Published in final edited form as:

Biochemistry. 2018 July 03; 57(26): 3925–3933. doi:10.1021/acs.biochem.8b00417.

Probing DNA base-dependent leaving group kinetic effects on the DNA polymerase transition state

Keriann Oertell¹, Boris A. Kashemirov², Corinne Minard², Pouya Haratipour², Khadijeh Alnajjar³, Joann B. Sweasy³, Vinod K. Batra⁴, William A. Beard⁴, Samuel H. Wilson⁴, Charles E. McKenna^{*,2}, and Myron F. Goodman^{*,1,2}

¹Department of Biological Sciences, University of Southern California, University Park Campus, Los Angeles, CA 90089

²Department of Chemistry, Dana and David Dornsife College of Letters, Arts, and Sciences, University of Southern California, University Park Campus, Los Angeles, CA 90089

³Department of Therapeutic Radiology and Department of Genetics, Yale University School of Medicine, New Haven, CT 06520

⁴Genome Integrity and Structural Biology Laboratory, National Institute of Environmental Health Sciences, National Institutes of Health, Research Triangle, NC 27709

Abstract

We examine the DNA polymerase β (pol β) transition state from a leaving group presteady-state kinetics perspective by measuring the rate of incorporation of dNTPs and corresponding novel β,γ -CXY-dNTP analogues, including individual β,γ -CHF and -CHCl diastereomers with defined stereochemistry at the bridging carbon, during the formation of right (R) and wrong (W) base pairs. Brønsted plots of $\log k_{\text{pol}}$ vs pK_{a4} of the leaving group bisphosphonic acids are used to interrogate the effects of base identity, dNTP analogue leaving group basicity, and the precise configuration of the C-X atom in *R* and *S* stereoisomers on the rate-determining step (k_{pol}). The dNTP analogues provide a range of leaving group basicity and steric properties by virtue of monohalogen, dihalogen or methyl substitution at the carbon atom bridging the β,γ -bisphosphonate which mimics the natural pyrophosphate leaving group in dNTPs. Brønsted plot relationships with negative slopes are revealed by the data, as was found for the dGTP and dTTP analogues, consistent with a bond-breaking component to the TS energy. However, greater multiplicity was shown in the LFER, revealing and unexpected dependence on the nucleotide base for both A and C. Strong base-dependent perturbations that modulate TS relative to ground state energies are likely to arise from electrostatic effects on catalysis in the pol active site. Deviations

*Corresponding Author To whom correspondence should be addressed: C.E.M, mckenna@usc.edu, (213)740-7007; M.F.G., mgoodman@usc.edu, (213)740-5190.

AUTHOR INFORMATION

Author Contributions

K.O. designed, performed and analyzed the kinetic experiments; C.M. and P.H. were involved in the synthesis of compounds used in this study; V.K.B., W.A.B., and S.H.W. provided pol β ; B.A.K., K.A., J.B.S., S.H.W., and C.E.M. provided insight and critiqued the manuscript; K.O. and M.F.G. wrote the manuscript.

Notes

The authors declare no competing financial interests.

from a uniform linear Brønsted plot relationship are discussed in terms of insights gained from structural features of the pre-chemistry DNA polymerase active site.

INTRODUCTION

Biochemical mechanisms regulating DNA polymerase (pol)-catalyzed deoxyribonucleotide incorporation and fidelity have been studied extensively since the early '80s using presteady state kinetic techniques to measure chemical rates (k_{pol}), primer-template (p/t) DNA and dNTP substrate binding parameters, and conformational steps along the reaction pathway^{1–4}. Concurrently, X-ray structural studies on ternary pol-dNTP-p/t DNA complexes have been used to map precise locations of right and wrong dNTP substrates in relation to active site amino acid side chain residues, template bases, and Mg^{2+} cofactors required for pol activity^{5–15}. These methods, supplemented by computational simulations^{16–19}, continue to play a central role in deducing the chemical mechanism and the structure of an otherwise “black box” transition state (TS).

Taking a complementary approach aimed at probing the TS more directly and dynamically, we are investigating pol-catalyzed incorporation and fidelity dependence on leaving-group structure. We synthesized a series of dNTP- β,γ -CXY analogues in which the β,γ -bridging oxygen is replaced by CXY substituents containing halogen or other substituents to modulate the dNTP analogue basicity^{13, 20–22}. The longer P-C bonds of dNTP- β,γ -CXY analogues are compensated by P_{β} -C- P_{γ} angles that are more acute than the P_{β} -O- P_{γ} angle, resulting in similar distances between P_{β} and P_{γ} atoms in natural dNTPs and their β,γ -CXY analogues²³. The corresponding bisphosphonate PP_i analogue leaving groups encompass a wide range of pK_{a4} values spanning 7.8 (CF_2) – 12.3 ($C(CH_3)_2$) (pyrophosphoric acid, the natural leaving group, has $pK_{a4} = 8.9$)^{21, 24}. Fifty-five β,γ -CXY analogues of all four bases and the parent nucleotides have been synthesized. Using human pol β , we have measured k_{pol} as a function of leaving group acidity^{20, 21, 24, 25}. A Linear Free Energy Relationship (LFER), in the form of a Brønsted plot, describes the linear dependence of $\log(k_{pol})$ on leaving group pK_{a4} . Nucleotide incorporation rates are reduced with increasing pK_{a4} values, with each dNTP analogue becoming less reactive to nucleophilic attack at P_{α} as the leaving group becomes increasingly harder to displace, if the rate-determining step (RDS) involves P-O bond-breaking primarily or else concerted with P-O bond-making. When $\log(k_{pol})$ is plotted vs pK_{a4} in such cases, a linear correlation should have a negative slope reflecting, at least partially, increased electrostatic charge stabilization demand in the TS. A horizontal Brønsted line is not consistent (no dependence of $\log(k_{pol})$ on pK_{a4}) with a chemical TS, indicating a pre-chemical RDS^{26–28}. Our previous LFER studies of pol β with dGTP- and dTTP- β,γ -CXY analogues generated nonzero correlations with negative Brønsted slopes, consistent with chemistry as the RDS^{20, 21, 24}, providing direct experimental support for kinetic models for pol β . The LFER approach has further revealed “hidden” TS features, such as profiles showing base-dependent slopes that differ significantly for incorporation of right (R) or wrong (W) deoxynucleotides. Distinct LFER profiles were also observed, one Brønsted line for dihalogen analogs, and a second line for monohalogen and non-halogen analogs.

Here we expand these studies by including kinetic results for novel dATP and dCTP β , γ -CXY analogues, along with separated stereoisomers for G, T, and C. The A and C analogues reveal new and surprising base-specific TS properties: an apparent break in the Brønsted slope, the appearance of separate “mixed” lines containing dihalo, and monohalo derivatives along with a “relocation” of the parent dNTP from monohalo lines, as observed for G and T analogues^{20, 21, 24}, to dihalo and mixed Brønsted lines. We also clarify the importance of CXY stereochemistry on k_{pol} effects by comparing incorporation kinetic rate constants for individual *R*- and *S*-CXY diastereomers, X = Y. Since the leaving group is the same for both isomers, any differences in k_{pol} for individual stereoisomers reflect interactions resulting from the nonequivalent orientations of C-X and C-Y within the pol active site. Taken together, the results for dNTP analogues with all four bases provide a richly detailed portrait illustrating how 1) subtle differences in triphosphate group stereochemistry and relative charge stabilization; 2) base structure; and, most importantly in terms of fidelity mechanisms, 3) base-base pairing vs mispairing affect the overall energy of the TS.

Methods

DNA synthesis, purification, radiolabeling, and annealing

Primer (5'-TAT TAC CGC GCT GAT GCG C), template, (5'-GCG TTG TTC CGA CMG CGC ATC AGC GCG GTA ATA, **M** = A, C, G or T), and 5'-phosphorylated downstream (5'-GTC GGA ACA ACG C) oligomers were synthesized on a solid phase DNA synthesizer and purified by 16% denaturing polyacrylamide gel electrophoresis, followed by desalting using oligonucleotide purification cartridges. Radiolabeling reactions consisted of 1 mol equiv primer which was 5'-end labeled with 0.4 U/ μ L T4 polynucleotide kinase and 0.7 mol equiv [γ -³²P]ATP with the supplied buffer at 37 °C for 30 minutes, followed by heat inactivation at 95 °C for 10 minutes. The primer was purified by size exclusion chromatography using a Bio-Spin 6 column, then annealed by mixing with 1.2 mol equiv template and 1.5 mol equiv downstream oligomers, heated to 95 °C and cooled slowly to room temperature to yield a 1 nt gapped DNA substrate.

Buffer and protein preparation

The reaction buffer consisted of 50 mM Tris-Cl, 20 mM KCl, 20 mM NaCl, 10 mM MgCl₂, 1 mM DTT, and 6% glycerol at pH 8.0. Wild type pol β was purified as previously reported²⁹.

Nucleotide analogue synthesis

All β , γ -bridging methylene dNTP analogues including the individual *R* and *S* β , γ -CHX isomers were synthesized and purified as previously described^{13, 20, 21, 24} or are novel and will be reported elsewhere: dATP analogues: CH₂, (*R/S*)-CHF, -CHCH₃, -CHCl; CF₂, CCl₂, CFCl, CBr₂; dCTP analogues: (*R/S*)-CHF, (*R*)- and (*S*)-CHF, (*R*)- and (*S*)-CHCl, CF₂, CCl₂, CFCl, CBr₂. Parent nucleotides (dGTP, dTTP, dATP, and dCTP) were purchased from New England Biolabs and used without further purification.

Presteady state single turnover reactions—Radiolabeled 1 nt gapped DNA (100 nM) was incubated with 600 nM pol β in reaction buffer (2x mixture) for 3 minutes at 37 °C. Equal volumes of the DNA/pol β mixture and a 2x solution of β,γ -CXY-dNTP in reaction buffer at different concentrations (0.5 – 200 μ M for dRTP and 100 – 2000 μ M for dWTP, 1x concentrations) were rapidly combined using a KinTek model RQF-3 quench flow apparatus. After the appropriate reaction time (0.09 – 20 s for R incorporations, 0.5 s – 45 min for W incorporations), the reaction was quenched with 0.5 M EDTA (pH 8.0). For times longer than 20 s, reactions were initiated and quenched by manual mixing. Reaction products were separated by 20% denaturing polyacrylamide gel electrophoresis (39 cm x 33 cm x 0.4 mm). Dehydrated gels were exposed to a phosphor screen and detected by phosphorescence emission. All reactions were carried out in triplicate. A detailed method for obtaining k_{pol} and apparent K_d kinetic parameters, as well as the generation of Brønsted plots ($\log k_{\text{pol}}$ vs. $\text{p}K_{\text{a4}}$) is given in the Supporting Information.

Results

Previous studies with dGTP and dTTP β,γ -CXY analogues produced similar results in the Brønsted LFER analysis of kinetic data^{20, 21, 24}. Two separate lines linearly correlating $\log k_{\text{pol}}$ with $\text{p}K_{\text{a4}}$ were observed, one grouping the mono-halogen + non-halogen analogues, a second the dihalogens, with similar negative slopes for right incorporations (G•C or T•A) and steeper negative slopes for wrong incorporations (G•T or T•G) (Table S2, References 20, 21, 24). We observed that faster rates of phosphodiester bond formation, for the most part, correlated with the acidity of the leaving bisphosphonate, although differences in the slopes and spacing of the Brønsted plot lines and spacing of the individual data points indicated that other factors also played a role in determining the TS energies and thus $\log(k_{\text{pol}})$. Our new data with incoming analogues of dATP (Fig. 1) and dCTP (Fig. 2) indicate that each base has a dramatic, specific influence on the relative position of the corresponding $\log(k_{\text{pol}})$ datum in the Brønsted plot.

LFER profiles depend on the dNTP analogue base

There are two Brønsted lines observed for the incorporation of A opposite template T, an upper line containing the parent dATP (O), which includes two monohalo and two non-halo dATP β,γ substituents, and a lower line containing dihalo substituents (Fig. 1A). However, in contrast to “well-behaved”, i.e., monophasic, Brønsted lines previously observed for dGTP (Fig. 3A)^{20, 21, 24} and dTTP β,γ -CXY analogues (Fig. 1B)²⁴, there is a break in the LFER (upper biphasic Brønsted line) for the correct pairing of dATP opposite templating base T (Fig. 1A). The slopes of the two lines are similar prior to the break (–0.37 upper, –0.44 lower; slopes given in Supporting Information, Table S2), but above $\text{p}K_{\text{a4}}$ 10.5 there is a sharp increase in slope of the upper line to –1.3. The RDS is chemistry, with Brønsted slopes significantly different from zero over the entire $\text{p}K_{\text{a}}$ range we have investigated^{26–28, 20, 21, 24}. The abrupt apparent change in slope for $\text{p}K_{\text{a4}} > 10.5$ suggests a *base*-dependent change in the chemical mechanism, however the demand for charge stabilization in the TS has suddenly *increased* by an order of magnitude, suggesting the loss of a stabilizing active site interaction that could be mediated indirectly by base pairing via a conformational

change. In the accompanying paper, a tighter fit with O3' is described, which may explain some of this change in terms of a different ground state for this complex.

In contrast to this picture with incoming dATP analogues, we do not see the same effect for the reverse base pairing, i.e., incoming dTTP opposite templating A (Fig. 1B). In this case, the line remains unbroken from pK_{a4} 8.9–11.5, implying the mechanism remains the same for the upper compounds, those containing one or no halogen atoms, throughout this range. Also notable is how close the dihalo and upper lines are, both spatially and in slope for dATP opp. T (Fig. 1A, Supporting Information Table S2), particularly when compared to dTTP opp. A (Fig. 1B). This indicates the incorporation properties for these compounds although somewhat similar, but nevertheless distinct from one another.

Two dihalogen lines for incorporation of A opposite C

For the mispairing of dATP opp. C (Fig. 1C), there is as for correct (Fig. 1A), a separation of the dihalo from the monohalo and non-halo compounds, but here we observe for the first time the presence of *two* distinct dihalo linear correlations, one for CF₂ and CCl₂, the other for CFCl and CBr₂. The lines are essentially parallel (slopes of –1.3 and –1.4, respectively), again demonstrating the similarities of the TS for these compounds. This TS is significantly different from the TS for the upper compounds (slope –0.75), although the RDS remains chemistry throughout the 7.8–12.3 pK_{a4} range. The two dihalo lines are relatively close together, indicating the mechanism is similar, but nonetheless distinct. As opposed to the correct pairing for these compounds (Fig. 1A), we do not see a break for the upper compounds, which suggests that for the A•C mispair, the same TS structure appears to be retained throughout the pK_{a4} range for each grouping of analogues.

dCTP β,γ analogue anomalies

As with all of the other pairings studied so far, the dCTP β,γ -CXY dihalo analogues are separate from the other compounds for both the correct and incorrect pairings. However, in the case of correct incorporation of dCTP opposite G (Fig. 2A), the parent compound is no longer located on the upper line. Note that the *k_{pol}* values for dGTP, dTTP and dATP are essentially coincident with their respective CHF analogues, which is to be expected based on the similarity in the bisphosphonic acid leaving group pK_{a4} values, which differ by only 0.1^{20, 21}. Unexpectedly, however, the parent dCTP has moved to a dihalo Brønsted line, and is now located in between CFCl and CBr₂ (Fig. 2A). The crystal structure of this ternary complex reveals a lack of contact between dCTP and R149 in the active site, as discussed in detail in the accompanying paper, which may be a factor in the change of lines for the parent compound.

This is the first example of a compound lacking two halogen atoms located on a Brønsted line containing dihalogen analogues. This “mixed” line is parallel to a second dihalo line containing CF₂ and CCl₂, with slopes of –0.85 and –0.91 respectively. The monohalo line, which contains CHF, CHCl, and CH₂ analogues, has a slightly shallower slope of –0.76. Since all three lines are of similar slope, the mechanism for each grouping is clearly similar.

The anomalous presence of three lines for C•G (Fig. 2A) is retained for C•A (Fig. 2B), but with one new “wrinkle”: CHF and dCTP are situated together on the middle “mixed”

Brønsted line, whereas CFCl is now located in between CF₂ and CCl₂ on a lower dihalo line (Fig. 2B). The slopes for each line differ significantly, -0.98 for the upper line, -2.3 for the “mixed” line, and -1.4 for the lower exclusively dihalo line, implying a different mechanism is responsible for each grouping despite chemistry being the RDS. A different line could be drawn connecting CF₂, CFCl, and CBr₂ with a shallow slope, however, that would leave CCl₂, O, and CHF as “orphan” points on the LFER. However, we note that for most of the data, especially for G (Figs. 3A, 3B; References ^{20, 21, 24}), and T (Figs. 1B, 3C; Reference ²⁴), the Brønsted lines appear unambiguous.

LFER alignments of separated dNTP β,γ stereoisomers

The synthesis of dNTP β,γ -bridging CXY analogues with X = Y results in 50:50 mixtures of *R* and *S* isomers at the bridging chiral center ^{13, 22, 25}. X-ray crystallographic studies of pol β -DNA- β,γ -dNTP analogue ternary complexes, and kinetic studies, show that the enzyme is able to discriminate between *R* and *S* stereoisomers in ground-state active site occupancies^{13, 22, 25} and between *R* and *S* isomers via differences in apparent K_d (prechemistry) and in k_{pol} (chemistry) in the TS ²⁵. Based on these structural and kinetic differences, we included data for selected plots.

The individual diastereomers are plotted on the same x-axis coordinate (pK_{a4}) as the mixture since the leaving group is the identical (achiral) compound. For the majority of the individual diastereomers, the *R* stereoisomer, which has the halogen atom (or F in the case of the CFCl analogue) pointed towards R183 in the active site, turns over faster than both the mixture and the *S* stereoisomer and therefore appears at the top of the three points (Figs. 1B, 2, 3), with the rate of the mixture being closer to the *R* isomer than to the *S*. A few differences are observed in this general pattern, especially when the mixture and *R* isomer are close together and may then be reversed, as seen for dT-CHF and dT-CHCl opposite A (Fig. 1B). Typically, the differences between the diastereomers are much larger for W (Fig. 2B, 3B, 3C) than for R base pairs (Figs. 1B, 2A, 3A).

In keeping with their anomalous behavior, the most pronounced variations in individual isomer patterns come from dCTP analogues (Fig. 2). For the pairing of C opp. G (Fig. 2A), the CHF compound follows the more usual pattern, the *R* isomer highest (fastest), followed closely by the mixture, and the *S* isomer lowest (slowest). The magnitude of k_{pol} for the mixture of CHCl is significantly closer to *S*, where the *R* and mix k_{pol} s tend to be much more similar for the non-C bases (Figs. 1B, 3), with one exception. The mispairing of the dC-CHCl compounds opposite A (Fig. 2B) conform to the behavioral patterns observed for dG and dT, where $k_{pol} R > mix > S$. The k_{pol} of the CHF mix is significantly below that for the *R* isomer, located closer to *S*.

Discussion

While the nucleotidyl transfer reaction responsible for template-dependent DNA synthesis is essentially the same for all pols, kinetic and structural studies have identified profound differences in individual steps in the pol catalytic cycle, from the binding of cognate ³⁰ and non-cognate ¹² substrates, to pol conformational changes and phosphodiester bond chemistry ^{1, 2, 4}. Each step may serve as a kinetic fidelity checkpoint that discriminates

between R and W in a pol-specific manner³. For example, replicative pols δ and ϵ exhibit distinct mutational signatures³¹, and repair pols ι and η have been shown to increase the incorporation of noncognate in specific template motifs^{32–34}. Key to understanding the none-too-subtle nuances of pol fidelity is to deduce how R and W base-pairing interactions modulate TS relative to ground-state energies. Approaches to infer the influence on base-pairing on TS energies include presteady state kinetic measurements^{20, 21, 24}, computer simulations^{16–18} and structural analysis^{5, 9, 10, 12, 14, 35} including recent visualization of TS active site geometries using time-lapse X-ray crystallography^{11, 36}.

We have taken a complementary approach to probe the TS, from the vantage point of PP_i and bisphosphonate analogue leaving-groups. The incorporation rate constants (k_{pol}) have been plotted against leaving group acidities, in which the chemically unmodifiable (except isotopically) dNTP β, γ -bridging O is replaced with CXY moieties with $\text{p}K_{\text{a4}}$ values of 7.8 to 12.3. We have synthesized a toolbox of 55 dNTP β, γ analogues distributed among the four natural dNTP substrates, including several separate *R* and *S* diastereomers. Each analogue undergoes a “standard” α, β -P-O-P cleavage reaction, in the sense that each atom involved in both bond making and breaking at the nucleotide P α remains the same as the natural dNTP. The chemical reaction mechanism is in principal conserved, although the chemical step as embodied in k_{pol} is sensitive to how each X, Y substituent may stabilize the triphosphate group negative charge. The leaving group (PP_i or CXY-bisphosphonate) acquires an additional negative charge in the course of the insertion reaction, accompanied by a linearly decreasing value of $\log(k_{\text{pol}})$ plotted as a function of increasing $\text{p}K_{\text{a4}}$ (Figs. 1 – 3).

TS properties revealed by LFER leaving group analysis

Polymerase structural studies often rely on substrate analogs that stop the reaction prior to catalysis resulting in noncatalytic “dead” complexes. To overcome this problem Freudenthal et al³⁶ and Nakamura et al³⁷ have recently trapped catalytic intermediates prior to and following nucleotide insertion of natural substrates by freezing polymerase (pol β or pol η)/DNA crystals at specific time intervals during catalysis. It was found that pyrophosphate and its associated metal remain stably bound in the active site and during the course of the reaction the active-site aspartate residues and metal ions remain in constant configuration. These findings are very important for analysis of our data. We have crystal structures of ternary complexes of pol β , DNA and practically every dNTP- β, γ -CXY analogue used in this study. From crystallographic studies we can conclude (see accompanying paper) that ternary complex structures with incoming dGTP, dATP, dTTP and dCTP analogs are not grossly altered by CXY modifications. This information, combined with the Freudenthal and Nakamura studies which show little to no movement of the P β -O-P γ portion of the incoming dNTP during the reaction, implies that the transition state conformation of the bisphosphonate leaving group as it is formed correlates with that of the ground state structure.

Since the conformation does not change dramatically during the reaction, the decrease in k_{pol} with a reduction in negative charge stabilization is consistent with an increase in the TS free energy in a chemical rate-determining step for pol β catalysis (for example, Fig. 4A,

step c, black, blue, red lines, References ^{20, 21, 24}). If the reason for the separation between the dihalo compounds and the rest was a result of the dihalo compounds having a substitution on both sides of the carbon, we would expect the $C(CH_3)_2$ compound to fall on the dihalo line. Since this is not the case, the linear fit of the Brønsted plot for both R and W (Figs. 1 – 3) ^{20, 21, 24} indicates that for the TS, the region of the active site proximal to the X and Y substituents is “accommodating”, i.e., showing little or no steric hindrance. This is similar to what is seen in the ground state where largely only subtle differences are seen in the ternary complex crystal structures between different substitutions (see accompanying paper). Notably, multiple distinct Brønsted profiles are observed for dihalo, monohalo and non-halo substituents (Figs. 1 – 3). In discussing LFER correlations observed for the dGTP β, γ -CXY analogues ^{20, 21, 24}, we speculated that repulsion between the *S* halogen and the α, β -bridging oxygen in the nucleotide analogue, or an unfavorable electrostatic/steric interaction between Arg183, could be more pronounced with base mispair in a chemical transition state. From our recent data, it appears that a large reduction in the magnitude of k_{pol} for the dihalo substitution is unlikely to result primarily from steric hindrance in the active site, because the two bulky moieties, $CHCH_3$ and $C(CH_3)_2$, both correlate with the upper monohalo-nonhalo lines for R and W base pairs (Figs. 1 – 3).

The LFER data can be represented on a schematic of a 2D free energy profile by increasing the heights of the activation energy for the chemical step for pol β with increasing pK_{a4} of the bisphosphonate leaving groups (Fig. 4A, step c). In contrast, dNTP- β, γ analogue K_d values are roughly similar in a 5 – 30 μM range for R and 400 – 1000 μM range for W (Supporting Information, Table S1; Fig. 4A, steps a and b (Refs ^{1, 2, 4}). Although prechemical conformational states cannot be identified from the Brønsted plots (Fig. 4A, step b), it is speculated that a type of “conformational-chemical” coupling ^{1, 24} involving each base differently causes significantly altered TS ΔG values.

Commonalities in this complex data set begin to emerge when combining the Brønsted plots into three subgroups (Figs. 4B, C). The upper subgroup composed of the monohalo and nonhalo derivatives is closely clustered with the stronger base stacking substrates (G and A) having higher k_{pol} values compared to the weaker stacking (T and C); this is the case over most of the pK_{a4} range for R (Fig. 4B) and W (Fig. 4C). The middle and lower subgroups contain the dihalos with the G and A situated above T and C for R (Fig. 4B). For W, G and A are also located above T and C, except that there are also two dihalos for A (CF₂ and CCl₂) that are also present in the lower subgroup (Fig. 4C). The parent dCTP is located “anomalously” on the upper dihalo line for R (Fig. 4B), likely due to the loss of a key contact in the ground state (see accompanying paper), and the mixed line for W (Fig. 4C). The LFER analysis reveals three distinct effects on RDS TS energies: 1) a large effect of increasing leaving-group pK_{a4} correlating with a decrease in k_{pol} ; 2) a reduction in k_{pol} for dihalo derivatives compared with mono- and non-halo derivatives with similar electronegativities; 3) TS energies that differ for G, T, A, and C. Each seems likely to reflect different electrostatic perturbations on TS ΔG values.

Mechanistic strategies in enzyme catalysis based at least in part on modulation of active site electrostatics have been extensively investigated using sophisticated computer modeling over several decades ^{38, 39}. Local electric fields in the active site of ketosteroid isomerase were

recently estimated by Boxer and coworkers to contribute $\sim 2/3$ of the total enzymatic rate enhancement, with positioning of a key active site residue implicated in catalysis (Asp40)^{40, 41}. We suggest that analogous electrostatic effects within the active site of pol β may at least partly account for the LFER profiles we have observed, driving the large displacements of the dihalo Brønsted lines relative to the monohalo-nonhalo lines for R and W, and triggered by base specific displacements resulting from differences in H-bonding and stacking interactions with base pairing or mispairing. Notably, the slopes of the lines are only slightly changed for R and W, suggesting that the reaction mechanism in terms of response of catalysis to increased demand for charge stabilization in the TS has not been altered. Overall, the complex structure-function correlations that emerge from our data pose an interesting and serious challenge to TS modeling to reproduce the combined leaving group + base effects on the TS energy (~ 1 to 2 kcal/mol) and may be amenable to computational analysis^{16–18} to generate an exquisitely detailed dynamic reaction pathway model.

TS energy perturbations were observed for individual diastereomers of β , γ -CHF, β , γ -CHCl, and β , γ -CFCl of dGTP and dTTP, compounds that are structurally identical except for the relative orientations of the C-F or C-Cl vs C-H bonds²⁵. These perturbations are confirmed in the dATP and dCTP analogues, showing that they are not base-dependent, but rather driven by the CXY stereochemistry. The *S*-CXY isomer is typically found to have a slightly smaller $\log(k_{\text{pol}})$, and the *S*- and *R*-CXY points bracket the data point obtained for the synthetic mixture ($\sim 1:1$) (Figs. 1B, 2, 3). We previously demonstrated by ¹⁹F NMR in competitive kinetics experiments using the of β , γ -CHF-dGTP *R/S* mixture that pol β more rapidly turns over the *R*- CHF isomer, which is also bound preferentially into crystals analyzed by X-ray crystallography and orients its halogen atom at 3 Å from Arg187^{13, 25}. The same kinetic preference held true for β , γ -CHCl-dGTP when a similar competition experiment was performed and analyzed by ³¹P NMR, although the ternary crystal structure does not show any ground state binding preference²⁵. In an accompanying paper, we compare ternary complex X-ray structures of pol β -p/t DNA bound to a selected set of dNTP β/γ analogues to determine how the bound nucleotide affects the ground-state protein structure (Ref). In conclusion, we have shown that LFER analysis of dNTP leaving group effects in pol β catalysis reveals the sensitivity of R and W TS energies to X,Y substituents, including stereospecific properties, and to base identity. We suggest that the β , γ - bridging oxygen dNTP analogues now representing all four natural bases and including key individual CXY diastereomers, provides a subtle and powerful tool to interrogate the TS of the incorporation reaction.

Supporting Information

Detailed descriptions of the kinetic assays and the generation of the Brønsted plots, along with the kinetic data and slopes of the Brønsted lines are provided in the Supporting Information. This content is available free of charge at <http://pubs.acs.org>.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding Sources

This research was supported by National Institute of Health grant number 1U19CA177547 and by the Division of Intramural Research of the NIH, National Institute of Environmental Health Sciences project numbers Z01 ES050158 and ES050159.

ABBREVIATIONS

Pol	DNA polymerase
p/t	primer-template
TS	transition state
LFER	linear free energy relationship
RDS	rate-determining step
R	right
W	wrong

REFERENCES

- [1]. Johnson KA. Conformational coupling in DNA polymerase fidelity. *Annu Rev Biochem.* 1993; 62:685–713. [PubMed: 7688945]
- [2]. Johnson KA. Role of induced fit in enzyme specificity: a molecular forward/reverse switch. *J Biol Chem.* 2008; 283:26297–26301. [PubMed: 18544537]
- [3]. Joyce CM, Benkovic SJ. DNA polymerase fidelity: kinetics, structure, and checkpoints. *Biochemistry.* 2004; 43:14317–14324. [PubMed: 15533035]
- [4]. Tsai YC, Johnson KA. A new paradigm for DNA polymerase specificity. *Biochemistry.* 2006; 45:9675–9687. [PubMed: 16893169]
- [5]. Arndt JW, Gong WM, Zhong XJ, Showalter AK, Liu J, Dunlap CA, Lin Z, Paxson C, Tsai MD, Chan MK. Insight into the catalytic mechanism of DNA polymerase beta: Structures of intermediate complexes. *Biochemistry.* 2001; 40:5368–5375. [PubMed: 11330999]
- [6]. Batra VK, Beard WA, Shock DD, Pedersen LC, Wilson SH. Structures of DNA polymerase beta with active-site mismatches suggest a transient abasic site intermediate during misincorporation. *Mol Cell.* 2008; 30:315–324. [PubMed: 18471977]
- [7]. Batra VK, Perera L, Lin P, Shock DD, Beard WA, Pedersen LC, Pedersen LG, Wilson SH. Amino acid substitution in the active site of DNA polymerase beta explains the energy barrier of the nucleotidyl transfer reaction. *J Am Chem Soc.* 2013; 135:8078–8088. [PubMed: 23647366]
- [8]. Batra VK, Shock DD, Beard WA, McKenna CE, Wilson SH. Binary complex crystal structure of DNA polymerase beta reveals multiple conformations of the templating 8-oxoguanine lesion. *Proc Natl Acad Sci U S A.* 2012; 109:113–118. [PubMed: 22178760]
- [9]. Beard WA, Wilson SH. Structure and mechanism of DNA polymerase beta. *Chem. Rev.* 2006; 106:361–382. [PubMed: 16464010]
- [10]. Biertumpfel C, Zhao Y, Kondo Y, Ramon-Maiques S, Gregory M, Lee JY, Masutani C, Lehmann AR, Hanaoka F, Yang W. Structure and mechanism of human DNA polymerase eta. *Nature.* 2010; 465:1044–1048. [PubMed: 20577208]
- [11]. Gao Y, Yang W. Capture of a third Mg(2)(+) is essential for catalyzing DNA synthesis. *Science.* 2016; 352:1334–1337. [PubMed: 27284197]
- [12]. Johnson SJ, Beese LS. Structures of mismatch replication errors observed in a DNA polymerase. *Cell.* 2004; 116:803–816. [PubMed: 15035983]

- [13]. McKenna CE, Kashemirov BA, Upton TG, Batra VK, Goodman MF, Pedersen LC, Beard WA, Wilson SH. (*R*)-beta,gamma-Fluoromethylene- dGTP-DNA Ternary Complex with DNA Polymerase beta. *J Am Chem Soc.* 2007; 129:15412–15413. [PubMed: 18031037]
- [14]. Wang W, Hellinga HW, Beese LS. Structural evidence for the rare tautomer hypothesis of spontaneous mutagenesis. *Proc Natl Acad Sci U S A.* 2011; 108:17644–17648. [PubMed: 22006298]
- [15]. Yang W, Weng PJ, Gao Y. A new paradigm of DNA synthesis: threemetal-ion catalysis. *Cell Biosci.* 2016; 6:51. [PubMed: 27602203]
- [16]. Florian J, Goodman MF, Warshel A. Computer simulation of the chemical catalysis of DNA polymerases: discriminating between alternative nucleotide insertion mechanisms for T7 DNA polymerase. *J Am Chem Soc.* 2003; 125:8163–8177. [PubMed: 12837086]
- [17]. Florian J, Goodman MF, Warshel A. Computer simulation studies of the fidelity of DNA polymerases. *Biopolymers.* 2003; 68:286–299. [PubMed: 12601790]
- [18]. Florian J, Goodman MF, Warshel A. Computer simulations of protein functions: searching for the molecular origin of the replication fidelity of DNA polymerases. *Proc Natl Acad Sci U S A.* 2005; 102:6819–6824. [PubMed: 15863620]
- [19]. Kamerlin SC, McKenna CE, Goodman MF, Warshel A. A computational study of the hydrolysis of dGTP analogues with halomethylene-modified leaving groups in solution: implications for the mechanism of DNA polymerases. *Biochemistry.* 2009; 48:5963–5971. [PubMed: 19391628]
- [20]. Sucato CA, Upton TG, Kashemirov BA, Batra VK, Martinek V, Xiang Y, Beard WA, Pedersen LC, Wilson SH, McKenna CE, Florian J, Warshel A, Goodman MF. Modifying the beta,gamma leaving-group bridging oxygen alters nucleotide incorporation efficiency, fidelity, and the catalytic mechanism of DNA polymerase beta. *Biochemistry.* 2007; 46:461–471. [PubMed: 17209556]
- [21]. Sucato CA, Upton TG, Kashemirov BA, Osuna J, Oertell K, Beard WA, Wilson SH, Florian J, Warshel A, McKenna CE, Goodman MF. DNA polymerase beta fidelity: Halomethylene-modified leaving groups in pre-steady-state kinetic analysis reveal differences at the chemical transition state. *Biochemistry.* 2008; 47:870–879. [PubMed: 18161950]
- [22]. Wu Y, Zakharova VM, Kashemirov BA, Goodman MF, Batra VK, Wilson SH, McKenna CE. beta,gamma-CHF- and beta,gamma-CHCl-dGTP diastereomers: synthesis, discrete 31P NMR signatures, and absolute configurations of new stereochemical probes for DNA polymerases. *J Am Chem Soc.* 2012; 134:8734–8737. [PubMed: 22397499]
- [23]. Zhang Z, Eloge J, Florian J. Quantum mechanical analysis of nonenzymatic nucleotidyl transfer reactions: kinetic and thermodynamic effects of beta-gamma bridging groups of dNTP substrates. *Biochemistry.* 2014; 53:4180–4191. [PubMed: 24901652]
- [24]. Oertell K, Chamberlain BT, Wu Y, Ferri E, Kashemirov BA, Beard WA, Wilson SH, McKenna CE, Goodman MF. Transition state in DNA polymerase beta catalysis: rate-limiting chemistry altered by base-pair configuration. *Biochemistry.* 2014; 53:1842–1848. [PubMed: 24580380]
- [25]. Oertell K, Wu Y, Zakharova VM, Kashemirov BA, Shock DD, Beard WA, Wilson SH, McKenna CE, Goodman MF. Effect of beta,gamma-CHF- and beta,gamma-CHCl-dGTP halogen atom stereochemistry on the transition state of DNA polymerase beta. *Biochemistry.* 2012; 51:8491–8501. [PubMed: 23043620]
- [26]. Hollfelder F, Herschlag D. The nature of the transition state for enzyme catalyzed phosphoryl transfer. Hydrolysis of O-aryl phosphorothioates by alkaline phosphatase. *Biochemistry.* 1995; 34:12255–12264. [PubMed: 7547968]
- [27]. Kirsch, JF. (1972) Linear Free Energy Relationships in Enzymology, In *Advances in Linear Free Energy Relationships* (Chapman, NB.; , and Shorter, J., Eds.), pp 369–400, Springer US, Boston, MA.
- [28]. Lassila JK, Zalatan JG, Herschlag D. Biological phosphoryl-transfer reactions: understanding mechanism and catalysis. *Annu Rev Biochem.* 2011; 80:669–702. [PubMed: 21513457]
- [29]. Beard WA, Wilson SH. Purification and domain-mapping of mammalian DNA polymerase beta. *Methods Enzymol.* 1995; 262:98–107. [PubMed: 8594388]
- [30]. Watson JD, Crick FH. The structure of DNA. *Cold Spring Harb Symp Quant Biol.* 1953; 18:123–131. [PubMed: 13168976]

- [31]. Rayner E, van Gool IC, Palles C, Kearsley SE, Bosse T, Tomlinson I, Church DN. A panoply of errors: polymerase proofreading domain mutations in cancer. *Nat Rev Cancer*. 2016; 16:71–81. [PubMed: 26822575]
- [32]. Choi JY, Lim S, Eoff RL, Guengerich FP. Kinetic analysis of basepairing preference for nucleotide incorporation opposite template pyrimidines by human DNA polymerase iota. *J Mol Biol*. 2009; 389:264–274. [PubMed: 19376129]
- [33]. Zeng X, Negrete GA, Kasmer C, Yang WW, Gearhart PJ. Absence of DNA polymerase eta reveals targeting of C mutations on the nontranscribed strand in immunoglobulin switch regions. *J Exp Med*. 2004; 199:917–924. [PubMed: 15051760]
- [34]. Zeng X, Winter DB, Kasmer C, Kraemer KH, Lehmann AR, Gearhart PJ. DNA polymerase eta is an A-T mutator in somatic hypermutation of immunoglobulin variable genes. *Nat Immunol*. 2001; 2:537–541. [PubMed: 11376341]
- [35]. Bebenek K, Pedersen LC, Kunkel TA. Replication infidelity via a mismatch with Watson-Crick geometry. *Proc Natl Acad Sci U S A*. 2011; 108:1862–1867. [PubMed: 21233421]
- [36]. Freudenthal BD, Beard WA, Shock DD, Wilson SH. Observing a DNA polymerase choose right from wrong. *Cell*. 2013; 154:157–168. [PubMed: 23827680]
- [37]. Nakamura T, Zhao Y, Yamagata Y, Hua YJ, Yang W. Watching DNA polymerase eta make a phosphodiester bond. *Nature*. 2012; 487:196–201. [PubMed: 22785315]
- [38]. Warshel A. Energetics of enzyme catalysis. *Proc Natl Acad Sci U S A*. 1978; 75:5250–5254. [PubMed: 281676]
- [39]. Warshel A, Sharma PK, Kato M, Xiang Y, Liu H, Olsson MH. Electrostatic basis for enzyme catalysis. *Chem Rev*. 2006; 106:3210–3235. [PubMed: 16895325]
- [40]. Fried SD, Bagchi S, Boxer SG. Extreme electric fields power catalysis in the active site of ketosteroid isomerase. *Science*. 2014; 346:1510–1514. [PubMed: 25525245]
- [41]. Fried SD, Boxer SG. Electric Fields and Enzyme Catalysis. *Annu Rev Biochem*. 2017; 86:387–415. [PubMed: 28375745]

Figure 1.

Plots correlating leaving group pK_{a4} and incorporation k_{pol} for A opposite T (A), T opposite A (B), and A opposite C (C). In the case of both correct pairings, the compounds are divided into two lines, one containing those compounds with two halogen atoms and the other with compounds containing one or no halogen atoms. In the case of dA-CHCH₃ opp. T, however, the rate is dramatically decreased from what is expected for A opp. T, causing a sharp increase in the slope between dA-CH₂ and dA-CHCH₃. The lines for the A opp. T plot are much closer together and the slopes are more shallow than for the T opp. A pairing. The separate diastereomers for T opp. A behave as expected, with the rate of the *R* isomer very close to the rate of the mixture and the *S* isomer significantly less than the mixture. In contrast to previous data with G and T, for the first time we see two dihalogen lines for A opp. C, with a difference between CF_2/CCl_2 and $CFCl/CBr_2$.

Figure 2.

Brønsted correlations for the dCTP β,γ -bridging analogues for both correct opposite G (**A**) and mispairing opposite A (**B**). In both cases, there are three lines with the parent and CBr_2 compounds on a separate, unexpected line. For the correct pairing, the CFCl analogue also falls on this middle line instead of the dihalo line, but falls on the expected dihalo line for the mispair while the CHF compound is on the intermediate line instead.

Figure 3.

Brønsted plots for dGTP analogues for both correct (A) and mispairing (B) as well as the mispair of dTTP opp. G (C). The diastereomer data is shown on the plots for the first time. The general pattern for rates of the individual diastereomers of CHF, CHCl, and CFCl is $R > \text{mixture} > S$. In most cases, the mixture is closer to the R than to the S , reflecting the preference of pol β for the R isomer. The rates of dT-CBr₂ and dT-CHCH₃ could not be determined for dTTP opp. G.

Figure 4.

Sketch of a simple reaction coordinate diagram (A) and overlay Brønsted plots for all eight pairings studied, excluding separate diastereomers (B, C). (A) Representative reaction coordinate diagram for the generic incorporation of a dNTP by a polymerase. The reaction is shown at the top with step a corresponding to the binding of the dNTP to the enzyme-DNA duplex, step b, a conformational change in the protein, and step c the incorporation reaction (chemistry). Three different activation energies (E_{a1} , E_{a2} , E_{a3}) are shown, representing compounds of increasing leaving group pK_{a4} , which correlates with decreased k_{pol} for pol β . (B) Overlay LFER plots for correct pairings. Black lines represent the monohalo and nonhalo containing lines from each of the four correct pairings. Red lines correspond to the mixed (C opp. G) and two of the dihalo lines (G opp. C and A opp. T). The green lines correspond to the dihalo lines of T opp. A and C opp. G. Solid lines represent the G opp. C data; dashed lines, T opp. A; dashed lines with one dot, A opp. T; dashed lines with two dots, C opp. G. (C) Overlay LFER plots for mispairings. As for the correct data, black lines represent the monohalo and nonhalo containing lines from each of the four incorrect pairs. Red lines correspond to the mixed (C opp. A) and two of the dihalo lines (G opp. T and the upper dihalo line for A opp. C). The green lines correspond to the dihalo lines of T opp. G, C opp. A, as well as the lower dihalo line for A opp. C. Solid lines represent the G opp. T data; dashed lines, T opp. G; dashed lines with one dot, A opp. C; dashed lines with two dots, C opp. A.