Role of a GAG Hinge in the Nucleotide-induced Conformational Change Governing Nucleotide Specificity by T7 DNA Polymerase*

Zhinan Jin1 and Kenneth A. Johnson2

From the Department of Chemistry and Biochemistry, Institute of Cellular and Molecular Biology, University of Texas, Austin, Texas 78712

A nucleotide-induced change in DNA polymerase structure governs the kinetics of polymerization by high fidelity DNA polymerases. Mutation of a GAG hinge (G542A/G544A) in T7 DNA polymerase resulted in a 1000-fold slower rate of conformational change, which then limited the rate of correct nucleotide incorporation. Rates of misincorporation were comparable to that seen for wild-type enzyme so that the net effect of the mutation was a large decrease in fidelity. We demonstrate that a presumably modest change from glycine to alanine 20 Å from the active site can severely restrict the flexibility of the enzyme structure needed to recognize and incorporate correct substrates with high specificity. These results emphasize the importance of the substrate-induced conformational change in governing nucleotide selectivity by accelerating the incorporation of correct base pairs but not mismatches.

DNA replication demands extraordinary specificity and efficiency in that polymerases have to recognize and incorporate one correct nucleotide of the pool of four nucleotides with similar structures and properties. T7 DNA polymerase replicates DNA at a rate of 300 bases/s, making only one error in every ~500,000 bases (1, 2). Like many enzymes, T7 DNA polymerase shows a large change in structure from an open to a closed state upon binding substrate (3). Kinetic data have shown that the conformational change governs nucleotide selectivity through a novel induced fit-type mechanism (4, 5).

For the past 2 decades, there has been considerable controversy regarding the role of induced fit in enzyme specificity (6, 7). In the DNA polymerase field, the debate has focused mainly on whether the substrate-induced conformational change is rate-limiting and whether it can contribute to specificity under any circumstances. Earlier studies based upon thio-elemental effects suggested that there is a rate-limiting conformational change step preceding chemistry (1, 2, 8–10). However, kinetic studies on DNA polymerase β using the fluorescence signal from 2-aminopurine and on Taq DNA polymerase I using FRET suggested that the open-to-closed conformational change was not rate-limiting (11–13). Studies on T7 DNA polymerase using a conformationally sensitive fluorophore revealed that the forward rate of the conformational change was faster than the chemistry step (4). However, the more important discovery from this study was that the reverse rate of the conformational change was much slower than the chemistry step, and therefore, only the forward rate of the conformational change and the ground state nucleotide binding $K_d$ dictated the specificity for correct nucleotide incorporation.

The important role of a conformational change in specificity and the proposed “new paradigm” for enzyme specificity may also apply to other enzymes (14). Work on HIV reverse transcriptase has also shown that the conformational change was fast and thermodynamically favorable (14). Because the subsequent chemistry step was faster than nucleotide release, the conformational change step committed the correct substrate to be incorporated, thereby defining the specificity constant. The rate of the chemistry step does not contribute to specificity as long as it is faster than the rate of nucleotide release.

It is largely unknown how enzyme dynamics at different levels and on different time scales are correlated with catalysis. There are several different modes of motion and flexibility observed in enzymes, such as molecular tumbling, “breathing,” rotation of amino acid side chains, shear and hinge motions of domains, flexible motions of loops, and thermally driven vibrations of atoms. Glycine residues provide the flexibility necessary for optimal enzyme activity and are often associated with flexible loops. There are eight glycine residues in the recognition domain of T7 DNA polymerase (residues 475–560), and all but one occur in loops connecting helices. We were particularly intrigued by two Gly-Ala-Gly (GAG) sequences flanking the O1 helix. Although each GAG is ~20 Å from the reaction center, the pair of glycine residues could contribute to the flexibility of the recognition domain and may be essential for nucleotide specificity. We speculated that these two GAG loops may serve as hinges to allow flexibility within the fingers domain.

In this study, the conformational change reported by a conformationally sensitive fluorophore was characterized in enzymes containing mutations in the GAG loops of the fingers domain of T7 DNA polymerase. Fig. 1 shows the structure of...
FIGURE 1. Conformational change in T7 DNA polymerase. The structures of T7 DNA polymerase in the open (blue) and closed (green) states are shown, drawn from Protein Data Bank files 1tk5 and 1tk0, respectively (15). The Gly542-Gly544 hinge is shown in red. Active site residues His500, Arg518, Lys522, Tyr526, and Tyr530 are shown in CPK green. The template strand is shown in cyan, the primer strand is green, and the incoming nucleotide is magenta. The site of labeling with MDCC (E514C) is shown in yellow, and the two metal ions are labeled A and B.

the T7 DNA polymerase in the open and closed states obtained in the absence and presence of nucleotide (15). The mutations of these glycine residues to alanine decreased the forward rate of the conformational change ~1000-fold. Our data suggest that the conformational change reported by the fluorophore attached to the fingers domain corresponds to the large rotation of three catalytically important helices, which contributes significantly to the specificity of correct nucleotide incorporation.

EXPERIMENTAL PROCEDURES

Construction, Expression, Purification, and Fluorescent Labeling of the Hinge Mutants—The double hinge mutant (G531A/G533A/G542A/G544A) and the single hinge mutant (G542A/G544A) were constructed by the PCR method from a plasmid, pG5X-T7-8Cys-light-exo-/, which encodes T7 DNA polymerase Cys-light mutant (C20S/C88A/C275A/C313A/C451S/C660A/C688A/C703A/E514C) as described by Tsai and Johnson (4, 16). The mutants also contained two mutations, D5A/E7A, to reduce exonuclease activity. The final constructs of both mutants were confirmed by DNA sequencing. The hinge mutants were expressed, purified, and fluorescently labeled with 7-diethylamino-3-(((2-maleimidyl-ethyl)amino)carbonyl) coumarin (MDCC)3 (Invitrogen), as described by Tsai et al. (16) with some modifications. After the addition of 0.4 mM isopropyl 1-thio-β-d-galactopyranoside and 12.5 μg/mL chloramphenicol to induce protein expression, the culture was grown overnight at 14 °C to increase protein solubility. Another change was that 0.5 mM NaCl was added to the cell lysate before the addition of 0.5% polyethyleneimine to increase the amount of soluble protein.

3 The abbreviations used are: MDCC, 7-diethylamino-3-(((2-maleimidyl-ethyl)amino)carbonyl) coumarin; dCTPαS, 1-thio-2′-deoxycytidine-5′-triphosphate; Sp isomer.

Synthetic Oligonucleotides—DNA substrates were obtained from Integrated DNA Technologies, Inc. (Coralville, IA) and purified by using 15% polyacrylamide, 7 M urea denaturing gel electrophoresis. A 27-mer primer (5′-GCC TCG CAG CCG TCC AAC CAA CTC AA-3′) and a 45-mer template (5′-GGA CGG CAT TGG ATC GAG GTT GAG TGG TGA GGA CGG CTG CGA GGC-3′) were adopted from a previous study to preserve the comparison between all relative studies (the underscore indicates the templating base). There were two versions of the 27-mer primer used in this study. For nucleotide binding studies, a ddCMP-terminated 27-mer primer was used to prevent incorporation of the next incoming nucleotide, dCTP. The dCMP-terminated primer was used for single nucleotide incorporation experiments and was 5′-32P-labeled using [γ-32P]ATP (PerkinElmer Life Sciences) and T4 polynucleotide kinase according to the manufacturer’s instructions (Invitrogen). The labeling reaction was terminated by heating at 95 °C for 5 min, and excess [γ-32P]ATP was removed with a Bio-spin 6 desalting column (Bio-Rad).

To make a DNA duplex, primer and template were mixed at a 1:1 molar ratio in a buffer containing 6 mM Tris-Cl (pH 7.5), 6 mM NaCl, and 0.2 mM EDTA. The mixture was heated at 95 °C for 2 min and allowed to cool slowly to room temperature.

Chemical Quench Experiments—Single turnover single nucleotide incorporation assays were performed at 20 °C in T7 reaction buffer containing 40 mM Tris (pH 7.5), 1 mM EDTA, 50 mM NaCl, 1 mM DTT, and 12.5 mM MgCl2. For dCTP and dCTPαS incorporation assays, an RQF-3 rapid quench-flow apparatus (KinTek Corp., Austin, TX) was used. One sample loop was loaded with preformed enzyme-DNA complex (600 nM enzyme, 12 μM thiorodoxin, and 200 nM 5′-32P-labeled 27-mer/45-mer DNA duplex) in T7 reaction buffer without MgCl2. Another loop was loaded with nucleotide in T7 reaction buffer containing 25 mM MgCl2. The reaction was started by rapidly mixing the two reactants and then was quenched by mixing with 0.5 mM EDTA after time intervals ranging from several ms to several s.

Nucleotide misincorporation assays were performed by manual mixing due to slower reaction. Preformed enzyme-DNA complex (600 nM enzyme, 12 μM thiorodoxin, and 200 nM 5′-32P-labeled 27-mer/45-mer DNA duplex) in T7 reaction buffer without MgCl2 was mixed with an equal volume of nucleotide solution in T7 reaction buffer containing 25 mM MgCl2. The reaction was started by rapidly mixing the two reactants and then was quenched by mixing with 0.5 mM EDTA after time intervals ranging from several ms to several s.

Pyrophosphorolysis assays were set up as described for nucleotide misincorporation assays, except that pyrophosphate was added instead of nucleotide. The products from the assays described above were resolved on 15% denaturing polyacrylamide sequencing gels, and then the dried gels were exposed to phosphor screens. The screens were scanned on a Storm 860 scanner (GE Healthcare). The amount of product formation at each time interval was analyzed using the ImageQuant software (GE Healthcare).

Fluorescence Emission Spectra—Emission spectra were recorded using a PTI fluorimeter (Photon Technology Interna-
Role of GAG Hinge in Polymerase Specificity

Conventional Data Fitting—The kinetic data of nucleotide incorporation were analyzed by nonlinear regression using GraFit 5 software (Erithacus Software, Surrey, UK). Each time course of single nucleotide incorporation at various nucleotide concentrations was fit to a single exponential equation, \( Y = A \exp(-k_{obs} t) + C \), where \( A \) is the amplitude, \( k_{obs} \) is the observed rate, and \( C \) is the end point. The concentration dependence of the rate of incorporation was fit to a hyperbola, \( Y = k_{pol} S_o/(K_d + S_o) \), where \( k_{pol} \) is the maximum incorporation rate, \( S_o \) is the variable starting nucleotide concentration, and \( K_d \) is the apparent equilibrium dissociation constant for nucleotide ground state binding. The pyrophosphorolysis data were analyzed similarly.

Each time course of fluorescence change from the stopped-flow experiments was initially fit to a single exponential equation to extract the rate and amplitude of the change. The data from nucleotide release were fit to a single exponential equation. The data from the titration experiment were fit to a hyperbolic equation, \( F = F_0 + \Delta F S_o/(K_d + S_o) \), where \( F_0 \) is the starting fluorescence, \( \Delta F \) is the fluorescence change, and \( S_o \) is the substrate concentration.

Global Data Fitting—Fluorescence stopped-flow data were fit globally by regression analysis based upon numerical integration of the rate equations using the program KinTek Explorer version 2.2 (KinTek Corp.) (17, 18). Data were fit to the reaction shown in Scheme 1 with the inclusion of DNA binding and release steps using known rates (10 \( \mu M^{-1} s^{-1} \) on rate and 0.2 \( s^{-1} \) off rate) (1). Fluorescence scaling factors were included in the data fitting process by defining an output expression \( F = scale(E + aED + bEDN + cFDN) + bkg \), where \( E, ED, EDN, \) and \( FDN \) are enzyme forms defined according to Scheme 1. The fluorescence factors for fitting the dCTP binding data were \( a = 1.44, b = 1.38, \) and \( c = 0.75 \). The fluorescence factors for fitting the dCTPαS binding data were \( a = 1.31, b = 1.28, \) and \( c = 0.7 \). The small decrease in fluorescence in the absence of nucleotide in each case was attributable to equilibration of the enzyme-DNA complex following a 1:1 dilution at the start of the reaction. In other words, the simulation was started by equilibrating the enzyme and DNA (600 nM enzyme, 200 nM DNA), and then the solution was diluted 2-fold upon mixing with the nucleotide. The rate of the nucleotide-induced conformational change \( (k_d) \) was initially constrained to be equal to the rate of incorporation \( (0.5 \text{ s}^{-1}) \) and was then adjusted in fitting once the scaling factors were established. A small normalization factor (1–2%) was applied to the family of curves describing the concentration dependence of nucleotide binding (see Figs. 4A and 5A), based upon the optimal global fit to the data as described in the instruction manual for the software. Estimates for errors on parameters were derived by nonlinear regression and by FitSpace confidence contour analysis using KinTek Explorer (17).

RESULTS

Generation of the Hinge Mutants—To characterize the conformational change during nucleotide binding and incorporation, mutants were constructed in the background of a Cys-light T7 DNA polymerase constructed by Tsai et al. (16) and labeled with MDCC. The double mutation, D5A/E7A, reduced the exonuclease activity. The labeled Cys-light mutant showed nearly full activity with only modest changes in the kinetic parameters compared with the wild type (exo−) T7 DNA polymerase (16).

We initially examined a double hinge mutant (G531A/G533A/G542A/G544A) in which four glycines within the GAG hinges were mutated to alanines. A single nucleotide incorporation assay showed that this mutant had lost all detectable catalytic ability (data not shown). Inspection of the structure suggested that the two mutations, G531A and G533A, might sterically interfere with the binding of the DNA template. Because this mutant was not active, further characterization was not performed. We made a single GAG hinge mutant (G542A/G544A), which was then labeled with MDCC at Cys514. Nucleotide incorporation burst assays of this mu-
tant before and after labeling showed that it was active and that the fluorescent labeling did not change the burst rate of this mutant significantly (data not shown). The remaining studies were performed using this G542A/G544A mutant, which we refer to as the “hinge mutant.”

Correct Nucleotide Incorporation—To measure the rate of incorporation of a correct nucleotide, single turnover quencher-flow experiments were performed by mixing an enzyme-DNA complex with a solution containing the correct nucleotide, dCTP. The time course of product formation at each dCTP concentration was fit to a single exponential equation to obtain the rate of nucleotide incorporation. The dCTP concentration dependence of the rate of incorporation was then fit to a hyperbola, yielding a maximum nucleotide incorporation rate, $k_{\text{pol}} = 0.52 \pm 0.02 \text{ s}^{-1}$, and an apparent nucleotide dissociation constant, $K_d = 1.7 \pm 0.2 \mu M$ (Fig. 2A). Compared with the previously reported values of $k_{\text{pol}} = 234 \pm 94 \text{ s}^{-1}$ and $K_d = 24 \pm 3.1 \mu M$ for the MDCC-labeled T7 DNA polymerase (4), the hinge mutant showed about a 450-fold decrease in nucleotide incorporation rate but only a 30-fold decrease in $k_{\text{cat}}/K_m (k_{\text{pol}}/K_d)$.

Because the mutations at residues 542 and 544 are remote from the active site and not directly involved in catalysis, we proposed that these mutations may slow the conformational change step. As one approach to examine whether the chemistry step is rate-limiting, we measured the thio-elemental effect on nucleotide incorporation. The thio-elemental effect was defined by the rate of incorporating a normal nucleotide divided by the rate of incorporating a corresponding α-thio-substituted nucleotide (with a pro-Sp non-bridging oxygen of the α-phosphate replaced by a sulfur) ($k_{\text{pol,dNTP}}/k_{\text{pol,dNTTP}}$). Because the sulfur decreases the rate of phosphoryl transfer due to its decreased electronegativity compared with oxygen, a thio-elemental effect will be observed if the chemistry step is rate-limiting (1, 2, 9, 19). We studied the dCTPαS concentration dependence of incorporation rates using single turnover quencher-flow methods (Fig. 2B). The data were fit to a hyperbola, defining a $k_{\text{pol}} = 0.60 \pm 0.04 \text{ s}^{-1}$, and a $K_d = 2.4 \pm 0.4 \mu M$. Because $k_{\text{pol,dCTP}}/k_{\text{pol,dCTPαS}} = 0.52/0.6 \approx 0.9$, no thio-elemental effect was evident. This suggested that the rate-limiting step is not the chemistry step but may be due to a conformational change step preceding chemistry.

Kinetics of Correct Nucleotide Binding—To test whether the conformational change was the rate-limiting step, we monitored the nucleotide-induced fluorescence change using MDCC-labeled enzyme. Fluorescence emission scans were performed to examine the MDCC-labeled hinge mutant with different ligands bound (Fig. 3). In these experiments, duplex DNA with dideoxy-terminated primer (DNA$_{dd}$) was used to prevent the chemistry step, only the binding of the nucleotide would be observed. Upon binding a correct nucleotide (dCTP) to the enzyme-DNA$_{dd}$ complex, the intensity of the fluorescence decreased about 30%. In contrast, after binding of an incorrect nucleotide (dGTP) an 18% increase of the fluorescence intensity was observed. The fluorescence emission spectra of the hinge mutant were similar to those of the Cys-light T7 DNA polymerase in various nucleotide-binding states (i.e. fluorescence decreased upon binding a correct nucleotide but increased upon binding an incorrect nucleotide). This suggested that the fluorescence signal that reported the conformational change for the Cys-light T7 DNA polymerase was not distorted significantly by the hinge mutations.

Role of GAG Hinge in Polymerase Specificity

**FIGURE 2. Kinetics of dCTP and dCTPαS incorporation by the hinge mutant.** A, an enzyme-DNA complex (600 nM enzyme and 200 nM DNA) was rapidly mixed (1:1) with dCTP solution at various concentrations to start the reaction. The time course of product formation at each dCTP concentration was fit to a single exponential to obtain the incorporation rate. The dCTP concentration dependence of incorporation rates was fit to a hyperbola, defining an apparent $K_d = 1.7 \pm 0.2 \mu M$ and a maximal incorporation rate $k_{\text{pol}} = 0.52 \pm 0.02 \text{ s}^{-1}$. B, the dCTPαS concentration dependence of nucleotide incorporation rates defined an apparent $K_d = 2.4 \pm 0.4 \mu M$ and a maximal incorporation rate $k_{\text{pol}} = 0.60 \pm 0.04 \text{ s}^{-1}$. Error bars, S.E.

**FIGURE 3. Fluorescence changes observed upon nucleotide binding.** Fluorescence emission spectra of MDCC-labeled hinge mutant enzyme with different substrate-bound states are shown. Fluorescence was excited at 425 nm. Concentrations are as follows: 200 nM enzyme, 4 μM thioredoxin, and 300 nM DNA$_{dd}$ (with ddCMP-terminated 27-mer primer), with and without 500 μM nucleotide as shown.
Role of GAG Hinge in Polymerase Specificity

The kinetics of the fluorescence change were measured by stopped-flow methods. The preformed enzyme-DNA_{dd} complex was rapidly mixed with a solution containing dCTP at various concentrations in the stopped-flow instrument, and the fluorescence changes following mixing were recorded. The time course of each stopped-flow reaction was fit to a single exponential to estimate the rate and amplitude of the reaction. At all concentrations, the time course was adequately fit to a single exponential function. Attempts to fit the data to a double exponential function did not produce consistent or meaningful results. The amplitude and rate of the observed transients were then plotted as a function of nucleotide concentration. The dCTP concentration dependence of the amplitude was fit to a hyperbola, defining an apparent dissociation constant, \( K_{d,\text{app}} = 0.52 \pm 0.04 \text{ \mu M} \) (Fig. 4B). The observed rate (Fig. 4C) first increased and then decreased as a function of increasing nucleotide concentration, but the extent of variation was small (1.3–1.6 s\(^{-1}\)). Regardless of the complexities underlying this unusual concentration dependence, the observation that the rate approaches a constant value, independent of concentration, implies that the binding reaction involves more than one step.

We measured the rate of release of bound dCTP from the preformed enzyme-DNA_{dd}-dCTP ternary complex (Fig. 5B). In the stopped-flow instrument, a solution of labeled enzyme-DNA_{dd}-dCTP complex was rapidly mixed with a solution containing a 4-fold excess of unlabeled enzyme-DNA complex, formed with an unlabeled wild type (exo\(^-\)) T7 DNA polymerase and a DNA with normal primer that can incorporate dCTP at about 300 s\(^{-1}\). This served as a trap so that all of the free dCTP could be removed rapidly from solution. We observed an increase in fluorescence due to release of dCTP from the labeled enzyme, which was fit to a single exponential equation, yielding a rate of 1.1 ± 0.02 s\(^{-1}\). This is similar to \( k_2 = 1.6 \text{ s}^{-1}\) for the Cys-light T7 DNA polymerase (4).

We fit the data to a two-step binding mechanism shown in Scheme 1, where \( ED_{dd} \) represents the open enzyme-DNA complex (with dideoxy-terminated DNA), and \( FD_{dd} \) represents the closed (reduced fluorescence) state that is formed after nucleotide binding.

\[
\begin{align*}
ED_{dd} + dCTP &\xrightarrow{k_1} ED_{dd}dCTP \\
&\xrightarrow{k_2} FD_{dd}dCTP
\end{align*}
\]

According to this scheme, the observed rate will approach a maximum value defined by the sum \( k_2 + k_{-2}\). With an observed rate of 1.5 s\(^{-1}\) and an estimate of \( k_{-2} = 1.1 \text{ s}^{-1}\), these data suggest a value of 0.4 s\(^{-1}\) for \( k_2\), which is comparable with the rate of nucleotide incorporation (0.5 s\(^{-1}\)). These data imply that the conformational change limits the rate of incorporation. Moreover, the concentration dependence of the observed rate and amplitude imply that the enzyme isomerization from the open to the closed state (step 2) is unfavorable thermodynamically.

The net equilibrium dissociation constant was measured by a titration experiment (Fig. 5C). A solution of the enzyme-DNA_{dd} complex was constantly stirred with the addition of a solution of dCTP at a slow constant rate, and the fluorescence intensity was recorded continuously. The dCTP concentration dependence of the fluorescence change from this titration experiment was fit to a hyperbola, defining the net equilibrium dissociation constant for two steps of nucleotide binding, \( K_{d,\text{net}} = 1.1 \pm 0.1 \text{ \mu M} \). This apparent dissociation constant is an estimate of a net equilibrium dissociation constant, \( K_{d,\text{net}} = K_{d,1}/(1 + K_2)\), where \( K_{d,1}\) is the equilibrium dissociation constant for nucleotide ground state binding, and...
Role of GAG Hinge in Polymerase Specificity

$K_2$ is the equilibrium constant of the second step according to a two-step binding mechanism shown in Scheme 1. With $k_2 = 0.5 \text{ s}^{-1}$, $k_{-2} = 1.1 \text{ s}^{-1}$, one can estimate $K_{d,1} = 1.6 \mu M$.

The net result from this conventional data fitting was that one could derive approximate estimates for $K_1$, $k_2$, and $k_{-2}$ and account for most aspects of the data. However, one could not fully explain the strange concentration dependence of the rate (Fig. 4C), and the rate constants, $k_1$ and $k_{-1}$, could not be resolved. A decade ago, this method of data fitting and analysis would have been considered to be adequate, but it falls short relative to the comprehensive data fitting afforded by methods based upon numerical integration of the rate equations and the fitting of raw data directly to the model.

Global Data Fitting by Computer Simulation—In order to more rigorously analyze the data and test the model, the time courses of the fluorescence changes were fit globally to the model shown in Scheme 1. Global data fitting directly to the model offers numerous advantages in that it bypasses all simplifying assumptions required to derive mathematical functions used in conventional fitting, and the concentration dependence of the rate and amplitude are fit simultaneously across the concentration series in modeling the data (18). In addition, we simultaneously fit the time course for nucleotide binding (Fig. 5A) and dissociation (Fig. 5B).

In fitting the data globally, we initially constrained $k_2$ to be greater than or equal to the observed rate of incorporation, $0.5 \text{ s}^{-1}$, in order to derive initial estimates for the fluorescence scaling factors. The fluorescence change was modeled as $F = s(E + aED + bEDdCTP + cFDdCTP)$, where $s$ is an arbitrary scaling factor that was different for each experiment, whereas $a = 1.44$, $b = 1.38$, and $c = 0.75$ were globally derived parameters defining the relative fluorescence change in forming $ED_{dd}$, $ED_{dd}dCTP$, and $FD_{dd}dCTP$, respectively.

Simultaneously fitting the data shown in Fig. 5A and B, yielded the rate constants summarized in Scheme 2 and Table 1.

$$
ED_{dd} + dCTP \rightleftharpoons 7.7 \mu M^{-1}s^{-1} ED_{dd}dCTP
$$

$$
ED_{dd}dCTP \rightarrow 0.5 s^{-1} FD_{dd}dCTP
$$

$$
ED_{dd}dCTP \rightarrow 1.2 s^{-1} ED_{dd}dCTP
$$

SCHEME 2

According to the mechanism shown in Scheme 2, the calculated $K_{d,net} = K_{d,1}/(1 + K_3) = 1.3/(1 + 0.5/1.2) = 0.9 \pm 0.1 \mu M$ and thus agrees very well with the results of the titration experiment (Fig. 5C) and is comparable with the estimate of $K_d \sim 0.5 \mu M$ from the concentration dependence of the amplitude of the stopped flow derived by conventional fitting (Fig. 4B).

Fitting the data by computer simulation involved a total of eight independent parameters (four rate constants and four fluorescence scaling factors). This is far fewer than the 30 independent parameters employed in conventional data fitting (a rate, amplitude, and end point for each of the nine traces plus $K_1$, $k_2$, and $k_{-2}$). Nonetheless, it is important to question...
Role of GAG Hinge in Polymerase Specificity

TABLE 1
Kinetics of nucleotide binding and enzyme isomerization for the hinge mutant

The rate constants governing nucleotide binding and enzyme isomerization were derived by globally fitting the time course of the fluorescence changes during nucleotide binding and release observed with dCTP and dCTPαS (Figs. 5 and 7, respectively) as described under “Experimental Procedures.” Error limits were derived by confidence contour analysis (Fig. 6) (17). The observed $K_d$ was obtained by equilibrium titration, whereas the calculated $K_d$ was defined by $K_d = k_{-1}/k_1$. The numbering is according to Scheme 1.

| dNTP  | $k_1$  | $k_{-1}$ | $k_2$  | $k_{-2}$ | $K_d$ | $K_d$net |
|-------|--------|----------|--------|----------|-------|----------|
| dCTP  | 7.7 ± 2 s$^{-1}$ | 10 ± 2.5 s$^{-1}$ | 0.5 ± 0.1 | 1.2 ± 0.02 s$^{-1}$ | 0.9 ± 0.1 | 1.1 ± 0.1 μM |
| dCTPαS | 6.3 ± 2 s$^{-1}$ | 26 ± 5 s$^{-1}$ | 0.5 ± 0.1 | 1.4 ± 0.04 s$^{-1}$ | 2.8 ± 0.3 | 2.4 ± 0.3 μM |

TABLE 2
Fidelity of nucleotide incorporation by the hinge mutant

The kinetics of nucleotide incorporation were quantified under single turnover conditions to define the apparent $K_d$ and maximum rate of incorporation, $k_{pol}$. The discrimination was calculated from the ratio of $k_{pol}/K_d$ for the correct nucleotide divided by that for the mismatched nucleotide. S.E. values were derived by nonlinear regression based upon global fits to the family of curves showing product versus time for various nucleotide concentrations.

| dNTP  | $k_{pol}$ | $k_d$ | $k_{pol}/K_d$ | Discrimination |
|-------|----------|-------|---------------|---------------|
| dCTP  | 0.52 ± 0.02 s$^{-1}$ | 1.7 ± 0.2 s$^{-1}$ | 0.31 ± 0.04 s$^{-1}$ | 1 |
| dATP  | 0.13 ± 0.02 s$^{-1}$ | 340 ± 75 s$^{-1}$ | 0.0004 ± 0.0001 s$^{-1}$ | 800 ± 200 |
| TTP   | 0.13 ± 0.02 s$^{-1}$ | 750 ± 170 s$^{-1}$ | 0.0002 ± 0.00005 s$^{-1}$ | 1800 ± 500 |
| dGTP  | 0.07 ± 0.007 s$^{-1}$ | 1020 ± 40 s$^{-1}$ | 0.000007 ± 0.0000007 s$^{-1}$ | 4500 ± 700 |

Although there are insufficient data to define a rate of the fast phase by conventional fitting methods.

Thus, the complete model shown in Scheme 2 with the rate constants described can adequately account for the data, including the concentration dependence for the rate and amplitude of the reaction. Moreover, the four rate constants are sufficiently constrained by the data with error limits as summarized in Table 1, derived by confidence contour analysis (17).

Kinetics of dCTPαS Binding—The kinetics of the conformational change upon binding dCTPαS were also examined by stopped-flow methods. Fig. 7A shows the recorded fluorescence transients after mixing the enzyme-DNA$\alpha d$ complex with dCTPαS at various concentrations. Similar to the binding of dCTP, the amplitude increased, and the observed rate of the fluorescence transient appeared to decrease with increasing dCTPαS concentration. We also measured the rate of dCTPαS release from the enzyme-DNA$\alpha d$-dCTPαS ternary complex (Fig. 7B). The time course of fluorescence increase corresponding to dCTPαS release was fit to a single exponential equation defining a dissociation rate of 1.36 ± 0.01 s$^{-1}$. Global fitting of the stopped-flow data gave the pathway for dCTPαS binding shown in Scheme 3.

$$ED_\alpha + dCTP\alpha S \overset{6.3 \mu M^{-1}s^{-1}}{\underset{26 s^{-1}}{\rightleftharpoons}} ED_\alpha dCTP\alpha S$$

$$ED_\alpha dCTP\alpha S \overset{0.5 s^{-1}}{\underset{1.4 s^{-1}}{\rightleftharpoons}} FD_\alpha dCTP\alpha S$$

Scheme 3

The net equilibrium dissociation constant was measured by a titration experiment under the same condition for dCTP (Fig. 7C). The dCTPαS concentration dependence of the fluorescence change from this titration experiment was fit to a hyperbola, defining the net equilibrium dissociation constant, $K_{d,net} = 2.4 ± 0.1 \mu M$. According to the mechanism shown in Scheme 3, the calculated $K_{d,net} = K_d/(1 + K_d) = 4.1/(1 + 0.5/1.4) = 3.0 \mu M$, which agrees very well with the results of the titration experiment.

Misincorporation and Discrimination—Misincorporation of dATP, dGTP, or dTTP opposite a dG on the template was examined under single turnover conditions using the hinge mutant, as described under “Experimental Procedures.” As
shown in Fig. 8A, the time course of incorporation of dGTP followed a single exponential function at each concentration. We fit the data globally to a minimal model to define the apparent $K_d = 1.02 \pm 0.04$ mM and $k_{pol} = 0.07 \pm 0.007$ s$^{-1}$. B, fluorescence was recorded after mixing a preformed E-DNA$_{ad}$ complex (200 nM enzyme, 300 nM dideoxy-terminated DNA) with various concentrations of dGTP (62.5, 125, 250, 500, 1000, 2000, and 4000 mM). Data were fit globally to Scheme 4 as described under "Results" to get the smooth curves. Both experiments were performed with the MDCC-labeled hinge mutant.

We also examined the kinetics of the fluorescence change upon binding the mismatch (dGTP) to dideoxy-terminated DNA as shown in Fig. 8B. The data required a two-step binding mechanism and were fit globally to the model shown in Scheme 4.

---

**SCHEME 4**

$ED_{dd} + dGTP \rightarrow ED_{dd}dGTP$

$400 \mu M$

$0.86$ s$^{-1}$

$FD_{dd}dGTP$

$4.0$ s$^{-1}$

**TABLE 1**

| Concentration (mM) | Rate Constant (s$^{-1}$) |
|--------------------|--------------------------|
| 62.5               | 0.86                     |
| 125                | 4.0                      |
| 250                |                           |
| 500                |                           |
| 1000               |                           |
| 2000               |                           |
| 4000               |                           |
Attempts to fit the fluorescence and quench-flow data simultaneously failed due to the different apparent dissociation constants. The fluorescence data required a $K_d = 400 \mu M$, whereas the quench-flow required a $K_d = 1 \mathrm{mM}$. This 2.5-fold difference could be due to the use of the dideoxy-terminated DNA in the fluorescence stopped-flow experiment.

These data show that there is a thermodynamically unfavorable isomerization after dGTP binding and preceding misincorporation. If the rate of isomerization preceding dGTP incorporation is the same as seen in the fluorescence data, then the intrinsic rate constant for incorporation would be $0.37 \, \mathrm{s}^{-1}$, to account for the observed rate of $0.07 \, \mathrm{s}^{-1}$. The rate of the nucleotide-induced isomerization is about 250-fold slower for the hinge mutant compared with wild type (4), although the chemistry step for misincorporation appears to be slower than the isomerization step and is largely rate-limiting. A summary of the results for each mismatch in terms of the $K_{app}$, specificity constant, and discrimination is reported in Table 2.

**Pyrophosphorolysis**—Pyrophosphorolysis is the reverse of the nucleotide incorporation reaction whereby the polymerase catalyzes the reaction of pyrophosphate with the 3'-terminal dNMP of the primer and then releases dNTP, so the primer is shortened by one nucleotide after a single turnover. An enzyme-DNA complex was mixed with a solution containing pyrophosphate at various concentrations to start the reaction. The reaction was quenched with EDTA, and the products were analyzed on a sequencing gel. Fig. 9 shows the time course of product formation at four concentrations of pyrophosphate. The data were fit globally to a minimal model involving pyrophosphate binding and reaction, but the data could not define an apparent second order rate constant for pyrophosphate.

A solution containing 600 nM enzyme and 200 nM 27/45-mer DNA was rapidly mixed (1:1) with pyrophosphate at various concentrations (1 (◼), 2 (■), 4 (▲), and 8 (▲) mM) to start the reactions. At the times indicated, the reactions were quenched with 0.5 mM EDTA and analyzed by sequencing gel. The time courses of product (26-mer) formation were fit globally to a minimal model involving pyrophosphate binding and reaction, but the data could not define an apparent second order rate constant for pyrophosphate reaction.

**DISCUSSION**

**Effect of the Hinge Mutant**—Our data show that the GAG hinge mutation slows the rate of the nucleotide-induced conformational change, which then becomes the rate-limiting step for single nucleotide incorporation of a correct base pair. The negligible thio-elemental effect observed for correct nucleotide incorporation suggests that a step preceding chemistry limits the rate of reaction at $0.5 \, \mathrm{s}^{-1}$. Global data fitting directly to the model provided a quantitative explanation for the observed fluorescence data. Combined with the data measured from the quench-flow experiments, we assigned the kinetic parameters to each step in the pathway shown in Scheme 2, which agrees with the observed net equilibrium dissociation constant derived from the titration experiment (Fig. 5). Therefore, all the data are consistent with the mechanism shown in Scheme 2, in which the conformational change reported by MDCC is rate-limiting. Other rate constants along the reaction pathway, such as the chemistry step and pyrophosphate release, are not defined by the data presented here and are presumed to be at least 10-fold faster than $0.5 \, \mathrm{s}^{-1}$.

**Global Data Fitting**—Fitting by numerical integration overcomes numerous limitations of conventional data fitting to analytically derived, simplified equations. Since programs for data fitting directly to a model of interest based upon numerical integration of rate equations were first developed by Carl Frieden (20) and have subsequently been refined and improved (18, 21, 22), there has been little justification for fitting to analytical functions by conventional methods. Rather, by forcing the data fitting to account for both the rate and amplitude of observed reactions and by fitting directly to the model, fitting by simulation avoids the mistakes due to conventional data fitting that are so common in the literature.

In this study, we illustrate the complexities in fitting data to a relatively simple model, but one where the combination of rate constants gave data that were difficult to interpret by conventional methods. Note especially that the conventional fitting fails to account for the apparent rise and fall in the observed rate of the reaction as a function of nucleotide concentration (Fig. 4C). This phenomenon is due to a small amplitude of a fast phase that cannot be reliably resolved in fitting to exponential functions but at moderate nucleotide concentrations leads to an overestimation of the rate because the observed transient is an unresolved mixture of two phases. This phenomenon was also observed in early studies on actomyosin before data fitting by computer simulation was available (23). We have also illustrated the resolution of the complexities caused by a readily reversible reactions in studies on the incorporation of 8-oxo-dGTP and AZT-triphosphate by the human mitochondrial DNA polymerase (24, 25). Our recent work on HIV reverse transcriptase also illustrates the utility of globally fitting fluorescence and rapid chemical quench-flow data simultaneously to derive a single unifying model (14).
It should be noted that by fitting the nucleotide dissociation experiment (Fig. 5B) by computer simulation, we avoid one of the pitfalls of such competition experiments when fit by conventional methods. In measuring the nucleotide dissociation rate after adding a 4-fold excess of unlabeled enzyme, one would require that the rate of rebinding of the nucleotide to the labeled enzyme was insignificant in order to assume that the observed rate provided an accurate measurement of the dissociation rate derived by fitting to a single exponential. However, in computer simulation, the concentrations of all reactants are followed throughout the time course of the reaction, and the data are fit by constraining nucleotide binding and release steps to be the same for labeled and unlabeled enzyme. In principle, one could derive the dissociation rate from an experiment using equimolar concentrations of the two enzyme forms. The fact that the globally fitted rate (1.2 s⁻¹) is 10% larger than the rate derived by the fit to a single exponential (1.1 s⁻¹) is consistent with the 4-fold excess of unlabeled enzyme used in these studies. Showing that the measured binding and dissociation rate agree with the overall equilibrium constant also provided an important additional check on the validity of the individual rate measurements.

Based upon our current analysis, there can be little question that our model quantitatively accounts for all aspects of the data. That does not establish whether a more complex model would account for the data equally well. Rather, we only define the minimal kinetic model to specify the kinetically significant steps leading up to nucleotide incorporation by the hinge mutant.

The Conformational Change Reported by MDCC—The 542–544 GAG hinge is 20–25 Å away from the α-phosphate of the incoming dNTP, yet mutations of the two glycines to alanine have a profound effect on the rate of the conformational change and the net rate of nucleotide incorporation. It must be stressed that we do not assume that the MDCC fluorophore defines the kinetics of the conformational change. Rather, the data fitting demonstrates that the fluorescence change is correlated with a step preceding chemistry, both for the wild-type (Cys-light) and hinge mutant forms of the enzyme. The end points of the reaction characterized structurally as open and closed are correlated with the fluorescence states seen in the absence and presence of nucleotide. Some investigators prefer the use of FRET because then it can be assumed that the FRET change is due to a distance change. However, the magnitude of the conformational change is only 15 Å along the vector of the largest movement, and FRET is subject to an error of ~10 Å due to the unknown orientation of the absorption and emission dipoles (26). Although popular, FRET introduces two labeling sites and two opportunities for unwanted secondary effects while offering little theoretical advantage over the use of a single reporter group. Akin to studies using native protein fluorescence, we may not know the precise molecular details that are leading to a fluorescence change, but that is less important than having a signal to monitor changes in structure. Indeed, it may be that the MDCC fluorescence provides an additional readout for conformational isomerization beyond the large structural transition seen between end points defined crystallographically. For example, our recent work on using MDCC to monitor conformational dynamics of HIV reverse transcriptase has resolved two conformational transitions preceding the incorporation of the anti-HIV drug lamivudine (14).

The fluorescence change characterized in the hinge mutant appears to be similar to the one reported for the wild-type (Cys-light) T7 DNA polymerase (4) but much slower. The profiles of the fluorescence emission of the enzymes bound with a correct or incorrect nucleotide are similar in magnitude. They both showed a decrease in fluorescence upon binding a correct nucleotide and an increase upon binding an incorrect nucleotide. Therefore, the hinge mutations did not change the direction of this fluorescence change but decreased the rate 1000-fold.

Structural data suggest that the conformational change from the open to the closed complex does not occur as a rigid body movement of the fingers domain; rather, the structural change from open to closed states involves the sliding and rotation of helices within the fingers domain (15), also seen in Tag polymerase (27). It is reasonable to suppose that the MDCC fluorophore is responding to changes in local environment during these conformational changes.

Fig. 1 shows the structural comparison of the fingers domain between the open and the closed complexes of T7 DNA polymerase. The most significant changes are in the O and O1 helices. As a result of all of these changes, the surface of the fingers domain appears more solvent-accessible in the closed complex than in the open complex as judged by the distance between the Glu514 (yellow, where the MDCC is located) and the O2 helix. This observation might explain the origin of the fluorescence change because the MDCC label sits on a loop, the environment of which appears to change between the open and closed states of the enzyme. The 542–544 GAG hinge connects the O1 helix with the O2 helix. Because the 542–544 GAG hinge is remote from the active site, the G/A mutations are not expected to perturb active site residues directly. Rather, our data support a model in which the open-to-closed conformational change is complex and involves the movements of several secondary structures within the fingers domain. The conformational change appears to occur as a highly cooperative process because hindering the rotation through the 542–544 GAG hinge slows the motion of the entire domain and the alignment of residues to promote catalysis.

The kinetics of the conformational change observed with dCTP and dCTPαS are surprisingly similar (Schemes 2 and 3 and Table 1). The major differences appear in the initial binding of nucleotide, which appears to be slightly weaker with dCTPαS. Although the rate constants \( k_1 \) and \( k_{-1} \) are determined with less certainty than \( k_2 \) and \( k_{-2} \) in the global fitting, the equilibrium constant, \( K_e \), is known with greater certainty than either \( k_1 \) or \( k_{-1} \). The 2-fold weaker binding of dCTPαS may be due to minor steric effects.

Pyrophosphate binding appears to be quite weak, and the reverse chemistry step is slow. Because of this, little is known about the reactions that govern pyrophosphorolysis. It is likely that pyrophosphate binding must overcome an unfavorable equilibrium constant for DNA translocation, which fa-
Role of GAG Hinge in Polymerase Specificity

...residues when binding a mismatch. Molecular details defining how this is accomplished are just beginning to be resolved.

REFERENCES

1. Patel, S. S., Wong, I., and Johnson, K. A. (1991) Biochemistry 30, 511–525
2. Wong, I., Patel, S. S., and Johnson, K. A. (1991) Biochemistry 30, 526–537
3. Doublié, S., Tabor, S., Long, A. M., Richardson, C. C., and Ellenberger, T. (1998) Nature 391, 251–258
4. Tsai, Y. C., and Johnson, K. A. (2006) Biochemistry 45, 9675–9687
5. Johnson, K. A. (2008) J. Biol. Chem. 283, 26297–26301
6. Herschlag, D. (1988) Biolo. Chem. 16, 62–96
7. Fersht, A. R. (1999) Enzyme Structure and Mechanism, 3rd Ed., pp. 370–373, Freeman, New York
8. Kuchta, R. D., Mizrahi, V., Benkovic, P. A., Johnson, K. A., and Benkovic, S. J. (1987) Biochemistry 26, 8410–8417
9. Mizrahi, V., Henrie, R. N., Marlier, J. F., Johnson, K. A., and Benkovic, S. J. (1985) Biochemistry 24, 4010–4018
10. Purohit, V., Grindley, N. D., and Joyce, G. (2003) Biochemistry 42, 10200–10211
11. Showalter, A. K., and Tsai, M. D. (2002) Biochemistry 41, 10571–10576
12. Bakhitina, M., Roettger, M. P., Kumar, S., and Tsai, M. D. (2007) Biochemistry 46, 5463–5472
13. Rothwell, P. J., Mitakov, S., and Waksman, G. (2005) Mol. Cell 19, 345–355
14. Kelling, M. W., and Johnson, K. A. (2010) Proc. Natl. Acad. Sci. U.S.A. 107, 7734–7739
15. Brieba, L. G., Kokoska, R. J., Bebenek, K., Kunkel, T. A., and Ellenberger, T. (2005) Structure 13, 1653–1659
16. Tsai, Y. C., Jin, Z., and Johnson, K. A. (2009) Anal. Biochem. 384, 136–144
17. Johnson, K. A., Simpson, Z. B., and Blom, T. (2009) Anal. Biochem. 387, 30–41
18. Johnson, K. A., Simpson, Z. B., and Blom, T. (2009) Anal. Biochem. 387, 20–29
19. Herschlag, D., Piccirilli, J. A., and Cech, T. R. (1991) Biochemistry 30, 4844–4854
20. Barshop, B. A., Wrenn, R. F., and Frieden, C. (1983) Anal. Biochem. 130, 134–145
21. Kuzmic, P. (1996) Anal. Biochem. 237, 260–273
22. Hoops, S., Sahle, S., Gauges, R., Lee, C., Pahle, J., Simus, N., Singhal, M., Xu, L., Mendes, P., and Kummer, U. (2006) Bioinformatics 22, 3067–3074
23. Johnson, K. A., and Taylor, E. W. (1978) Biochemistry 17, 3432–3442
24. Hanes, J. W., Thal, D. M., and Johnson, K. A. (2006) J. Biol. Chem. 281, 36241–36248
25. Hanes, J. W., and Johnson, K. A. (2007) Nucleic Acids Res. 35, 6973–6983
26. Dolghih, E., Ortiz, W., Kim, S., Krueger, B. P., Krause, J. L., and Roitberg, A. E. (2009) J. Phys. Chem. A 113, 4639–4646
27. Li, Y., Korolev, S., and Waksman, G. (1998) EMBO J. 17, 7514–7525