 \geq 6000 \text{ s}^{-1}$ as well as a well-defined value for $K_f k_2$ (Fig. 2C) shown by the linear relationship between $k_2$ and $1/K_f$ in the red line of the 2D confidence contour plot. With the equation-based data fitting shown here, the raw data were fit using a total of 17 parameters and the trace at the highest concentration was omitted because of the low amplitude and large errors. In contrast, with global fitting the full set of primary data was fit using only two parameters. Global data fitting provides a more realistic assessment of errors in fitting and gives lower limits for parameters that are not well constrained.

This is the first of many examples of the advantages of using global data fitting to reveal the information content of kinetic data. As experiments get more complex and interrelated, we emphasize the importance of interpreting experimental data globally using a single unifying model to overcome the traditional piecemeal approach of fitting using simplified equations for complex mechanistic studies. Although we favor the fit by simulation, here we initially present plots of rate versus concentration followed by quantitative analysis of the concentration dependence of the observed rate (when possible) to reveal the underlying mechanism. The equation-based data fitting provides a rationale to explain our interpretation of the results and provides the basis for developing a model that is then used to fit data globally.

**dATP dissociation rate experiments**

To more accurately measure the rate constant for nucleotide dissociation (limited by enzyme opening, $k_{-2}$), a stopped flow experiment was performed where a solution containing a ternary $E.DNAdd.dATP$ complex was mixed with a large excess of unlabeled wildtype $E.DNA$ complex to start the reaction. The large excess of wildtype $E.DNA$ complex (with a normal 3’OH primer) was used to trap dATP released from the labeled $E.DNAdd.dATP$ ternary complex. This mixture works efficiently as a trap for free dATP because the rate of nucleotide binding and incorporation by the large excess of wildtype enzyme is much faster than the slow release of the nucleotide from the labeled $E.DNAdd.dATP$ ternary complex. The decrease in fluorescence (Fig. 2D) fits a single exponential function with an observed decay rate of 1.39 ± 0.002 s⁻¹. Since the kinetics of nucleotide binding and incorporation are known for the wildtype enzyme (21), these parameters can be used when fitting by simulation to determine the nucleotide dissociation rate, even if the nucleotide trap is not 100% effective.

To ensure the dissociation rate measured in the stopped-flow experiment indeed reflected the nucleotide dissociation rate and not an artifact of the fluorescence signal, a similar experiment was performed using rapid-quench methods to measure the nucleotide dissociation rate (Fig. 2E).

An $E.DNAdd[\alpha^{32P}]dATP$ ternary complex was mixed with a large excess of wildtype $E.DNA$ complex to start the reaction. Dissociation of the $[\alpha^{32P}]-dATP$ from the $E.DNAdd[\alpha^{32P}]dATP$ complex was monitored by the decrease in $[\alpha^{32P}]-dATP$ concentration as it binds to the wildtype enzyme and is converted to DNA product. The reaction was quenched by mixing with EDTA from the quench syringe at various times, and samples were analyzed by TLC on polyethyleneimine (PEI)-cellulose plates to separate the free nucleotide from the nucleotide incorporated into DNA. The initial fast phase is due to the small amount of unbound nucleotide from the $E.DNAdd.dATP$ mixture that is rapidly incorporated by the wildtype polymerase upon mixing, while the slower phase reflects the nucleotide dissociation rate, which is limited by the reverse of the conformational change step ($k_{-2}$) from the labeled enzyme. The slower phase has a rate of 1.25 ± 0.15 s⁻¹, which is in agreement with the stopped flow dissociation rate measurement.

**Equilibrium dATP titration**

An equilibrium titration of the labeled $E.DNA_{dd}$ complex was performed with an increasing dATP concentration to get an independent measurement of the $K_d$ for dATP binding (Fig. 2F). A solution of labeled $E.DNA_{dd}$ was incubated in a cuvette inside a temperature-controlled cuvette in the titration module for the stopped flow. The titrant dATP was added to the cuvette over the course of 5 min with constant stirring at 4 °C while continuously monitoring the fluorescence (Fig. 2F). After correcting the fluorescence intensity for the small dilution during the titration, the resulting curve was fit using a quadratic equation (below):

$$y = F_0 + \Delta F \cdot \frac{[ED]_0 + [N] + K_d - \sqrt{([ED]_0 + [N] + K_d)^2 - 4 \cdot [ED]_0 \cdot [N]}}{2 \cdot [ED]_0},$$

where $F_0$ is the initial fluorescence, $\Delta F$ is the change in fluorescence upon saturation of dATP binding to the enzyme, $[ED]_0$ is the initial concentration of the labeled enzyme $E.DNA_{dd}$ complex, and $[N]$ is the concentration of dATP added during the titration. The titration defines the active enzyme concentration, 400 ± 0.2 nM and $K_d = 10^7 \pm 1$ nM.

The net $k_{d, net}$ determined from the rate constants for the fit to the on- and off-rate experiments for a two-step binding
model is defined by Equation 5.

\[ K_{d,net} = \frac{1}{K_1(1+K_2)} \approx \frac{1}{K_1 K_2} = \frac{k_{-2}}{K_1 k_2} = 96 \pm 2 \text{ nM} \]  

(5)

Because \( K_2 >> 1 \), the net \( K_{d,net} \) is simply the ratio of the measured dissociation rate \( (k_{-2}) \) divided by the observed second-order rate constant for nucleotide binding \( (K_1 k_2) \). The calculated value is close to the value determined from the direct equilibrium titration experiment, indicating that all kinetically significant steps have been included for nucleotide binding.

**Kinetics of polymerization**

Next, experiments were performed using the DNA 27/45 substrate containing a 3'-OH on the primer strand to allow measurement of the rates of the covalent bond formation between the 3'-OH group of the primer strand and the α-phosphate of the incoming dNTP. A pre-steady-state burst experiment was performed using rapid-quench-flow methods (Fig. 3A). A preformed E.DNA complex was mixed with varying concentrations of dATP (2.5–110 μM) and Mg\(^{2+}\) to start the reaction. At various times, the reaction was quenched by mixing with EDTA, and the products were resolved by denaturing PAGE. For conventional data analysis, data at each dATP concentration were fit to a burst equation (Equation 6):

\[ y = A_1 \cdot (1 - e^{-\lambda_1 t}) + k_{ss} t \]  

(6)

where \( A_1 \) is the amplitude of the exponential phase, \( \lambda_1 \) is the rate of the exponential phase, and \( k_{ss} \) is the slope of the linear phase. A hyperbolic fit of the concentration dependence of the rate of the exponential phase (Fig. 3B) provided estimates for the parameters from the maximum observed rate, \( k_{pol} = 210 \pm 15 \text{ s}^{-1} \), and half-maximal nucleotide concentration to define an apparent \( K_{d,app} = 15.5 \pm 2 \text{ μM} \). The ratio \( k_{pol}/K_{d,app} = 13 \pm 2 \text{ μM}^{-1} \text{ s}^{-1} \) provides an estimate of \( k_{cat}/K_m \) the specificity constant for the first incorporation.

Note that the \( K_{d,app} \) measured in this experiment is ∼150 times larger than the value reported from the equilibrium titration experiment \( (K_d \sim 100 \text{ nM}) \), reflecting the fact that this experiment gives a value that is not the true \( K_d \) for dATP binding. This feature is attributable to the two-step binding kinetics that do not reach equilibrium on the timescale of a single turnover (17). The meaning of \( k_{pol} \) and \( K_{d,app} \) with respect to individual rate constants cannot be determined without more information, as described below in global data fitting including multiple experiments.

**Time dependence of fluorescence changes during dATP incorporation**

The next experiment was designed to monitor the changes in fluorescence in the stopped flow to measure the conformational changes involving enzyme closing after nucleotide binding and reopening after chemistry. A labeled E.DNA complex was mixed with various concentrations of dATP (5–100 μM) and Mg\(^{2+}\) to start the reaction. As shown in Figure 3C, we observed an increase in fluorescence during dATP binding followed by a slower decrease in fluorescence back to the starting value.

At all concentrations, the traces can be fit using a double exponential function (Equation 7):

\[ y = A_0 + A_1 \cdot (1 - e^{-\lambda_1 t}) + A_2 \cdot (1 - e^{-\lambda_2 t}) \]  

(7)

where \( A_0 \) is the initial fluorescence, whereas \( A_1 \) and \( \lambda_1 \), and \( A_2 \) and \( \lambda_2 \) are the amplitudes and rates of the fast and slow phases, respectively.

The dATP concentration dependence of the rates of the fast and slow phases (Fig. 3D) shows an almost linear increase in the rate of the fast phase with increasing dATP concentrations and a constant observed rate for the slower phase at 126 ± 4 s\(^{-1}\). Fitting the concentration dependence of the fast phase to a hyperbola in an attempt to extract the maximum rate of the conformational change \( (k_2) \) has large errors (7800 ± 3400 s\(^{-1}\)) but is comparable with the values obtained for the nucleotide binding-rate experiment with dNDA\(^{dd}\) in Figure 2A.

The increase in fluorescence is a function of the transition from open to closed enzyme \( (ED_n \rightarrow FD_nN) \), whereas the subsequent decrease is due to the closed to open transition \( (FD_{n+1},PP_i \rightarrow ED_{n+1}) \) after the chemical reaction (Equation 1). We know that the increase in fluorescence is correlated with enzyme structural changes seen after nucleotide binding to deoxyterminated DNA (Fig. 2). Moreover, global data fitting in correlating the fluorescence transients with the measured rates of the chemical reaction establish the pathway. Since \( k_{pol} \) measured in the quench flow burst experiment is approximately 200 s\(^{-1}\), the slower phase of the fluorescence transient (approximately 125 s\(^{-1}\)) must be attributed to an isomerization step \( (k_d) \) following chemistry that may be coincident with PP\(_i\) release (Equation 1).

Globally fitting these data along with nucleotide binding kinetics can afford estimates of \( k_3 \), \( k_{3,b} \), and \( k_d \) from the concentration dependence of the rate and amplitude of the pre-steady-state burst of incorporation (Fig. 3A) and the fluorescence transients (Fig. 3C). Simultaneous fitting of rates and amplitudes of the reactions allows resolution of \( k_3 \), \( k_{3,b} \), and \( k_d \), which we will show after all experiments have been included. Understanding global data fitting does not translate easily to equation-based data fitting where the rates and amplitudes are derived independently and using five independent parameters per trace lead to increased errors. Global data fitting affords direct estimates of the intrinsic rate constants while accounting for rates and amplitudes simultaneously. Moreover, the most accurate estimates are based on the inclusion of data from pyrophosphorolysis. Therefore, we postpone showing the global data fitting until all data have been presented.

The model invoking slow PP\(_i\) release following fast chemistry raises two testable hypotheses: 1) the slow PP\(_i\) release may limit the rate of polymerization during multiple nucleotide incorporations, which can be measured directly; 2) the reverse rate of chemistry \( (k_{3,b}) \) can be extracted from analysis of the pyrophosphorolysis reaction to define the internal equilibrium
Conformational dynamics for T7 DNA polymerase

constant for chemistry, $K_3$. These hypotheses will be tested by pyrophosphorolysis experiments in the following section.

**Kinetics of pyrophosphorolysis**

Initial attempts to measure pyrophosphorolysis using 6-carboxyfluorescein labeled DNA (FAM-DNA) resolved on DNA sequencing gels (not shown) produced many bands, both above and below the starting 27-nt primer, making the experiment difficult to interpret. Owing to the high efficiency of polymerization, any dNTP produced during the slower pyrophosphorolysis reaction becomes a substrate for the primer extension reaction. To circumvent the difficulties with this experimental approach, we used a previously reported strategy (14) that required a DNA substrate containing [$\alpha$-$^{32}$P]-dNMP incorporated enzymatically at the 3’ end of the primer strand of DNA 27/45 (named [$\alpha$-$^{32}$P]-28/45) as described in the experimental procedures section. This afforded direct measurement of a single round of pyrophosphorolysis.

A preformed E-[α-$^{32}$P]-28/45 DNA complex with varying concentrations of PP$_i$ (1–12 mM) was mixed with Mg$^{2+}$ and unlabeled dATP (20 μM) to start the reaction. Reactions were quenched at various times by adding 0.3 M EDTA, and reaction products were analyzed by TLC on PEI-cellulose plates to resolve the formation of the product [$\alpha$-$^{32}$P]-dATP (Fig. 4A).

Unlabeled dATP served as a trap to prevent the rebinding of [α-$^{32}$P]-dATP and to fill in any primer where the radiolabeled dAMP has been removed. This prevents multiple rounds of pyrophosphorolysis, which could produce dNTPs that would then serve as substrates for correct incorporation on top of the radiolabeled 3’ dAMP of the primer strand. The low

![Figure 3. Kinetics of dATP incorporation.](image)

**Figure 3. Kinetics of dATP incorporation.** A, quench flow pre-steady-state burst experiment. T7 DNA polymerase E514Cou (120 nM), BSA (0.1 mg/ml), thioredoxin (2.4 μM), and FAM-27/45 DNA (200 nM) were mixed with dATP (2.5–110 μM) and Mg$^{2+}$ (12.5 mM) to start the reaction. Data at various dATP concentrations are shown as different colors, fit to a burst equation (black line) at each concentration. B, rate versus dATP concentration. Rates are from the fit of the exponential phase in (A) and are shown fit to a hyperbola (black line), estimating the maximum rate, $k_{\text{max}} = 210 \pm 15$ s$^{-1}$ and $K_{\text{d,app}} = 15.5 \pm 2$ μM. C, stopped flow dATP incorporation. T7 DNA polymerase E514Cou (750 nM), thioredoxin (15 μM), and DNA 27/45 (1 μM) were mixed with dATP (5–100 μM) and Mg$^{2+}$ (12.5 mM) to start the reaction. Data at each dATP concentration were fit to a double exponential function (black lines). D, rate versus dATP concentration for stopped flow fluorescence. Rates are from the data in (C). The fast phase data (red points, $k_{\text{obs}} 1$) were fit to a hyperbola to estimate the maximum rate of the conformational change ($k_2$), $7800 \pm 3400$ s$^{-1}$. The slow phase data (green points, $k_{\text{obs}} 2$) were fit to a line, setting a lower limit on the maximum rate of the isomerization following chemistry ($k_4$) of $126 \pm 4$ s$^{-1}$.

6 J. Biol. Chem. (2021) 296 100143
concentration of dATP used in the experiment was chosen to preclude incorporation of a dATP:dA mismatch on top of the \([\alpha^{32}\text{P}]-\text{labeled primer strand}\) but is sufficient to rapidly replace the \([\alpha^{32}\text{P}]-\text{dAMP}\) after its removal from the primer. These details allow the experiment to be interpreted without complications. In addition, although Mg\(^{2+}\)-PPi precipitates at concentrations above \(~5\) mM, the precipitation occurs on the timescale of approximately 30 s, whereas the pyrophosphorylation reactions at concentrations of Mg\(^{2+}\)-PPi \(\geq 5\) mM occurs on a shorter timescale. Therefore, measurements can be made before precipitation by mixing PPi (without Mg\(^{2+}\)) and the E-[\(\alpha^{32}\text{P}]-28/45\) complex with Mg\(^{2+}\) to start the reaction. A control stopped flow experiment showed PPi does not bind to the enzyme in the absence of Mg\(^{2+}\) (not shown).

The time dependence of forming \([\alpha^{32}\text{P}]-\text{dATP}\) (Fig. 4A) at each PPi concentration was adequately fit using a single exponential function (Equation 3). The PPi concentration dependence of the observed rate (Fig. 4B) was fit to a hyperbola to give an apparent \(K_{d,\text{app}} = 5.6 \pm 1.2\) mM for PPi, and a maximum rate of pyrophosphorylation, \(k_{PPi} = 0.2 \pm 0.02\) s\(^{-1}\). The weak apparent binding affinity and slow rate of pyrophosphorylation indicate that the reverse reaction is highly unfavorable relative to the forward dATP incorporation reaction. However, these parameters underestimate the efficiency of pyrophosphorylation owing to additional aspects of the reaction, such as translocation, which can be resolved by global data fitting including additional experiments.

**Figure 4. Kinetics of pyrophosphorylation.** A, hand quench pyrophosphorylation with \(3'-[\alpha^{32}\text{P}]-\text{labeled DNA}\). A solution of T7 DNA polymerase E514C (500 nM), thioredoxin (10 \(\mu\)M), BSA (0.1 mg/ml), \(3'-[\alpha^{32}\text{P}]-28/45\) DNA (100 nM), and PPi (1–12 mM) was mixed with unlabeled dATP (20 \(\mu\)M) and Mg\(^{2+}\) (12.5 mM) to start the reaction. Data at each PPi concentration are shown fit to a single exponential function (black lines). B, rate versus PPi concentration for pyrophosphorylation experiment. Rates are from the data in (A). Data are shown fit to a hyperbola (black line), estimating an apparent \(K_{d,\text{app}} = 5.6 \pm 1.2\) mM and \(k_{PPi} = 0.20 \pm 0.02\) s\(^{-1}\). C, stopped flow pyrophosphorylation T7 DNA polymerase E514C (500 nM), thioredoxin (10 \(\mu\)M), 28/45 DNA (600 nM), and PPi (1–8 mM) were mixed with Mg\(^{2+}\) (12.5 mM) to start the reaction. Data at each PPi concentration were fit to a single exponential function (black lines). D, rate and amplitude versus PPi concentration for stopped flow pyrophosphorylation. Rates and amplitudes are from data in (C). Data are shown fit to a hyperbola (black line) estimating an apparent \(K_{d,\text{app}} = 950 \pm 390\) \(\mu\)M. Inset: Amplitude versus PPi concentration. Shown fit to a hyperbola (black line), estimating an apparent \(K_{d} = 2.2 \pm 0.3\) mM. Increasing amplitude with increasing PPi concentration is indicative of a "conformational selection" mechanism.
Conformational dynamics for T7 DNA polymerase

We reasoned that, since there is a fluorescence signal for the forward reaction, there should also be a fluorescence signal associated with PPi binding following the \( FD_{e+1,PP} \) states that can be followed by stopped-flow methods (Equation 1). We performed this experiment (Fig. 4C) by mixing a solution containing the \( ED \) complex and varying PPi concentrations (1–8 mM) with \( Mg^{2+} \) to start the reaction. Data collection was limited to 250 ms to restrict the observed portion of the reaction to PPi binding. On a longer timescale, a subsequent forward reaction after dATP is produced leads to a more complex signal.

The stopped flow traces at all PPi concentrations fit a single exponential function (Equation 3). Fitting the rate and amplitude versus PPi concentration using a hyperbola is shown in Figure 4D. Of interest, the observed rate of the fluorescence transient decreased with increasing PPi concentrations with an apparent \( K_d \) of 950 ± 390 \( \mu \text{M} \). However, the amplitude increased. This type of behavior is the kinetic signature for a “conformational selection” mechanism (25, 26), where a rate-limiting isomerization of the enzyme occurs before binding of the substrate (PPi; in this case). In particular, it is important that attempts to fit the data by simulation with all experiments up to this point using the model in Equation 1 were unsuccessful because they failed to provide a model containing two enzyme states prior to PPi binding (i.e., translocation).

A reasonable fit to the data was obtained by adding a translocation step (\( ED' \rightarrow ED \) or \( EP' \rightarrow EP \)) after PPi release but before binding of the next dNTP (Fig. 5).

As the translocation step is the only known step in the polymerization cycle that is not included in the model in Equation 1, we hypothesized that translocation could modulate the binding of dNTP or PPi, as first described in prior work on the T7 DNA polymerase (11, 14). The concept of binding of PPi or dNTP to a pre-existing equilibrium between pre- and posttranslocation states, respectively, was later given the name of a “Brownian ratchet” (27). In this model, the release of the PPi product is separated from dATP binding by a reversible translocation step. After translocation, ATP binding traps the DNA at the active site. This contrasts with the “power-stroke” model where PPi and dNTP bind to the same enzyme state, and subsequent to dNTP binding translocation and PPi release occur simultaneously (28).

Our updated model including translocation is shown in Figure 5, where \( k_5 \) is the rate constant governing translocation from the pretranslocated states (\( EP', ED' \)) that are able to bind PPi) to the posttranslocated states (\( ED, EP \) that are able to bind the next incoming nucleotide). Because the equilibrium favors the translocated state, dATP binds to the predominant state in the population and the subsequent conformational change follows an induced-fit model. In contrast, PPi binds only to the minor (untranslocated) state following a pattern of conformational selection.

With this more complete model, extrapolation of the observed rate versus concentration plot to zero [PPi] (Fig. 4D) provides an estimate for the sum \( k_{-5} + k_5 \). Although there are large errors on this y-intercept because of the decreasing amplitude with decreasing concentrations of PPi, the data are sufficient to set a lower limit of \( \sim 100 \text{ s}^{-1} \) for the sum of these two rate constants (\( k_{-5} + k_5 \)). The rate constant for reverse translocation (\( k_{-5} \)) is better defined because it is obtained using data at higher PPi concentrations, extrapolated to give \( k_{-5} \sim 10 \text{ s}^{-1} \).

The experiments in the next section were designed to more rigorously distinguish the model in Equation 1 from the model in Figure 5.

Inhibition of dATP incorporation by PPi

The model in Figure 5 has some important and testable implications that distinguish it from the power stroke model where the nucleotide and PPi bind to the same enzyme state. The following experiments were designed using KinTek Explorer with the two different models to determine the best reaction conditions to differentiate the two kinetic schemes. The first experiment involves performing the forward nucleotide incorporation reaction in the presence of varying concentrations of PPi.

Stopped flow PPi inhibition on dATP incorporation

With the model in Figure 5, inhibition by rebinding PPi directly to the EP' state formed after chemistry should appear to be much tighter than the affinity derived from the experiment in Figure 4A where PPi binding must overcome the unfavorable translocation equilibrium. To test this hypothesis, we monitored the stopped-flow fluorescence signal observed at a fixed dATP concentration and varying concentrations of PPi. In this experiment a preformed \( ED \) complex with varying concentrations of PPi (0–8 mM) was mixed with \( Mg^{2+} \) and a fixed concentration of dATP (50 \( \mu \text{M} \)) in the stopped flow to start the reaction (Fig. 6A). Traces at each PPi concentration were fit to a double exponential function, and the observed rates for the fast and slow phases were plotted versus PPi concentration (Fig. 6B). It is surprising that the rate of the fast phase, corresponding to dATP binding steps, showed only a slight inhibition with increasing PPi concentrations (step 6 in Fig. 5). This inhibition of the fast phase fits a linear function

\[
\begin{align*}
\text{ED', PPi} & \xrightarrow{k_5} \text{EP', PPi} \\
& \xrightarrow{k_6} \text{EP, PPi}
\end{align*}
\]

Figure 5. Full kinetic model for nucleotide incorporation by T7 DNA polymerase. This pathway is expanded from Equation 1 to Include translocation steps (\( k_5, k_3 \)) and competitive PPi binding to the post-translocated state (\( k_5 \)).
with a slope of $-0.027 \pm 0.003 \text{ mM}^{-1} \text{s}^{-1}$. The concentration dependence on the rate of the isomerization accompanying PP$_i$ release following chemistry showed an apparent $K_d = 775 \pm 75 \text{ μM}$, which is comparable with that obtained with the stopped flow pyrophosphorolysis experiment in Figure 4C, consistent with the translocation model shown in Figure 5.

Fitting this experiment globally in KinTek Explorer with all experiments presented up to this point also required an additional rapid equilibrium PP$_i$ inhibition step ($K_0$), competitive with dNTP, that binds to the posttranslocated state (ED, EP) with a $K_i$ of approximately 6 mM (Figure 5). This accounts for the small linear inhibition of the rate of the fast phase, in addition to the translocation step proposed earlier. Addition of this competitive inhibition step did not significantly change the rate constants for nucleotide binding ($K_1, K_2, K_3$), chemistry ($k_3$), or the reverse of translocation ($k_{-3}$); however, the reverse of chemistry ($k_{-3}$) was increased $\sim$3 fold to 136 s$^{-1}$ and an increase in $k_5$ to $\sim$200 s$^{-1}$ to achieve an optimal global fit of the data, which is described below.

**Quench flow measurement of PP$_i$ inhibition of dATP incorporation**

A parallel experiment was performed in the quench flow by mixing a preformed E.DNA complex and varying concentrations of PP$_i$ (0–8 mM) with Mg$^{2+}$ and dATP (20 μM) to start

![Figure 6. Kinetics of PP$_i$ inhibition on dATP incorporation.](image)
the reaction (Fig. 6C). The data at each PP₈ concentration were fit to a burst equation (Equation 6), and the observed rate of the exponential phase versus PP₈ concentration is shown fit to a hyperbola in Figure 6D. Minimal inhibition is observed as the majority of the enzyme species is in the state ready to bind dATP (ED) with only weak competitive PP₈ inhibition, not the form (ED') that tightly binds PP₈. PP₈ does not significantly inhibit steps until after the product DNA has been formed and is committed to going forward to form product in steps that are not observed in this experiment.

This experiment was performed in the original series of papers on this enzyme (14) and is consistent with the work presented here where PP₈ acts as a weak competitive inhibitor for nucleotide binding with an observed effect on the rate of the burst. This experiment does not provide information to justify one translocation model over the other. Rather, it is included for consistency with previously published work and shows the additional information obtained from using a labeled protein to dissecting the reaction to give the complete story.

**Kinetics of processive polymerization**

The model proposed in Figure 5 has important implications for processive polymerization. The first prediction, mentioned earlier, is that incorporation of nucleotides after the first incorporation may be slower than the first incorporation owing to rate limiting PP₈ release and translocation in the reaction pathway. The second prediction is that PP₈ will inhibit the second nucleotide incorporation in a series to a much greater extent than the first incorporation. After the first incorporation, the ED complex will pass directly through the state that is able to bind PP₈ (ED₂ in Equation 8) with relatively high affinity. These two predictions are tested with the following experiment.

\[
ED_{27}^{+dATP} \rightarrow FD_{28} \rightarrow PP_{i} \rightarrow ED_{28}^{trans} \rightarrow ED_{28} \\
ED_{28}^{+dTTP} \rightarrow FD_{29} \rightarrow PP_{i} \rightarrow ED_{29}^{trans} \rightarrow ED_{29}
\]

(Equation 8)

**Processive polymerization and inhibition by PP₈**

To test these two hypotheses, we performed a rapid quench experiment where an EDNA complex, in the presence or absence of 8 mM PP₈, was mixed with dATP (100 µM), dTTP (100 µM), and Mg²⁺ to start the reaction (Fig. 7). Fitting the data for the experiment without PP₈ (closed circles) gives net rates of 182 ± 10 and 41 ± 1 s⁻¹ for the first and second incorporations, respectively, based on a simple two-step reaction: ED₇ → ED₈ → ED₉. The second incorporation is much slower than the first incorporation, owing to the rate-limiting PP₈ release and translocation steps that occur after fast chemistry for the first incorporation (Equation 8). The experiment performed in the presence of 8 mM PP₈ (open circles) shows some inhibition of the rate of the first incorporation, with a rate of 104 ± 8 s⁻¹, whereas the rate of the second incorporation, 9.3 ± 0.6 s⁻¹, was significantly reduced. For the simple model in Equation 1, even at 8 mM, one would expect only slight inhibition on the second incorporation (less than 2-fold). The large inhibition of the second incorporation reaction provides further support for the model in Figure 5 and Equation 8.

**Global fitting of data and free-energy profile**

All experiments were globally fit in KinTek Explorer (29, 30) using the unified model shown in Figure 5. The rate constants for each step in the pathway derived from the global fitting are given in the scheme at the top of Figure 8, and the error estimates are listed in Table 1. The global fit for all experiments is shown as the solid-colored lines in Figure 8. The confidence contour FitSpace analysis (24) was used to provide the most rigorous statistical test to determine whether each parameter in the model was well constrained by the experimental data. The value of $\chi^2$ is plotted versus the systematically varied value for each rate constant as shown in Figure 9. Of importance, the error profile for each rate constant is symmetric showing that each is well constrained by the data.

The new standard for finding a minimal model is based on defining a pathway that is necessary and sufficient to account for all available data. Achieving a good fit establishes that the minimal model is sufficient to account for the data. Showing that all kinetic parameters are well constrained by the data establishes that the model is not overly complex (30).

Table 1 shows the best-fit values for each rate constant along with upper and lower limits derived from the confidence

![Figure 7](image-url)
Figure 8. Global fit of all experimental data. The scheme at the top contains the rate constants derived from the global fit of the data shown in this figure. Error estimates for each rate constant are given in Table 1. Rate constants in red were locked during the fitting, and the numbers in parenthesis above the steps are the step numbers. The experiments in Figures 2–4, 6 and 7 are shown with solid colored lines through the data showing the global fit to the rate constants in the scheme at the top of the figure. A, stopped flow dATP binding rate. B, stopped flow dATP dissociation rate. C, equilibrium dATP binding titration. D, stopped flow fluorescence during dATP incorporation. E, quench flow measurement of dATP incorporation. F, quench flow measurement of PPi inhibition of processive polymerization. G, stopped flow measurement during PPi inhibition of dATP incorporation. H, quench flow measurement of PPi inhibition of PPi incorporation. I, kinetics of pyrophosphorolysis. J, stopped flow pyrophosphorolysis.
Contour analysis with a $\chi^2$ threshold of 0.99. From the mechanism and rate constants in Figure 8 and physiological dATP and PPi concentrations, a free-energy profile was created in KinTek Explorer (Fig. 10). Most importantly, the specificity-determining step is the highest barrier relative to the starting material, which is the conformational change step. There is no single rate-determining step because $k_3$, $k_4$, and $k_5$ are comparable.

The complete solution to derive steady-state parameters from intrinsic rate constants for the model shown in Figure 5 is:

$$k_{\text{cat}} = \frac{k_3 k_4 k_5}{k_{-3} (k_2 + k_{-2} + k_6 (k_2 + k_{-2} + k_3) + k_2 k_5 (k_4 + k_5))}$$

$$\frac{k_{\text{cat}}}{K_m} = \frac{k_1 k_2 k_5}{(k_{-3} k_{-2} + k_6 (k_{-3} + k_4) + k_2 k_4 (k_{-1} + k_2))}$$

Using these equations and the rate constants in Table 1, we calculate $k_{\text{cat}} = 75 \text{ s}^{-1}$ and $k_{\text{cat}}/K_m = 16.5 \mu\text{M}^{-1} \text{ s}^{-1}$ for processive synthesis, partially rate limited by PPi release and translocation. Based on a rapid equilibrium nucleotide binding model the specificity constant can be approximated as:

$$K_{\text{cat}} / K_m = K_1 k_2 k_5 / (k_5 + k_{-5})$$

The conformational change after weak ground-state nucleotide binding is the primary determinant of specificity via this induced-fit mechanism but is attenuated by the fraction of ED complex that is in the translocated state (95%).

### Table 1
| Parameter | Best-fit value | 95% Confidence interval |
|-----------|---------------|------------------------|
| $K_1 k_2$ | 18 ± 0.6 $\mu\text{M}^{-1} \text{ s}^{-1}$ | 17.4–18.6 |
| $1/K_1$  | 360 ± 12 $\mu\text{M}$ | 351–377 |
| $k_2$    | 6500 $\text{s}^{-1}$ | ≥6000 |
| $k_3$    | 1.7 ± 0.04 $\text{s}^{-1}$ | 1.67–1.74 |
| $k_4$    | 310 ± 40 $\text{s}^{-1}$ | 273–359 |
| $k_{-3}$ | 140 ± 30 $\text{s}^{-1}$ | 115–170 |
| $k_5$    | 220 ± 18 $\text{s}^{-1}$ | 210–246 |
| $k_{-4}$ | 0.22 ± 0.03 $\mu\text{M}^{-1} \text{ s}^{-1}$ | 0.201–0.261 |
| $k_6$    | 190 ± 25 $\text{s}^{-1}$ | 171–222 |
| $k_{-6}$ | 10.7 ± 0.6 $\text{s}^{-1}$ | 10.1–11.3 |
| $K_6$    | 6.1 ± 0.3 $\text{mM}$ | 5.74–6.43 |

Lower and upper limits are reported from a 0.99 $\chi^2$ threshold to give a 95% confidence interval. The value of $k_2$ (in brackets) was locked during fitting, which determines the value for $K_1$. For each of the two equilibrium constants ($K_1$ and $K_6$), the binding rate constant was locked and fitting was used to define the dissociation rate constant was used to estimate the equilibrium constant. For example, $k_1$ was locked at 100 $\mu\text{M}^{-1} \text{ s}^{-1}$ and $K_1 = k_1/k_{-1}$ was calculated from estimates for $k_{-1}$ derived in fitting. Similarly, $K_6 = k_6/k_{-6}$ with $k_6$ locked in fitting.

![Figure 9. Confidence contour plots for rate constants in the dATP incorporation pathway.](image-url)
The apparent $K_{d,PPi}$ from the rate versus concentration plot in Figure 4D is $950 \pm 390 \mu M$, which is defined in terms of rate constants as $K_{d,PPi} = k_d(k_5 + k_{-5})/k_{-d}k_5$, which approximates the true $K_d$ for PPi binding to the pretranslocated state. Given estimates for $k_d$ (125 s$^{-1}$ from Fig. 3), $k_{-d}$, or the second-order rate constant for PPi binding to the pretranslocated state, can be estimated around 0.2 $\mu M^{-1}$ s$^{-1}$. When compared with the value for $k_{-d}$ in fitting to Equation 1 ($\sim$0.02 $\mu M^{-1}$ s$^{-1}$), with unchanged $k_d$, this model predicts PPi binding $\sim$10-fold tighter to the enzyme than previously measured from the hand quench experiment owing to pre-equilibration of the enzyme into the posttranslocated state (EP) that is unable to bind PPi that must be overcome before the reverse reaction can proceed.

We show that, for our fluorescent DNA polymerase (21), initial weak ground-state nucleotide binding ($1/K_i \geq 360 \mu M$) induces a conformational change ($k_{-d} \geq 6000$ s$^{-1}$) that is much faster than the rate of chemistry ($k_3 = 310$ s$^{-1}$). Although neither $K_i$ nor $k_3$ is well defined, the apparent second-order rate constant for nucleotide binding is well defined as $K_i/k_2 = \pm 0.6$ $\mu M^{-1}$ s$^{-1}$ (Fig. 2C). Of importance, the rate constant for reverse of the conformational change step allowing nucleotide release ($k_{-2} > 1.7$ s$^{-1}$) is much slower than chemistry (Fig. 2D). Thus, kinetic partitioning of the FDN state to either release nucleotide or form product for the correct nucleotide greatly favors the forward reaction. Specificity is a function of all steps leading up to the first largely irreversible step, which can be identified as the highest overall barrier in the free-energy profile (Fig. 10). Accordingly, the specificity constant ($k_{cat}/K_m$) is determined solely by the net second-order rate constant for binding defined by product of the equilibrium constant for nucleotide binding to the open state times the rate of the conformational change (1, 17).

Prior studies on an MDCC-labeled 8-Cys lite T7 DNA polymerase variant (31) provided an estimate of the rate of the conformational change of 700 s$^{-1}$ at 20°C, a similar nucleotide dissociation rate of 2 s$^{-1}$, a similar rate of chemistry around 300 s$^{-1}$, but a higher-affinity ground-state binding at 28 $\mu M$ (1). The extensive mutagenesis required for site-specific cysteine labeling decreased the fidelity of this enzyme variant by greatly increasing the ground-state binding affinity of mismatches. In comparison, a single unnatural amino acid variant of T7 DNA polymerase maintains the fidelity of the wildtype enzyme (21). Although the ground-state binding affinity is approximately 10-fold weaker for our variant, the rate of the conformational change is at least 10 times faster, in spite of the lower temperature in this study. One study using microfluidics provided an estimate of the substrate-induced conformational change of 4700 s$^{-1}$ but the effect of the Cy3 label in this study was not examined (32).

In studies on HIVRT, another family A polymerase, a fluorescently labeled variant (17) showed similar rate constants with weak ground-state binding (200 $\mu M$), a fast conformational change (2000 s$^{-1}$), but a slower nucleotide dissociation rate (0.06 s$^{-1}$) and chemistry (15 s$^{-1}$). In both HIVRT and T7 DNA polymerase, the conformational change is the specificity-
Conformational dynamics for T7 DNA polymerase

determining step for correct nucleotide incorporation, as it is the highest barrier relative to the starting material (Fig. 10).

Studies on polymerase β (33) also revealed a nucleotide-induced conformational change at 1400 s\(^{-1}\) but a much slower rate of chemistry (6.4 s\(^{-1}\)) and faster nucleotide dissociation rate (12 s\(^{-1}\)). For this low-fidelity repair enzyme, chemistry is the predominant specificity-determining step as it represents the highest overall barrier, but \(k_{\text{cat}}/K_{m}\) is also affected by the two-step rapid-equilibrium nucleotide binding. Because \(k_{\text{cat}} > k_{b}\), the specificity constant is approximately defined by \(k_{\text{cat}}/K_{m} \approx K_{p}/K_{F}\).

In this report a single unnatural coumarin amino acid afforded an optimal signal to monitor conformational dynamics during DNA polymerization. The rise and fall of the fluorescence transient during a single round of nucleotide incorporation (Fig. 3C) revealed the closing of the enzyme after binding nucleotide followed by chemistry (measured by rapid quench in Fig. 3A) then reopening to release the PP\(_{i}\) product. Globally fitting fluorescence transients with rapid-chemical quench flow data allows resolution of each step in the pathway by correlating conformational changes with measurable chemical reactions. Under our experimental conditions at 4 °C, we showed that enzyme reopening (and presumably PP\(_{i}\) release) occurred at a rate that is slower than the chemistry and therefore limits the rate of processive DNA replication. This conclusion was confirmed by measurements of the effect of added pyrophosphate on processive synthesis (Fig. 7).

Similar studies on DNA polymerase β measuring the fluorescence transient gave a rate of PP\(_{i}\) release that was coincident with chemistry, signifying that the chemistry was the rate-limiting step in the pathway and PP\(_{i}\) release is fast (33). Similarly, when HIVRT is copying a DNA template, the kinetics of the chemical reaction, enzyme opening (measured by fluorescence), and PP\(_{i}\) release are all coincident, suggesting that the rate-limiting chemistry is followed by rapid enzyme opening and PP\(_{i}\) release (17, 34). However, for an RNA template, chemistry is fast (250 s\(^{-1}\)) and PP\(_{i}\) release is slow (58 s\(^{-1}\)), which limits the rate of processive synthesis (35), as seen for T7 DNA polymerase at 4 °C.

To derive the most complete model we also performed experiments on the reverse (pyrophosphorolysis) reaction using both chemical quench methods as well as stopped flow methods to monitor conformational change using our labeled T7 DNA polymerase variant. Interestingly, when fluorescence changes were measured during pyrophosphorolysis, the traces showed an increasing amplitude but decreasing rate with increasing PP\(_{i}\) concentrations (Fig. 4C), indicative of a “conformational selection” mechanism. We showed that these results could be explained by an unfavorable reverse translocation step required before PP\(_{i}\) could bind to the less populated ED\(^{\prime}\) state. These experiments help to define the kinetics of the translocation step from ED\(^{\prime}\) to ED with an equilibrium constant of \(K_{S} = 18\) in the forward direction.

In this model, PP\(_{i}\) release occurs with enzyme opening to form ED\(^{\prime}\) and is followed by translocation. Similar ensemble measurements could potentially be performed on other labeled DNA polymerases to extract translocation rates. This sequential reaction is counter to the predictions of the power stroke model. Initial structural studies suggested a power stroke model for T7 RNA polymerase (36), but direct measurement of the kinetics of translocation by ensemble methods has proven difficult.

Of importance, our new data define the rate constants for translocation during DNA polymerization. Prior measurements of translocation have relied on single molecule methods, using optical tweezer experiments or experiments using an α-hemolysis nanopore (22, 23) at various forces or potential differences, respectively. These data suggest a rapid equilibrium constant for translocation with rate constants on the order of 1000 s\(^{-1}\) and an equilibrium constant near unity. This represents an example of a Brownian ratchet in which translocation is driven by selective binding to the posttranslocation state, which is in rapid equilibrium with the pretranslocation state. In contrast, our data present a different view in which translocation is favored thermodynamically (\(K_{\alpha} = 18\), \(\Delta G = -6.7 \text{ kJ/mol at 4 °C}\)) but is also rate limiting during processive polymerization. Thus, measured rates of translocation during polymerization do not support either the Brownian ratchet or the power stroke model, each of which invokes nucleotide binding to drive translocation. Although nucleotide binding to the posttranslocation state could drive the translocation reaction to completion (14), we now know that the translocation is not in rapid equilibrium. Rather, the translocation proceeds in the absence of nucleotide to 95% completion to facilitate rapid nucleotide binding to the post-translocation state. In our experiments performed at 4 °C, translocation was rate limiting, allowing direct measurement of the rate. Experiments remain to be performed to see how the rate constants change at higher temperatures, where the rate of translocation is expected to increase.

All experiments were performed at 4 °C, which is a lower temperature than optimal, but certainly within the physiological range for Escherichia coli. Work on other polymerases has shown that fidelity does not seem to change much as a function of temperature (37). However, the rates of the conformational change for HIVRT showed a significant temperature dependence, which afforded accurate estimates of the rate of the conformational change by extrapolation to 37 °C (17, 38). Differences in temperature dependence are likely to reflect the physiologically important range of temperatures for a particular enzyme. While HIVRT evolved to infect human cells at 37 °C, presumably T7 bacteriophage must have evolved to infect E. coli over a wide range of temperatures. The rate of nucleotide incorporation is too fast to measure accurately at 20 °C, but the temperature dependence of the experiments presented here would allow fast rates to be better estimated by extrapolation to higher temperatures. Such measurements would also provide valuable enthalpy and activation energies for each step in the pathway.
The data reported here provide the first complete pathway and free-energy profile for a correct nucleotide incorporation for a high-fidelity DNA polymerase. This work extends our understanding of the induced-fit mechanism described for the HIVRT DNA polymerase, an enzyme with moderate fidelity. For the full understanding of DNA polymerase fidelity, the misincorporation and exonuclease proofreading reactions must also be quantified. Similar studies on the misincorporation reaction with our labeled T7 DNA polymerase are complex and will be the subject of a future paper.

**Experimental procedures**

**Enzyme labeling and purification**

We site specifically labeled and purified T7 DNA polymerase with a fluorescent unnatural amino acid, (7-hydroxy-4-coumarin-yl) ethylglycine as described (21). In this report we showed that the labeled enzyme retained activity indistinguishable from that of the wildtype enzyme.

**Preparation of oligonucleotides for kinetics assays**

Synthetic oligonucleotides, including oligos containing 2',3'-dideoxy or 5'-[6-FAM] modifications, were custom synthesized by Integrated DNA Technologies with standard desalting and were further purified in house by denaturing PAGE to >99% full length oligo. Purified oligos were stored in 66.2 buffer (6 mM Tris-HCl pH 7.5, 6 mM NaCl, 0.2 mM EDTA) at −20 °C. Concentrations of oligos were determined by absorbance at 260 nm, using extinction coefficients calculated from the sequence of the oligonucleotide. Extinction coefficients and sequences for each oligo used in this study are listed in **Table 2**. Primer and template DNA strands were mixed in a 1:1.05 ratio in annealing buffer (10 mM Tris-HCl pH 7.5, 50 mM NaCl, 1 mM EDTA), heated to 95 °C, and cooled slowly to room temperature over approximately 2 h. Double stranded primer/template DNA substrates and their short names used in this paper are shown in **Table 3**. A control pre-steady-state burst experiment was performed using 32P-labeled DNA and FAM-DNA to check for artifacts due to the fluorescent label (not shown). No difference was detected, so the FAM-DNA substrate was used for convenience.

**Preparation of 3'-α-[32P]-dAMP labeled 28/45 DNA substrate**

The following components were mixed together in T7 Reaction Buffer at 4 °C for 5 min: 50 nM wildtype T7 DNA polymerase, 1 μM thioredoxin, 0.1 mg/ml bovine serum albumin (BSA), 1 μM DNA 27/45, 130 nM [α-32P]-dATP (Perkin Elmer), and 12.5 mM Mg2+. Unlabeled dATP was then added to 10 μM, and the reaction was incubated at room temperature for another 12 min before quenching with the addition of EDTA to 100 mM final concentration. Protein, unincorporated dATP, and [α-32P]-dATP were removed using Oligo Clean and Concentrator Spin Columns (Zymo Research). The final product was checked for purity (>95%) and concentration by TLC on PEI-cellulose plates (Sigma) and phosphorimaging.

**Kinetics experiments**

All dNTP solutions and BSA were purchased from New England Biolabs. All chemicals used in the reaction buffer were purchased from Fisher Scientific or Sigma-Aldrich. The exonuclease-deficient D5A E7A variant of T7 gene product 5 was used in all experiments and enzymatic synthesis reactions in this paper to prevent exonuclease degradation of the DNA. Wildtype and E514Cou T7 DNA polymerase as well as thioredoxin were expressed and purified as described (21). All reactions were carried out in T7 Reaction Buffer at 4 °C to better resolve the individual rate constants in the pathway. In some experiments the enzyme-DNA complex was preincubated without Mg2+ and Mg2+-dNTP was added to start the reaction. In other experiments, both reactant mixtures were preincubated in the presence of Mg2+ before mixing with dNTP to start the reaction. In both cases, the final total MgCl2 concentration after mixing was 12.5 mM and there was no observed difference in the kinetics between the two experimental protocols. All concentrations given are those after mixing.

**Quench flow experiments**

Rapid-quench experiments were performed on an RQF-3 instrument (KinTek) with a circulating water bath set to 4 °C. Unless otherwise noted, all experiments were performed with 0.6 M EDTA in the quench syringe and T7 Reaction Buffer without magnesium in the drive syringes. Formamide Loading Buffer (5% sucrose [w/v], 90% formamide [v/v], 0.1% bromophenol blue [w/v], 0.1% xylene cyanol [w/v], 10 mM EDTA) was added to each sample, followed by denaturation at 98 °C for 2 min and separation on 15% polyacrylamide sequencing gels containing 7 M urea for approximately 3 h at 50 °C. Gels were scanned on a Typhoon FLA 9500 laser scanner (GE Healthcare) with the FAM fluorescence filter, and the resulting images were quantified with Image Quant software (GE Healthcare). Concentrations of reaction components

---

**Table 2**

| Oligonucleotides used with extinction coefficients | Sequence (5’→3’) | Extinction coefficient, 260 nm (M⁻¹ cm⁻¹) |
|---|---|---|
| 27 | CCGTCGCAGCCGGTCACAACCACCTACAC | 245,700 |
| 27_d | CCGTGCAGCCGGTCACAACCACCTACAC_d | 245,700 |
| FAM-27 | [6-FAM]-CGGTCGCAGCCGGTCACAACCACCTACAC | 266,660 |
| 28 | CCGTGCAGCCGGTCACAACCACCTACAC | 259,500 |
| [α-32P]-28 | CCGTGCAGCCGGTCACAACCACCTACAC | 259,500 |
| 45-18F | GGACGGCATTTTGATGATGTTGAGTGGATGTTGGACCGGCTGCGACGG | 453,700 |

α32P.
given in the text are final concentrations after mixing unless otherwise noted. Quench flow experiments were repeated on separate occasions at least two to three times to ensure reproducibility.

Stopped flow measurements

All stopped flow experiments were performed on a KinTek SF-300X instrument (dead time of 1.3 ms) with a circulating water bath set to 4 °C. A 150-W Xenon lamp (Hamamatsu) was used as the light source. Samples were excited at 295 nm, and emission was monitored at 445 nm using a filter with a 45-nm band pass (Semrock). All stopped-flow data shown in the main text are an average of at least eight individual traces, and all stopped flow experiments were reproduced at least three times to ensure reproducibility. The TMX titration module for the SF-300X (KinTek Corp) was used for equilibrium titration experiments. Titrations were performed with 280 μl of a solution of enzyme and DNA/dd in the cuvette and 20.5 μl of dATP titrant added from a Hamilton syringe over the course of 5 min. The reaction was mixed by constant stirring from a micro-stir bar in the cuvette and the titration data were corrected for the small dilution before fitting in the KinTek Explorer. The titration experiment was performed on three separate occasions to ensure reproducibility.

Other experiments

For pyrophosphorolysis experiments, a 100 mM stock solution of Na-PPi (Sigma-Aldrich) was prepared, aliquoted, and stored at −80 °C. PPi was incubated in the absence of Mg2+ in the mixture containing T7 DNA polymerase for solubility reasons. Control stopped flow experiments (not shown) indicated that PPi does not bind to the enzyme without Mg2+, so mixing Mg2+ from the other syringe starts the reaction and allows higher concentrations to be reached for short reaction times before Mg2+-PPi precipitates. For the experiment with [α-32P]-28/45 DNA in Figure 4A and the nucleotide dissociation rate experiment in Figure 2E, the formation of [α-32P]-dATP was monitored by TLC on PEI-cellulose plates. To remove contaminants, TLC plates were first developed in ultrapure H2O. Plates were then dried, the contaminants at the top cut off, and the plates stored at 4 °C. Samples were spotted on 1-cm lanes and developed in 0.3 M KPO4 pH 7. Plates were then dried, visualized by phosphorimaging, and quantified in Image Quant.

Kinetics schemes and figures

Structure figures were prepared with the Pymol Molecular Graphics System, Version 2.0 (Schrodinger, LLC). Molecular structures were prepared with the computer program ChemDraw (Perkin Elmer Informatics). Reactions, rate constants, and equilibrium constants in each model are mostly numbered sequentially from left to right, where \( k_i \) is the forward rate constant and \( k_j \) is the reverse rate constant for the \( i^{th} \) step. The convention for rate constants from previous DNA polymerase kinetics papers (1, 4, 11, 14, 17, 34, 35) was used in Equation 1, where \( K_i \) is the equilibrium constant for ground-state nucleotide binding (rapid equilibrium, \( k_{-1} \gg k_j \)), \( k_2 \) is the conformational change, \( k_3 \) is the rate of chemistry, and \( k_4 \) is pyrophosphate release. Translocation steps added to the final model in Figures 5 and 8 are numbered as step 5 and competitive PPi inhibition is modeled as step 6 to maintain these conventions, although they appear earlier in the model than ground-state nucleotide binding.

Conventional data fitting

Initially, data from experiments were fit by nonlinear regression with the analytical fit function in KinTek Explorer to single exponential, double exponential, or burst equations given in the main text. Experiments performed at multiple substrate concentrations were fit to one of these equations and the concentration dependence of the observed rate was then either fit to a hyperbola or linear function. The hyperbolic function defines a maximum rate, \( k_{\text{att}} \), and an apparent \( K_d \) whereas the linear function defines the apparent second-order rate constant for substrate binding. The patterns in the data and concentration dependence of the rates were important to guide the development of a model for more refined data fitting based on numerical integration of rate equations derived from the full model using the KinTek Explorer software (30). All kinetic data in Figures 2–4, 6, 7 are shown fit to equations, defined in the main text, where the black line through the data is the fit to the analytical function.

Data fitting in KinTek explorer

Global data fitting was performed by fitting all experiments in KinTek Explorer to the reaction schemes in the main text, with experimental details of mixing steps and reactant concentrations input for each experiment. In fitting data by simulation, each experiment is modeled exactly as it was.
performed, including allowing pre-equilibrium of the ED complex. Rapid-quench experiments were modeled simply as the sum of species containing product (FP, PP + EP + P) including product DNA (P) release, not shown in the model. Fluorescence transients were modeled by using fluorescence scaling factors. For example, the experiment shown in Figure 3C (7D) was modeled as:

\[ a \cdot (E + ED^\prime + ED + EDA + EP^\prime + EP + b \cdot (FDA + FP \cdot PP))\]

where \( a \) scales the overall signal relative to enzyme concentration and \( b \) represents the fractional change in fluorescence in forming the closed state. In the process of fitting data, a value of \( b = 1.28 \) was derived indicating a 28% increase in fluorescence.

For second-order binding steps, only the equilibrium constant for binding was defined by the data (steps 1 and 6 in the scheme shown in Figure 8), so the binding rate constant was locked at 100 μM⁻¹ s⁻¹ and the reverse rate constant was allowed to float in the fitting to calculate an equilibrium constant. This second-order rate constant represents diffusion-limited substrate binding and is locked at this value to provide a constraint on the rate constants for that step. The confidence contours were derived using the FitSpace function (24). These confidence contour plots are calculated by systematically varying a single rate constant and holding it fixed at a particular value while refitting the data allowing all other rate constants to float. The goodness of fit was scored by the resulting \( \chi^2 \) value. The confidence interval is defined based on a threshold in \( \chi^2 \) calculated from the F-distribution based on the number of data points and number of variable parameters to give the 95% confidence limits. For the data given in Figure 8, this threshold was of 0.99 and was used to estimate the upper and lower limits for each rate constant.

The free-energy profile in Figure 9 was created in KinTek Explorer using the rate constants given in Table 1 using simple transition state theory.

\[ \text{rate} = \frac{A \cdot k_B T}{h} \exp(-\Delta G^\ddagger/RT) \]

\[ 0 < A \leq 1 \]

where \( k_B \) is the Boltzmann constant, \( h \) is the Planck’s constant, and \( R \) is the gas constant. The free-energy profile was created using a transmission coefficient of \( A = 0.01 \) to better show the relationships. Values for the physiological concentrations of dATP (175 μM) (39) and PPi (1.3 mM) (40) were used for each of the second-order rate constants.

Data availability

All data are contained within the manuscript.

Author contributions—T. L. D. and K. A. J. conceptualization, methodology, formal analysis, writing – original draft, writing – review & editing, visualization; T. L. D. validation, investigation; K. A. J. resources, supervision, project administration, funding acquisition.

Conformational dynamics for T7 DNA polymerase

Funding and additional information—This work was supported by grants from NIH (1 R01 GM114223 and R01AI110577 to K. A. J.) and the Welch Foundation (F-1604 to K. A. J.). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Conflict of interest—K. A. J. is president of KinTek Corporation, which provided the SF300x stopped-flow and RQF-3 rapid-quench-flow instruments, and KinTek Explorer software used in this study.

Abbreviations—The abbreviations used are: BSA, bovine serum albumin; PEI, polyethyleneimine.

References

1. Tsai, Y.-C., and Johnson, K. A. (2006) A new paradigm for DNA polymerase specificity. Biochemistry 45, 9675–9687
2. Oertell, K., Harcourt, E. M., Mohsen, M. G., Petruska, J., Kool, E. T., and Goodman, M. F. (2016) Kinetic selection vs. free energy of DNA base pairing in control of polymerase fidelity. Proc. Natl. Acad. Sci. U. S. A. 113, E2277
3. Lee, H. R., Helquist, S. A., Kool, E. T., and Johnson, K. A. (2008) Importance of hydrogen bonding for efficiency and specificity of the human mitochondrial DNA polymerase. J. Biol. Chem. 283, 14402–14410
4. Donlin, M. J., Patel, S. S., and Johnson, K. A. (1991) Kinetic partitioning between the exonuclease and polymerase sites in DNA error correction. Biochemistry 30, 538–546
5. Benkovic, S. J., and Cameron, C. E. (1995) [20] Kinetic analysis of nucleotide incorporation and misincorporation by klenow fragment of Escherichia coli DNA polymerase I. In Methods in Enzymology, Academic Press, Cambridge, MA: 279–269
6. Xu, C., Maxwell, B. A., and Suo, Z. (2014) Conformational dynamics of Thermus aquaticus DNA polymerase I during catalysis. J. Mol. Biol. 426, 2901–2917
7. Bakhtina, M., Roettger, M. P., and Tsai, M.-D. (2009) Contribution of the reverse rate of the conformational step to polymerase β fidelity. Biochemistry 48, 3197–3208
8. Zhong, X., Patel, S. S., Werneburg, B. G., and Tsai, M.-D. (1997) DNA polymerase β: multiple conformational changes in the mechanism of catalysis. Biochemistry 36, 11891–11900
9. Washington, M. T., Johnson, R. E., Prakash, L., and Prakash, S. (2004) Human DNA polymerase iota utilizes different nucleotide incorporation mechanisms dependent upon the template base. Mol. Cell. Biol. 24, 936–943
10. Johnson, A. A., and Johnson, K. A. (2001) Exonuclease proofreading by human mitochondrial DNA polymerase. J. Biol. Chem. 276, 38097–38107
11. Wong, L., Patel, S. S., and Johnson, K. A. (1991) An induced-fit kinetic mechanism for DNA replication fidelity: direct measurement by single-turnover kinetics. Biochemistry 30, 526–537
12. Fernandez-leiro, R., Conrad, I., Yang, J.-c., Freund, S. M. V., Scheres, S. H. W., and Lamers, M. H. (2017) Self-correcting mismatches during high-fidelity DNA replication. Nat. Struct. Mol. Biol. 24, 140–143
13. Doublé, S., Tabor, S., Long, A. M., Richardson, C. C., and Ellenberger, T. (1998) Crystal structure of a bacteriophage T7 DNA replication complex at 2.2 Å resolution. Nature 391, 251–258
14. Patel, S. S., Wong, L., and Johnson, K. A. (1991) Pre-steady-state kinetic analysis of processive DNA replication including complete characterization of an exonuclease-deficient mutant. Biochemistry 30, 511–525
15. Herschlag, D., Piccirilli, J. A., and Cech, T. R. (1991) Ribozyme-catalyzed and nonenzymatic reactions of phosphate diesters: rate effects upon substitution of sulfur for a nonbridging phosphor oxygen atom. Biochemistry 30, 4844–4854
16. Showalter, A. K., and Tsai, M.-D. (2002) A reexamination of the nucleotide incorporation fidelity of DNA polymerases. Biochemistry 41, 10571–10576
17. Kellinger, M. W., and Johnson, K. A. (2010) Nucleotide-dependent conformational change governs specificity and analog discrimination
Conformational dynamics for T7 DNA polymerase

by HIV reverse transcriptase. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 7734–7739

18. Ram Prasad, B., Kamerlin, S. C. L., Florián, J., and Warshel, A. (2012) Prechemistry barriers and checkpoints do not contribute to fidelity and catalysis as long as they are not rate limiting. *Theor. Chem. Acc.* **131**, 1288

19. Kimzilatilin, S., Johnson, K. A., and Elber, R. (2015) Enzyme selectivity of HIV reverse transcriptase: conformations, ligands, and free energy partition. *J. Phys. Chem. B* **119**, 11513–11526

20. Kimzilatilin, S., Nguyen, V., Johnson, K. A., and Elber, R. (2012) How conformational dynamics of DNA polymerase select correct substrates: experiments and simulations. *Structure* **20**, 618–627

21. Dangerfield, T. L., and Johnson, K. A. (2020) Optimized incorporation of an unnatural fluorescent amino acid affords measurement of conformational dynamics governing high-fidelity DNA replication. *J. Biol. Chem.* **295**, 17265–17280

22. Dahl, J. M., Mai, A. H., Cherf, G. M., Jetha, N. N., Garalde, D. R., Marziali, A., Akeson, M., Wang, H., and Lieberman, K. R. (2012) Direct observation of translocation in individual DNA polymerase complexes. *J. Biol. Chem.* **287**, 13407–13421

23. Lieberman, K. R., Dahl, J. M., Mai, A. H., Cox, A., Akeson, M., and Wang, H. (2013) Kinetic mechanism of translocation and dNTP binding in individual DNA polymerase complexes. *J. Am. Chem. Soc.* **135**, 9149–9155

24. Johnson, K. A., Simpson, Z. B., and Blom, T. (2009) *FitSpace explorer: an algorithm to evaluate multidimensional parameter space in fitting kinetic data*. *Anal. Biochem.* **387**, 30–41

25. Hammes, G. G., Chang, Y.-C., and Oas, T. G. (2009) Conformational dynamics for T7 DNA polymerase. *Biochemistry* **48**, 6275–6286

26. Hammel, M., Ellenberger, T., Gainey, M. D., and Wallen, J. R. (2019) Combined solution and crystal methods reveal the electrostatic tethers that provide a flexible platform for replication activities in the bacteriophage T7 replisome. *Biochemistry* **58**, 4466–4479