JBC RESEARCH ARTICLE

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 (>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 (ED_n) binds nucleotide (N) in the open state. A transition to the closed state ($ED_nN \rightarrow FD_nN$) then facilitates the chemical reaction followed by enzyme opening and release of pyrophosphate (PP_i).

$$ED_n + N \overset{K1}{\rightleftarrows} {}^{K_1}ED_nN \underset{k_{-2}}{\overset{k_2}{\rightleftarrows}} FD_nN \underset{k_{-3}}{\overset{k_3}{\rightleftarrows}} FD_{n+1}.PP_i \underset{k_{-4}}{\overset{k_4}{\rightleftarrows}} ED_{n+1}+PP_i \tag{1}$$

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.

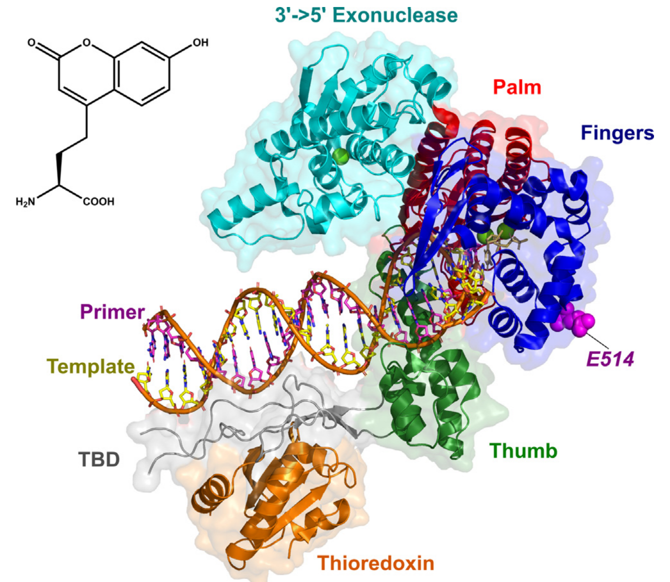

Figure 1. Structure of fluorescent unnatural amino acid and T7 DNA polymerase. *Top left*, structure of the fluorescent unnatural amino acid, L-(7-hydroxycoumarin-4-yl) ethylglycine (7-HCou). *Right*, structure of T7 DNA polymerase: The overall structure resembles a right hand with the polymerase active site situated at the interface of the palm domain (*red*), fingers domain (*blue*), and thumb domain (*green*). 7-HCou is inserted at position 514, on the back side of the fingers domain and is shown as *magenta spheres*. Thioredoxin (*orange*) is a 12-kDa host redox protein that associates with T7 gene product 5 through a unique extension of the thumb domain called the thioredoxin binding domain (TBD, *gray*) and helps with stable primer (*magenta*)/template (*yellow*) DNA binding. The 3′-5′ proofreading exonuclease domain (*cyan*) increases the fidelity of the enzyme by at least a factor of 1000. Magnesium ions are shown as *green spheres* in the exonuclease and polymerase active sites. The structure was drawn using Pymol from PDBID: 6p7e (41).

Initial attempts to look for a rate-limiting conformational change were based on examining the effect of sulfur substitution of one of the nonbridging oxygens of the incoming nucleotide, ostensibly to slow the rate of chemical reaction. The absence of a thio-elemental effect on the measured rate of incorporation was used to argue for a rate-limiting conformational change preceding chemistry (11, 14). However, subsequent studies showed a much smaller-than-anticipated inherent elemental effect (15) and studies on DNA polymerase β provided evidence for a fast conformational change preceding chemistry (16). Accordingly, the rate of the conformational change was deemed to be unimportant for fidelity other than contributing to the fraction of enzyme in the closed state to facilitate catalysis. However, subsequent measurements using higher-fidelity enzymes led to a new paradigm for understanding the role of fast conformational changes in fidelity based on kinetic partitioning of the closed FD_nN state (1). Essentially, the conformational change locks a correct nucleotide at the active site to commit it to the forward reaction, but the enzyme rapidly opens to afford release rather than incorporation of a mismatch. Further studies on HIV reverse transcriptase provided more complete data to support this model (17), but it still remained controversial (18) in spite of support from molecular dynamics simulation methods (19, 20).

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, and k_{-2} in Equation 2) using a 2′,3′ dideoxy-nucleotide-terminated primer to allow dATP binding but prevent the chemical reaction.

$$ED_{dd}+dATP \overset{K1}{\rightleftarrows} {}^{K_1}ED_{dd}dATP \underset{k_{-2}}{\overset{k_2}{\rightleftarrows}} FD_{dd}dATP \tag{2}$$

A 27-nt (nucleotide) 2′,3′-dideoxy-terminated primer strand was annealed to a 45-nt template strand ($27_{dd}/45$, 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

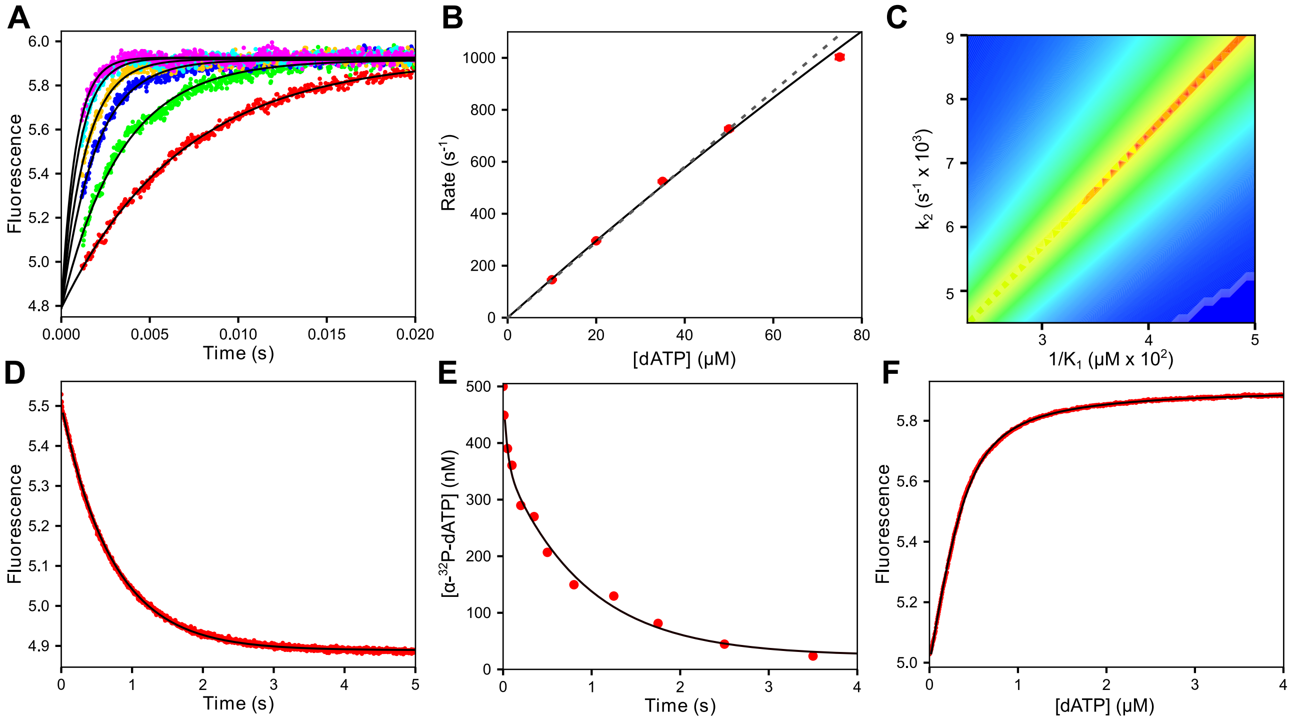

Figure 2. Kinetics and equilibrium of nucleotide binding steps preceding chemistry. *A*, stopped flow dATP binding rate. T7 DNA polymerase E514Cou (750 nM), thioredoxin (15 μM), and DNA$_{dd}$ (1 μM) were mixed with dATP (5–100 μM) and Mg^{2+} (12.5 mM) to start the reaction. *Colored dots* are stopped 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_{max}: 12,200 ± 6850 s^{-1}, $K_{d,app}$: 805 ± 470 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_1) and the conformational change step (k_2) 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 χ^2 confidence threshold and shows the linear relationship between k_2 and 1/K_1. *D*, fluorescence measurement of dATP dissociation rate. T7 DNA polymerase E514Cou (500 nM), thioredoxin (10 μM), DNA$_{dd}$ (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-DNA$_{dd}$-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), DNA$_{dd}$ (500 nM), BSA (0.1 mg/ml), and [α-^{32}P]-dATP (500 nM) were mixed with wildtype T7 DNA polymerase (5 μM), thioredoxin (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-^{32}P]-dATP incorporated rapidly into the trap DNA and the second slower phase corresponding to the nucleotide dissociation rate has a rate of 1.25 ± 0.15 s^{-1}. *F*, equilibrium dATP binding titration. A solution of T7 DNA polymerase E514Cou (400 nM), thioredoxin (8 μM), DNA$_{dd}$ (600 nM), and Mg^{2+} (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 $K_{d,dATP}$: 107 ± 1 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-DNA$_{dd}$ complex with various concentrations of dATP (5–100 μM) and Mg^{2+} in the stopped flow (Fig. 2*A*). 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:

$$y = A_0 + A_1 \cdot (1 - e^{-\lambda_1 t}) \tag{3}$$

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

$$rate = \frac{k_{max}[dATP]}{K_d + [dATP]} = \frac{K_1 k_2 [dATP]}{1 + K_1 [dATP]} + k_{-2} \tag{4}$$

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 ($K_1 k_2$) leading to the $FD_{dd}N$ state (Equation 2). A slightly better fit to the data was obtained assuming a rapid-equilibrium ground-state dATP binding (K_1) 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 ± 6850 s^{-1}) and the $K_d = 1/K_1$ (805 ± 470 μM). The initial slope of the concentration dependence provides an accurate estimate for K_1k_2 for the two-step model, even though neither K_1 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 $k_2 \geq 6000\ s^{-1}$ as well as a well-defined value for K_1k_2 (Fig. 2*C*) shown by the linear relationship between k_2 and $1/K_1$ 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 $E.DNA_{dd}.dATP$ 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 $E.DNA_{dd}.dATP$ 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 $E.DNA_{dd}.dATP$ ternary complex. The decrease in fluorescence (Fig. 2*D*) 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. 2*E*).

An $E.DNA_{dd}.[\alpha\text{-}^{32}P]dATP$ ternary complex was mixed with a large excess of wildtype *E.DNA* complex to start the reaction. Dissociation of the [α-^{32}P]-dATP from the $ED_{dd}.[\alpha\text{-}^{32}P]dATP$ complex was monitored by the decrease in [α-^{32}P]-ATP 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 27/45 by the wildtype enzyme in the trap mixture. The data fit a double exponential—modeling the experiment by simulation provided a rationale for understanding the biphasic data. The small initial fast phase is due to the small amount of unbound nucleotide from the $E.DNA_{dd}.dATP$ 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. 2*F*). 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. 2*F*). After correcting the fluorescence intensity for the small dilution during the titration, the resulting curve was fit using a quadratic equation (below):

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

where F_0 is the initial fluorescence, ΔF 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 = 107 ± 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.

$$K_{d,net} = \frac{1}{K_1(1+K_2)} \simeq \frac{1}{K_1K_2} = \frac{k_{-2}}{K_1k_2} = 96 \pm 2\ nM \qquad (5)$$

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_1k_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. 3*A*). 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):

$$y = A_1 \cdot (1 - e^{-\lambda_1 t}) + k_{ss}t \qquad (6)$$

where A_1 is the amplitude of the exponential phase, λ_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. 3*B*) provided estimates for the parameters from the maximum observed rate, k_{pol} = 210 ± 15 s^{-1}, and half-maximal nucleotide concentration to define an *apparent* K_d, $K_{d,app}$ = 15.5 ± 2 μM. The ratio $k_{pol}/K_{d,app}$ = 13 ± 2 $\mu M^{-1}\ s^{-1}$ 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 ∼100 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 3*C*, 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):

$$y = A_0 + A_1 \cdot (1 - e^{-\lambda_1 t}) + A_2 \cdot (1 - e^{-\lambda_2 t}) \qquad (7)$$

where A_0 is the initial fluorescence, whereas A_1 and λ_1, and A_2 and λ_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. 3*D*) 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 DNA_{dd} in Figure 2*A*.

The increase in fluorescence is a function of the transition from open to closed enzyme ($ED_nN \rightarrow FD_nN$), 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 dideoxy-terminated 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_4) 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}, and k_4 from the concentration dependence of the rate and amplitude of the pre-steady-state burst of incorporation (Fig. 3*A*) and the fluorescence transients (Fig. 3*C*). Simultaneous fitting of rates and amplitudes of the reactions allows resolution of k_3, k_{-3}, and k_4, 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}) can be extracted from analysis of the pyrophosphorolysis reaction to define the internal equilibrium 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 [α-^{32}P]-dNMP incorporated enzymatically at the 3’ end of the primer strand of DNA 27/45 (named [α-^{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 [α-^{32}P]-dATP (Fig. 4*A*).

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

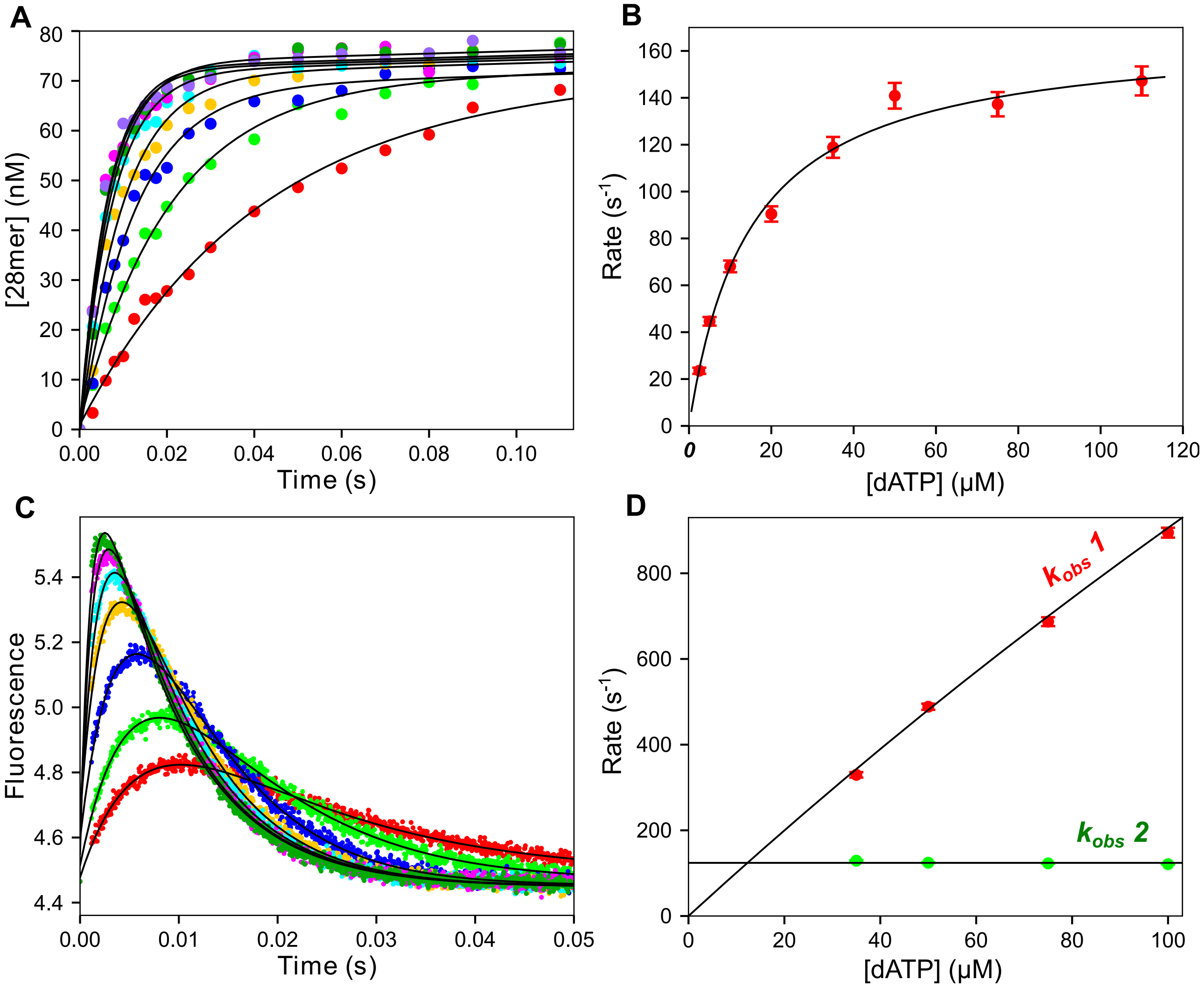

Figure 3. Kinetics of dATP incorporation. *A*, quench flow pre-steady-state burst experiment. T7 DNA polymerase E514Cou (120 nM), BSA (0.1 mg/ml), thioredoxin (2.4 μM), and FAM-27/45 DNA (200 nM) were mixed with dATP (2.5–110 μM) and Mg^{2+} (12.5 mM) to start the reaction. Data at various dATP concentrations 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_{pol} = 210 ± 15 s^{-1} and $K_{d,app}$ =15.5 ± 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_{obs} 1) were fit to a hyperbola to estimate the maximum rate of the conformational change (k_2), 7800 ± 3400 s^{-1}. The slow phase data (*green points*, k_{obs} 2) were fit to a line, setting a lower limit on the maximum rate of the isomerization following chemistry (k_4) of 126 ± 4 s^{-1}.

concentration of dATP used in the experiment was chosen to preclude incorporation of a dATP:dA mismatch on top of the [α-^{32}P]-labeled primer strand (11) but is sufficient to rapidly replace the [α-^{32}P]-dAMP after its removal from the primer. These details allow the experiment to be interpreted without complications. In addition, although Mg^{2+}-PP$_i$ precipitates at concentrations above ~5 mM, the precipitation occurs on the timescale of approximately 30 s, whereas the pyrophosphorolysis reactions at concentrations of Mg^{2+}-PP$_i$ ≥ 5 mM occurs on a shorter timescale. Therefore, measurements can be made before precipitation by mixing PP$_i$ (without Mg^{2+}) and the E-[α-^{32}P]-28/45 complex with Mg^{2+} to start the reaction. A control stopped flow experiment showed PP$_i$ does not bind to the enzyme in the absence of Mg^{2+} (not shown).

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

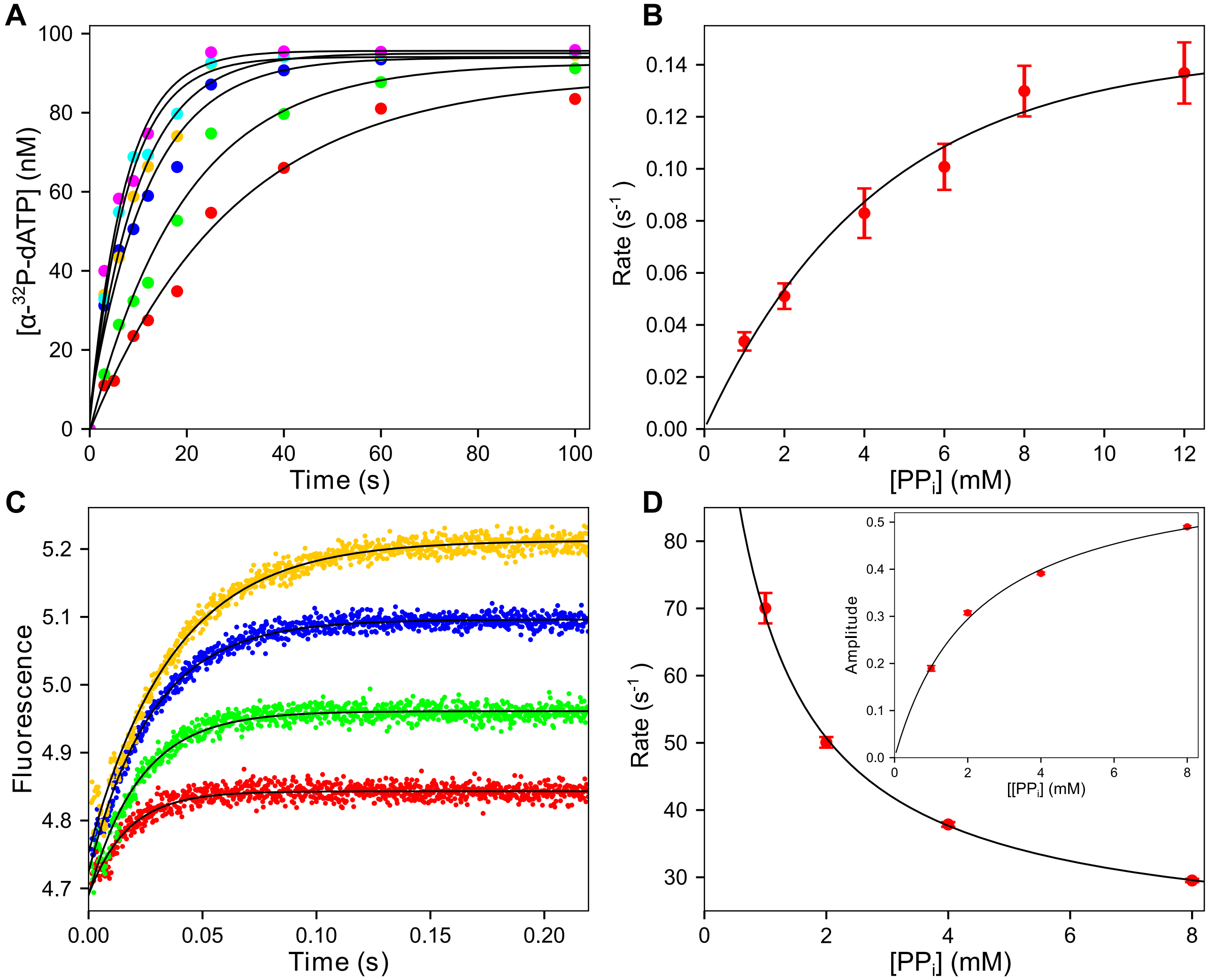

Figure 4. Kinetics of pyrophosphorolysis. *A*, hand quench pyrophosphorolysis with 3′-[α- ^{32}P]-labeled DNA. A solution of T7 DNA polymerase E514Cou (500 nM), thioredoxin (10 μM), BSA (0.1 mg/ml), 3′-[α-^{32}P]-28/45 DNA (100 nM), and PP$_i$ (1–12 mM) was mixed with unlabeled dATP (20 μM) and Mg^{2+} (12.5 mM) to start the reaction. Data at each PP$_i$ concentration are shown fit to a single exponential function (*black lines*). *B*, rate *versus* PP$_i$ concentration for pyrophosphorolysis experiment. Rates are from the data in (*A*). Data are shown fit to a hyperbola (*black line*), estimating an apparent $K_{d,PPi}$: 5.6 ± 1.2 mM and k_{PPi} 0.20 ± 0.02 s^{-1}. *C*, stopped flow pyrophosphorolysis T7 DNA polymerase E514Cou (500 nM), thioredoxin (10 μM), 28/45 DNA (600 nM), and PP$_i$ (1–8 mM) were mixed with Mg^{2+} (12.5 mM) to start the reaction. Data at each PP$_i$ concentration were fit to a single exponential function (*black lines*). *D*, rate and amplitude *versus* PP$_i$ concentration for stopped flow pyrophosphorolysis. Rates and amplitudes are from data in (*C*). Data are shown fit to a hyperbola (*black line*) estimating an apparent $K_{d,PPi}$: 950 ± 390 μM. *Inset:* Amplitude *versus* PP$_i$ concentration. Shown fit to a hyperbola (*black line*), estimating an apparent K_d of 2.2 ± 0.3 mM. Increasing amplitude with increasing PP$_i$ concentration is indicative of a "conformational selection" mechanism.

Conformational changes during pyrophosphorolysis

We reasoned that, since there is a fluorescence signal for the forward reaction, there should also be a fluorescence signal associated with PP$_i$ binding upon forming the FD_nN and $FD_{n+1}.PP_i$ states that can be followed by stopped-flow methods (Equation 1). We performed this experiment (Fig. 4*C*) by mixing a solution containing the *E.DNA* complex and varying concentrations of PP$_i$ (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 PP$_i$ 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 PP$_i$ concentrations fit a single exponential function (Equation 3). Fitting the rate and amplitude *versus* PP$_i$ concentration using a hyperbola is shown in Figure 4*D*. Of interest, the observed rate of the fluorescence transient decreased with increasing PP$_i$ concentrations with an *apparent* K_d = 950 ± 390 μM. 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 (PP$_i$ 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 PP$_i$ binding (*i.e.*, translocation).

A reasonable fit to the data was obtained by adding a translocation step (ED′→ ED or EP′→ EP) after PP$_i$ 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 PP$_i$ as first described in prior work on the T7 DNA polymerase (11, 14). The concept of binding of PP$_i$ 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 PP$_i$ 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 PP$_i$ and dNTP bind to the same enzyme state, and subsequent to dNTP binding translocation and PP$_i$ 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 PP$_i$) 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, PP$_i$ 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 [PP$_i$] (Fig. 4*D*) 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 PP$_i$, the data are sufficient to set a lower limit of ~100 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 PP$_i$ concentrations, extrapolated to give k_{-5} ~10 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 PP$_i$

The model in Figure 5 has some important and testable implications that distinguish it from the power stroke model where the nucleotide and PP$_i$ 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 PP$_i$.

Stopped flow PP$_i$ inhibition on dATP incorporation

With the model in Figure 5, inhibition by rebinding PP$_i$ directly to the EP' state formed after chemistry should appear to be much tighter than the affinity derived from the experiment in Figure 4*A* where PP$_i$ 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 PP$_i$. In this experiment a preformed *E.DNA* complex with varying concentrations of PP$_i$ (0–8 mM) was mixed with Mg^{2+} and a fixed concentration of dATP (50 μM) in the stopped flow to start the reaction (Fig. 6*A*). Traces at each PP$_i$ concentration were fit to a double exponential function, and the observed rates for the fast and slow phases were plotted *versus* PP$_i$ concentration (Fig. 6*B*). It is surprising that the rate of the fast phase, corresponding to dATP binding steps, showed only a slight inhibition with increasing PP$_i$ concentrations (step 6 in Fig. 5). This inhibition of the fast phase fits a linear function

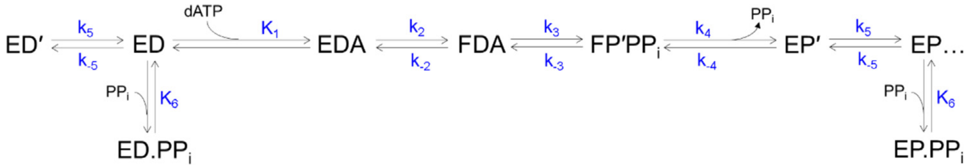

Figure 5. Full kinetic model for nucleotide incorporation by T7 DNA polymerase. This pathway is expanded from Equation 1 to Include translocation steps (k_5, k_{-5}) and competitive PP$_i$ binding to the post-translocated state (K_6).

with a slope of −0.027 ± 0.003 mM^{-1} s^{-1}. The concentration dependence on the rate of the isomerization accompanying PP_i release following chemistry showed an apparent K_d = 775 ± 75 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_6), 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_{-2}), chemistry (k_3), or the reverse of translocation (k_{-5}); however, the reverse of chemistry (k_{-3}) was increased ∼3 fold to 136 s^{-1}, requiring also a small increase in k_4 to 220 s^{-1} and an increase in k_5 to ∼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

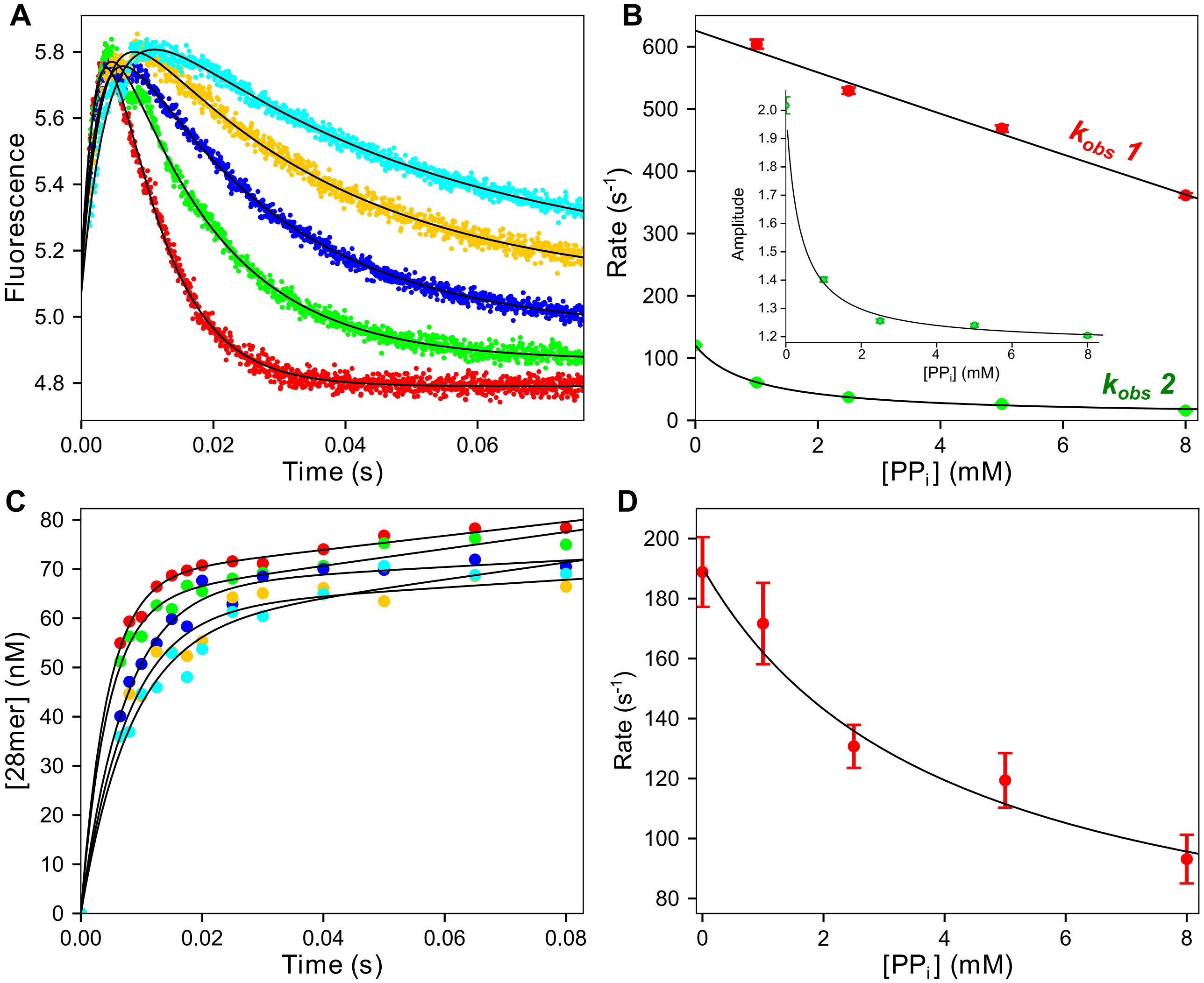

Figure 6. Kinetics of PP_i inhibition on dATP incorporation. *A*, fluorescence measurement during PP_i inhibition on dATP incorporation. T7 DNA polymerase E514Cou (500 nM), thioredoxin (10 μM), 27/45 DNA (600 nM), and PP_i (0–8 mM) were mixed with dATP (50 μM) and Mg^{2+} (12.5 mM) to start the reaction. Data at each concentration were fit to a double exponential function (*black lines*). *B*, rate *versus* PP_i concentration for inhibition of dATP incorporation. Rates are from data in (*A*). Data are shown for the fast phase (*red*, k_{obs} 1) and slow phase (*green*, k_{obs} 2), fit to a line and a hyperbola, respectively (*black lines*). The slope of the linear fit of the slow phase derived from the fitting is −0.028 ± 0.003 mM^{-1} s^{-1}, which suggests weak competitive inhibition by PP_i for binding to the enzyme. The hyperbolic fit to the slow phase gives an apparent $K_{d,PPi}$: 775 ± 75 μM. *Inset:* Amplitude *versus* PP_i concentration for the slow phase, estimating an apparent K_d of 350 ± 80 μM. *C*, quench flow measurement of PP_i inhibition of dATP incorporation. T7 DNA polymerase E514Cou (120 nM), thioredoxin (2.4 μM), FAM-27/45 DNA (200 nM), BSA (0.1 mg/ml), and PP_i (0–12 mM) from one syringe were mixed with dATP (20 μM) and Mg^{2+} (12.5 mM) from the other syringe to start the reaction. Data at each PP_i concentration are shown fit to a burst equation (*black lines*). *D*, rate *versus* PP_i for inhibition on dATP incorporation. Rates are from data in (*C*). Data for the exponential phase of the burst at each concentration are shown fit to a hyperbola (*black line*) and gives an apparent $K_{i,PPi}$: 5.9 ± 3.1 mM.

the reaction (Fig. 6*C*). The data at each PP_i concentration were fit to a burst equation (Equation 6), and the observed rate of the exponential phase *versus* PP_i concentration is shown fit to a hyperbola in Figure 6*D*. Minimal inhibition is observed as the majority of the enzyme species is in the state ready to bind dATP (*ED*) with only weak competitive PP_i inhibition, not the form (*ED*′) that tightly binds PP_i. PP_i 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 PP_i 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 PP_i release and translocation in the reaction pathway. The second prediction is that PP_i 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 PP_i (ED'_{28} in Equation 8) with relatively high affinity. These two predictions are tested with the following experiment.

$$\begin{aligned} ED_{27} &\xrightarrow{+dATP} FD_{28}{}'PP_i \underset{PP_i}{\rightleftarrows} ED'_{28} \overset{trans.}{\rightleftarrows} ED_{28} \\ ED_{28} &\xrightarrow{+dTTP} FD_{29}{}'PP_i \underset{PP_i}{\rightleftarrows} ED'_{29} \overset{trans.}{\rightleftarrows} ED_{29} \end{aligned} \tag{8}$$

Processive polymerization and inhibition by PP_i

To test these two hypotheses, we performed a rapid quench experiment where an *E.DNA* complex, in the presence or absence of 8 mM PP_i, was mixed with dATP (100 μM), dTTP (100 μM), and Mg^{2+} to start the reaction (Fig. 7). Fitting the data for the experiment without PP_i (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: $ED_{27} \rightarrow ED_{28} \rightarrow ED_{29}$. The second incorporation is much slower than the first incorporation, owing to the rate-limiting PP_i release and translocation steps that occur after fast chemistry for the first incorporation (Equation 8). The experiment performed in the presence of 8 mM PP_i (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.

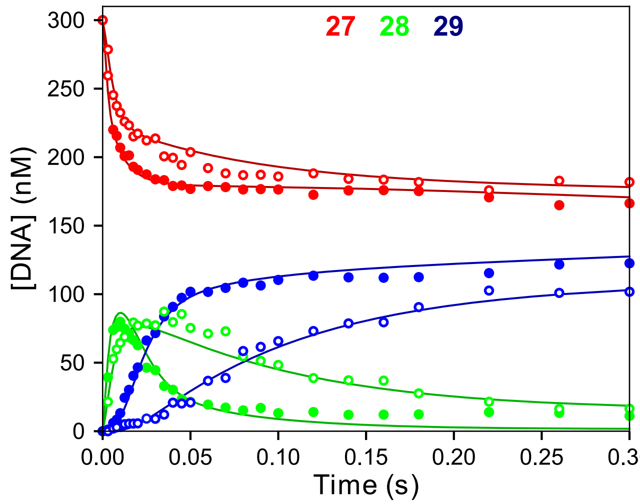

Figure 7. PP_i inhibition on processive polymerization. T7 DNA polymerase E514Cou (120 nM), thioredoxin (4 μM), FAM-27/45 DNA (300 nM), BSA (0.1 mg/ml), and PP_i (0 or 8 mM) were mixed with dATP (100 μM), dTTP (100 μM), and Mg^{2+} (12.5 mM) to start the reaction. DNA lengths for each color are given at the top of the panel. The closed data points were collected without added PP_i, and the open data points had 8 mM PP_i in the reaction. The rates for the first and second incorporations for the simple model $ED_{27} \rightarrow ED_{28} \rightarrow ED_{29}$ for the reaction without PP_i were 182 ± 10 s^{-1} and 41 ± 1 s^{-1}, respectively. The rates derived from the fit by simulation for the two incorporations for the reaction with 8 mM PP_i are 104 ± 8 s^{-1} and 9.3 ± 0.6 s^{-1}, respectively. Lines through the data were computed based on global data fitting shown in Figure 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 χ^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

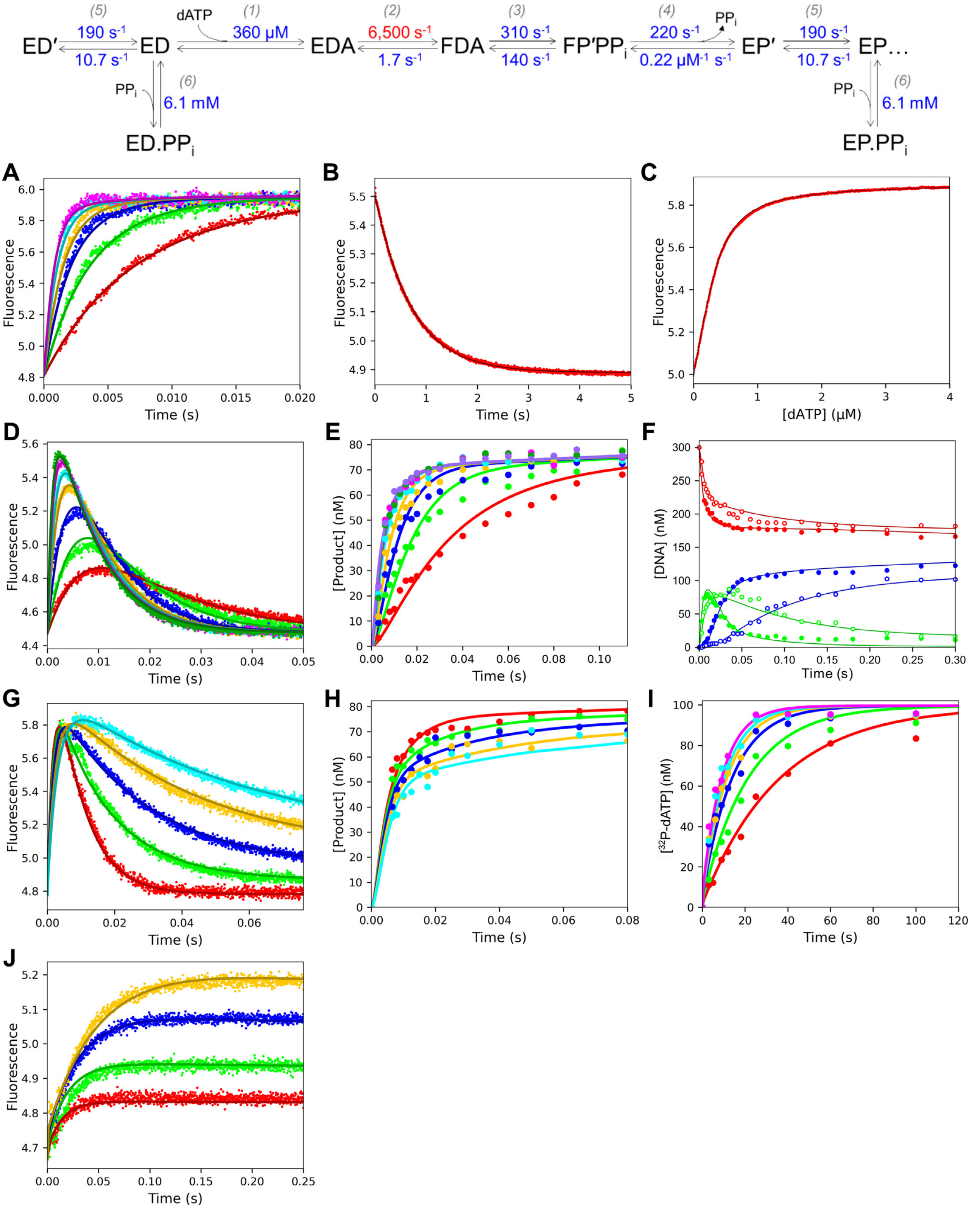

Figure 8. Global fit of all experimental data. The scheme at the top contains the rate constants derived from the global fit of the data shown in this figure. Error estimates for each rate constant are given in Table 1. Rate constants in *red* were locked during the fitting, and the numbers in parenthesis 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 PP_i inhibition of processive polymerization. *G*, stopped flow measurement during PP_i inhibition of dATP incorporation. H, quench flow measurement of PP_i inhibition of dATP incorporation. *I*, kinetics of pyrophosphorolysis. *J*, stopped flow pyrophosphorolysis.

Table 1
Rate constants derived by global data fitting

Parameter	Best-fit value	95% Confidence interval
K_1k_2	$18 \pm 0.6\ \mu M^{-1}\ s^{-1}$	17.4–18.6
$1/K_1$	$360 \pm 12\ \mu M$	351–377
k_2	$[6500\ s^{-1}]$	≥6000
k_{-2}	$1.70 \pm 0.04\ s^{-1}$	1.67–1.74
k_3	$310 \pm 40\ s^{-1}$	273–359
k_{-3}	$140 \pm 30\ s^{-1}$	115–170
k_4	$220 \pm 18\ s^{-1}$	210–246
k_{-4}	$0.22 \pm 0.03\ \mu M^{-1}\ s^{-1}$	0.201–0.261
k_5	$190 \pm 25\ s^{-1}$	171–222
k_{-5}	$10.7 \pm 0.6\ s^{-1}$	10.1–11.3
K_6	$6.1 \pm 0.3\ mM$	5.74–6.43

Lower and upper limits are reported from a 0.99 χ^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 to define the dissociation rate constant was used to estimate the equilibrium constant. For example, k_1 was locked at 100 $\mu M^{-1}\ s^{-1}$ 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.

contour analysis with a χ^2 threshold of 0.99. From the mechanism and rate constants in Figure 8 and physiological dATP and PP_i 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:

$$k_{cat} = \frac{k_2k_3k_4k_5}{k_{-3}k_5(k_2+k_{-2})+k_4k_5(k_2+k_{-2}+k_3)+k_2k_3(k_4+k_5)}$$

$$k_{cat} \Big/ K_m = \frac{k_1k_2k_3k_4k_5}{(k_5+k_{-5})\{k_{-1}k_{-2}(k_{-3}+k_4)+k_3k_4(k_{-1}+k_2)\}}$$

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

$$K_{cat} / K_m \approx K_1k_2k_5 / (k_5 + k_{-5})$$

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%).

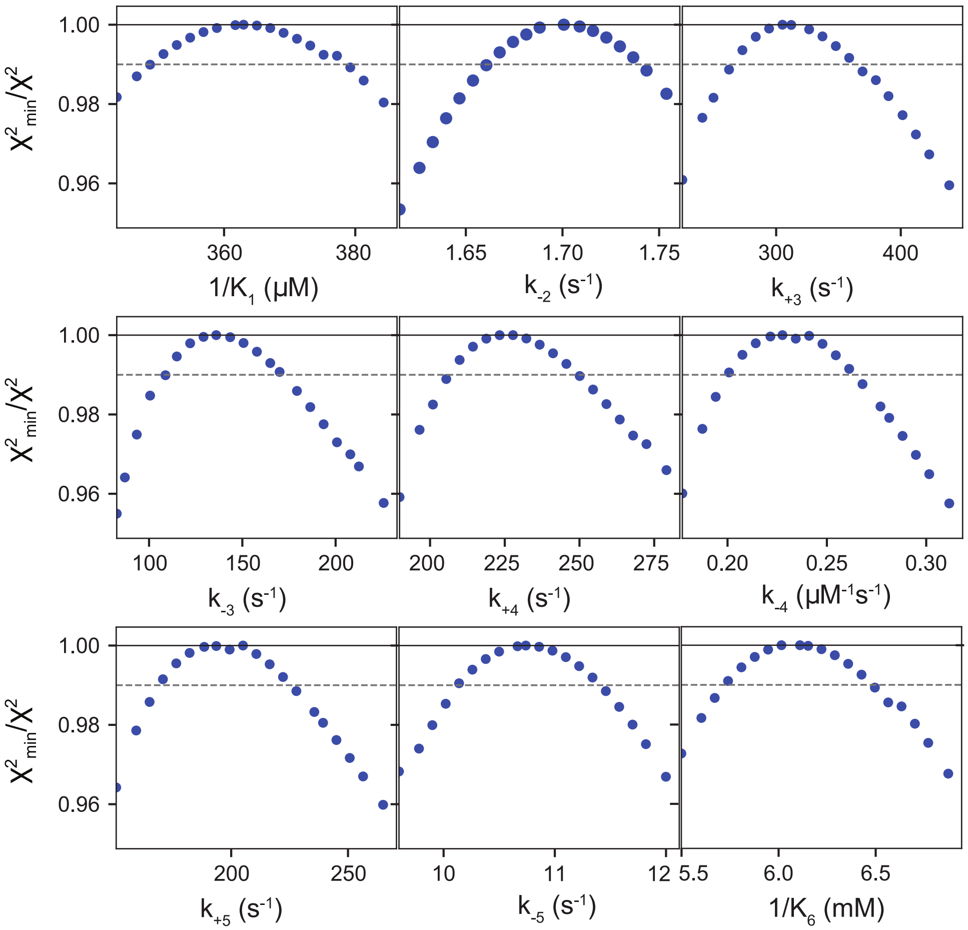

Figure 9. Confidence contour plots for rate constants in the dATP incorporation pathway. The confidence contours from the 1D FitSpace calculation in KinTek Explorer are shown as *blue data points* for each rate constant. The *gray dashed line* shows the 0.99 χ^2 threshold to establish 95% confidence intervals for each rate constant, calculated with KinTek Explorer and summarized in Table 1.

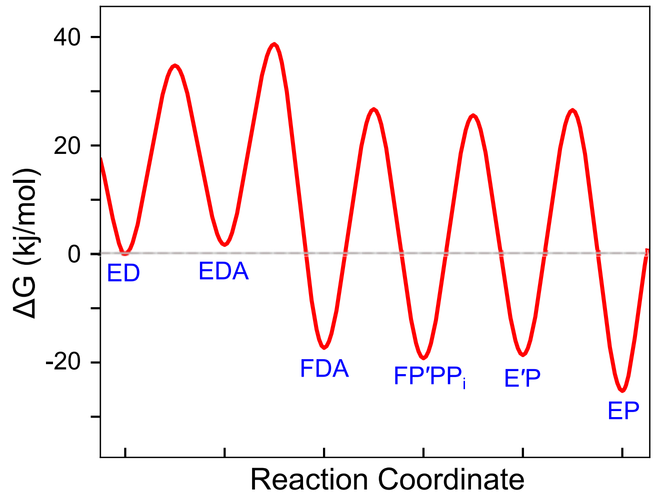

Figure 10. Free-energy profile for dATP incorporation at 4 °C. The free-energy profile was generated using the rate constants and model in the scheme at the top of Figure 8, at 277 K, and physiological concentrations of substrate and product (175 μM dATP, 1.3 mM PP$_i$). A transmission coefficient of 0.01 was used to better display the relationships.

Discussion

The work presented here provides the first fully complete kinetic analysis of correct nucleotide incorporation by a high-fidelity DNA polymerase. Global data fitting was required to achieve a complete fit of the full data set to support a single unifying model. It is difficult to reconcile the results of conventional equation-based data fitting with the intrinsic rate constants given in the scheme in Figure 8 because the simplifying assumptions required to derive equations are inaccurate for this complex model. For example, the nucleotide concentration dependence of the rate of incorporation (Fig. 3, *A*–*B*) gave a maximum polymerization rate, k_{pol} = 210 s^{-1}, and an apparent $K_{d,app}$ value of 15.5 μM providing an estimate of $k_{pol}/K_{d,app}$ = 13.5 μM^{-1} s^{-1}. The specificity constant defined by global fitting was somewhat higher, k_{cat}/K_m = 17 μM^{-1} s^{-1}, whereas the rate constant for the chemical reaction was significantly higher than the observed rate of the burst, k_3 = 312 s^{-1}. The solution of a simple burst equation implies that the observed decay rate should be equal to the sum of all rate constants contributing to the endpoint. However, based on global data fitting we now know that the pre-steady-state burst of incorporation will not follow a single exponential; rather, it is a sum of two or three exponentials with comparable rate constants that blend together to give an observed decay rate that can be adequately fit using a single exponential function with a decay rate that is slower than the individual rate constants. Certainly, quench flow data with few data points are not sufficient to resolve multiple exponentials. Similarly, measurements of the kinetics of pyrophosphorolysis yielded results that are attenuated by translocation, the conformational change, and the reversible chemical reaction. In globally fitting data from all experiments, the underlying model and intrinsic rate constants are resolved by taking into account the way in which different aspects of the pathway affect a given measurement designed to measure a single step.

The apparent $K_{d,PPi}$ from the rate *versus* concentration plot in Figure 4*D* is 950 ± 390 μM, which is defined in terms of rate constants as $K_{d,PPi} = k_4(k_5 + k_{-5})/k_{-4}k_5$, which approximates the true K_d for PP$_i$ binding to the pretranslocated state. Given estimates for k_4 (125 s^{-1} from Fig. 3), k_{-4}, or the second-order rate constant for PP$_i$ binding to the pretranslocated state, can be estimated around 0.2 μM^{-1} s^{-1}. When compared with the value for k_{-4} in fitting to Equation 1 (~0.02 μM^{-1} s^{-1}), with unchanged k_4, this model predicts PP$_i$ binding ~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 PP$_i$ 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 ($1/K_1$ ≥ 360 μM) induces a conformational change (k_2, ≥ 6000 s^{-1}) that is much faster than the rate of chemistry (k_3 = 310 s^{-1}). Although neither K_1 nor k_2 is well defined, the apparent second-order rate constant for nucleotide binding is well defined as K_1k_2 = 18 ± 0.6 μM^{-1} s^{-1} (Fig. 2*C*). 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. 2*D*). 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 μM (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 μM), 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-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_{cat}/K_m is also affected by the two-step rapid-equilibrium nucleotide binding. Because $k_{-2} > k_3$, the specificity constant is approximately defined by $k_{cat}/K_m \approx K_1K_2k_3$.

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. 3*C*) revealed the closing of the enzyme after binding nucleotide followed by chemistry (measured by rapid quench in Fig. 3*A*) 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. 4*C*), 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′ state. These experiments help to define the kinetics of the translocation step from ED′ to ED with an equilibrium constant of K_5 = 18 in the forward direction.

$$ED' \underset{11\ s^{-1}}{\overset{188\ s^{-1}}{\rightleftarrows}} ED$$

In this model, PP_i release occurs with enzyme opening to form ED′ 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_6 = 18, ΔG = −6.7 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 posttranslocation 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 ^{32}P-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-^{32}P]-dAMP labeled 28/45 DNA substrate

The following components were mixed together in T7 Reaction Buffer (40 mM Tris-HCl, pH 7.5, 50 mM NaCl, 1 mM EDTA, 1 mM DTT) (14) and incubated at room temperature (25 °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-^{32}P]-dATP (Perkin Elmer), and 12.5 mM Mg^{2+}. 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-^{32}P]-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 Mg^{2+} and Mg^{2+}-dNTP was added to start the reaction. In other experiments, both reactant mixtures were preincubated in the presence of Mg^{2+} before mixing with dNTP to start the reaction. In both cases, the final total $MgCl_2$ 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

Oligo name	Sequence (5'→3')	Extinction coefficient, 260 nm (M^{-1} cm^{-1})
27	CCGTCGCAGCCGTCCAACCAACTCAAC	245,700
27_{dd}	CCGTCGCAGCCGTCCAACCAACTCAAC$_{dd}$	245,700
FAM-27	[6-FAM]-CCGTCGCAGCCGTCCAACCAACTCAAC	266,660
28	CCGTCGCAGCCGTCCAACCAACTCAACA	259,500
[α-^{32}P]-28	CCGTCGCAGCCGTCCAACCAACTCAAC*A	259,500
45-18T	GGACGGCATTGGATCGATGTTGAGTTGGTTGGACGGCTGCGACGG	433,700

*^{32}P.

Table 3
dsDNA substrates used in kinetics experiments

Name	Sequence
27/45 (DNA)	5'-CCG TCG CAG CCG TCC AAC CAA CTC AAC-3' 3'-GGC AGC GTC GGC AGG TTG GTT GAG TTG TAG CTA GGT TAC GGC AGG-5'
27_{dd}/45 (DNA_{dd})	5'-CCG TCG CAG CCG TCC AAC CAA CTC AAC_{dd}-3' 3'-GGC AGC GTC GGC AGG TTG GTT GAG TTG TAG CTA GGT TAC GGC AGG-5'
28/45	5'-CCG TCG CAG CCG TCC AAC CAA CTC AAC A-3' 3'-GGC AGC GTC GGC AGG TTG GTT GAG TTG TAG CTA GGT TAC GGC AGG-5'
[α-^{32}P]-28/45	5'-CCG TCG CAG CCG TCC AAC CAA CTC AAC* A-3' 3'-GGC AGC GTC GGC AGG TTG GTT GAG TTG TAG CTA GGT TAC GGC AGG-5'
FAM-27/45 (FAM-DNA)	[6-FAM]-CCG TCG CAG CCG TCC AAC CAA CTC AAC-3' 3'-GGC AGC GTC GGC AGG TTG GTT GAG TTG TAG CTA GGT TAC GGC AGG-5'

*^{32}P.

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-PP_i (Sigma-Aldrich) was prepared, aliquoted, and stored at −80 °C. PP_i was incubated in the absence of Mg^{2+} in the mixture containing T7 DNA polymerase for solubility reasons. Control stopped flow experiments (not shown) indicated that PP_i does not bind to the enzyme without Mg^{2+}, so mixing Mg^{2+} from the other syringe starts the reaction and allows higher concentrations to be reached for short reaction times before Mg^{2+}-PP_i precipitates. For the experiment with [α-^{32}P]-28/45 DNA in Figure 4*A* and the nucleotide dissociation rate experiment in Figure 2*E*, the formation of [α-^{32}P]-dATP was monitored by TLC on PEI-cellulose plates. To remove contaminants, TLC plates were first developed in ultrapure H_2O. 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 KPO_4 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_{-i} is the reverse rate constant for the ith 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_1 is the equilibrium constant for ground-state nucleotide binding (rapid equilibrium, $k_{-1} >> k_2$), 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 PP_i 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_{cat}, 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 *E.D* complex. Rapid-quench experiments were modeled simply as the sum of species containing product ($FP.PP_i$+EP'+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 3*C* (7D) was modeled as:

$$a * (E + ED' + ED + EDA + EP' + EP + b * (FDA+FP.PP_i))$$

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 χ^2 value. The confidence interval is defined based on a threshold in χ^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.

$$rate = A \cdot \frac{k_B T}{h} \exp(-\Delta G^{\ddagger}/RT)$$
$$0 < A \leq 1$$

where k_B is the Boltzman 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 PP_i (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.

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: 257–269
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, I., 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, J., 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. Doublié, 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 A resolution. *Nature* **391**, 251–258
14. Patel, S. S., Wong, I., 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 phosphoryl 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

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. Kirmizialtin, 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. Kirmizialtin, 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 selection or induced fit: a flux description of reaction mechanism. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 13737–13741
26. Rothwell, P. J., and Waksman, G. (2007) A pre-equilibrium before nucleotide binding limits fingers subdomain closure by Klentaq1. *J. Biol. Chem.* **282**, 28884–28892
27. Guo, Q., and Sousa, R. (2006) Translocation by T7 RNA polymerase: a sensitively poised Brownian ratchet. *J. Mol. Biol.* **358**, 241–254
28. Gelles, J., and Landick, R. (1998) RNA polymerase as a molecular motor. *Cell* **93**, 13–16
29. Johnson, K. A., Simpson, Z. B., and Blom, T. (2009) Global kinetic explorer: a new computer program for dynamic simulation and fitting of kinetic data. *Anal. Biochem.* **387**, 20–29
30. Johnson, K. A. (2009) Fitting enzyme kinetic data with KinTek global kinetic explorer. *Methods Enzymol.* **467**, 601–626
31. Tsai, Y.-C., Jin, Z., and Johnson, K. A. (2009) Site-specific labeling of T7 DNA polymerase with a conformationally sensitive fluorophore and its use in detecting single-nucleotide polymorphisms. *Anal. Biochem.* **384**, 136–144
32. Luo, G., Wang, M., Konigsberg, W. H., and Xie, X. S. (2007) Single-molecule and ensemble fluorescence assays for a functionally important conformational change in T7 DNA polymerase. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 12610–12615
33. Balbo, P. B., Wang, E. C.-W., and Tsai, M.-D. (2011) Kinetic mechanism of active site assembly and chemical catalysis of DNA polymerase β. *Biochemistry* **50**, 9865–9875
34. Atis, M., Johnson, K. A., and Elber, R. (2017) Pyrophosphate release in the protein HIV reverse transcriptase. *J. Phys. Chem. B* **121**, 9557–9565
35. Li, A., Gong, S., and Johnson, K. A. (2016) Rate-limiting pyrophosphate release by HIV reverse transcriptase improves fidelity. *J. Biol. Chem.* **291**, 26554–26565
36. Yin, Y. W., and Steitz, T. A. (2004) The structural mechanism of translocation and helicase activity in T7 RNA polymerase. *Cell* **116**, 393–404
37. Fiala, K. A., Sherrer, S. M., Brown, J. A., and Suo, Z. (2008) Mechanistic consequences of temperature on DNA polymerization catalyzed by a Y-family DNA polymerase. *Nucleic Acids Res.* **36**, 1990–2001
38. Li, A., Ziehr, J. L., and Johnson, K. A. (2017) A new general method for simultaneous fitting of temperature and concentration dependence of reaction rates yields kinetic and thermodynamic parameters for HIV reverse transcriptase specificity. *J. Biol. Chem.* **292**, 6695–6702
39. Bochner, B. R., and Ames, B. N. (1982) Complete analysis of cellular nucleotides by two-dimensional thin layer chromatography. *J. Biol. Chem.* **257**, 9759–9769
40. Heinonen, J. (1974) Intracellular concentration of inorganic pyrophosphate in the cells of Escherichia coli: a method for its determination. *Anal. Biochem.* **59**, 366–374
41. Foster, B. M., Rosenberg, D., Salvo, H., Stephens, K. L., Bintz, B. J., 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