HHS Public Access

Author manuscript

Biochemistry. Author manuscript; available in PMC 2019 July 03.

Published in final edited form as:

Biochemistry. 2018 July 03; 57(26): 3934–3944. doi:10.1021/acs.biochem.8b00418.

Mapping functional substrate-enzyme interactions in the pol β active site through chemical biology: Structural responses to acidity modification of incoming dNTPs

Vinod K. Batra[1], **Keriann Oertell**[2], **William A. Beard**[1], **Boris A. Kashemirov**[3], **Charles E. McKenna**[3], **Myron F. Goodman**[2], and **Samuel H. Wilson**[1,4]

[1]Genome Integrity and Structural Biology Laboratory, National Institute of Environmental Health Sciences, 111 T.W. Alexander Drive, Research Triangle Park, NC, 27709.

[2]Department of Biological Sciences, University of Southern California, Los Angeles, CA.

[3]Department of Chemistry, University of Southern California, Los Angeles, CA.

Abstract

We report high-resolution crystal structures of DNA polymerase (pol) β in ternary complex with a panel of incoming dNTPs carrying acidity-modified 5′-triphosphate groups. These novel dNTP analogues have a variety of halomethylene substitutions replacing the bridging oxygen between Pβ and Pγ of the incoming dNTP, whereas other analogues have alkaline substitutions at the bridging oxygen. Use of these analogues allows the first systematic comparison of effects of 5′-triphosphate acidity modification on active site structures and the rate constant of DNA synthesis. These ternary complex structures with incoming dATP, dTTP and dCTP analogues reveal the enzyme's active site is not grossly altered by the acidity modifications of the triphosphate group. Yet, with analogues of all three incoming dNTP bases, subtle structural differences are apparent in interactions around the nascent base pair and at the guanidinium groups of active site arginine residues. These results are important in understanding how acidity modification of the incoming dNTP's 5′-triphosphate can influence DNA polymerase activity and the significance of interactions at arginines 183 and 149 in the active site.

Graphical Abstract

[4]Corresponding Author: wilson5@niehs.nih.gov.

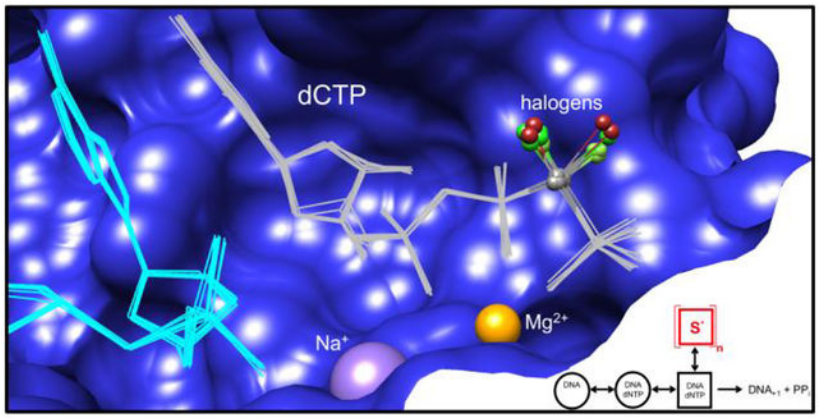
Author Contributions VKB performed crystallographic studies. The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript. ‡

Supporting Information.

Supporting information file contains crystallographic statistics (Table S1, S4 and S6), and distance tables for dNTP analogues (Table S3, S5, and S7) and the significance of distances (Table S2) and figures S1–S8

Accession Numbers

The coordinates and structure factors for structures reported in this manuscript have been deposited with Protein Data Bank.



Introduction

DNA polymerase mediated DNA synthesis is essential in all forms of life for genome replication and DNA repair. DNA polymerases incorporate nucleoside 5'-triphosphates into the 3'-end of a nascent DNA strand releasing pyrophosphate, and the enzymes are able to vastly lower the activation energy of nucleotidyl transfer compared to the reaction in water[1]. Many different isoforms of DNA polymerase are found in nature, and the enzyme isoforms are proposed to have specialized capacities depending upon their biological functions[2]. Some aspects of polymerase function, however, are shared across all enzyme isoforms, including primer strand elongation with 5' to 3' polarity, use of divalent metal ion cofactors, strong preference for nucleoside 5′-triphosphate substrates over other phosphate forms, capacity to release and cycle away from the DNA product, and the ability to conduct both the forward and reverse enzymatic reactions[1]. The requirements for template-directed primer strand elongation by DNA polymerases include precise active site alignment of substrates enabling the chemical reaction where the primer O3' oxyanion attacks Pα of the incoming nucleotide[3]. The reaction releases pyrophosphate (PPi) from the incoming nucleotide and results in covalent bond formation between the incoming nucleotide and primer strand. Proper alignment of the O3' and Pα atoms depends on the enzyme's ability to position the substrates in a precise orientation for chemistry that is promoted through a divalent metal ion mediated charge transfer[4].

Crystal structures of DNA polymerase (pol) β revealed the architecture of this enzyme's active site with bound pre-chemistry substrate molecules, divalent metal ions and coordinating water molecules[567]. The active site interactions include the enzyme's main-chain, side chains and water molecules that coordinate oxygens of the incoming nucleotide's 5'-triphosphate group. The interactions among these components are interdependent such that chemistry depends upon hydrogen bonding between the template base and incoming nucleotide base, as well as interactions between the enzyme and the triphosphate group of the incoming nucleotide, among others. A key component in this network of interactions is the integrity of a charge neutralization system between two closely positioned divalent metal ions, water molecules, enzyme side chains, and the triphosphate group of the incoming nucleotide[8].

Information from crystal structures[5] of the pol β active site and attendant computational analyses[9] has been instrumental in assigning the interatomic angle, optimal distance, and charge transfer for the O3' attack on Pα, for both correct and incorrect incoming nucleotides. From this work, O3' is positioned to make an in-line attack on Pα over a distance of ~ 3.4 Å and the angle of the O3' – Pα – O alignment is 160°. It is also known that mismatched incoming nucleotides are associated with distortions in both the angle for forming the O3' and Pα bond and the distance between O3' and Pα[7, 10]. These distortions correlate with lower efficiency of the forward nucleotidyl transferase reaction.

Despite extensive study and the advances noted above, the atomic level mechanisms of controlling the rate and accuracy the DNA polymerase enzymatic reaction have only recently begun to emerge. The combined use of enzyme kinetics along with structural and chemical biology approaches provides important new insight and reveals topics for future study. For example, McKenna and coworkers[1112–16] recently synthesized a novel panel of dNTP analogues with 5′-triphosphate group substitutions at the oxygen bridging the β-phosphate and γ-phosphate groups. These analogues enable study of a range of charge donating and charge withdrawing properties in the triphosphate group of the incoming nucleotide, and the analogues provide a systematic approach to understand the influence of triphosphate group acidity on the polymerase reaction. To enhance the potential of these modified nucleotides for use in DNA polymerase enzymology, we have obtained crystal structures of pre-catalytic ternary complexes of the analogues bound in the pol β active site with a single-nucleotide gapped DNA substrate.

These ternary complex pre-catalytic structures reveal spatial differences between the enzyme side chains, coordinating water molecules, and the oxygen atoms of the modified triphosphate group. In a subset of cases, correlations are evident between acidity in the triphosphate group and differences in active site contacts. These results are discussed and enable analyses of catalytic activity as a function of acidity of the incoming nucleotide's triphosphate group with both purine and pyrimidine bases[14].

Experimental Procedures

Protein expression and substrates.

Human pol β was overexpressed in *Escherichia coli* and purified as previously described[17]. DNA consisted of a 16-mer template, a 9-mer upstream primer and a 5-mer downstream oligonucleotide. The DNA sequences were similar (except at the template position) to that used in previous studies[5] in order to minimize sequence-dependent structural differences. The sequence of the downstream oligonucleotide was 5′-GTCGG-3′ and the 5′-terminus was phosphorylated. The template sequence was 5′-CCGACXXCGCATCAGC-3′ and the upstream primer sequence was 5′-GCTGATGCG-3′. Proper annealing of these oligonucleotides resulted in a 2-nucleotide gapped DNA substrate with guanosine or thymidine serving as the first (X) templating nucleotide in the gap. Oligonucleotides were dissolved in 20 mM $MgCl_2$, 100 mM Tris-HCl, pH 7.5. Template, upstream, and downstream oligonucleotides were mixed in a 1:1:1 ratio and annealed using a thermal cycler by heating 10 min at 90 °C and cooling to 4 °C (1 °C/min) resulting in a mixture of 1 mM of 2-nucleotide gapped duplex DNA. This solution was then mixed with an equal

volume of pol β at 4 °C. The mixture was warmed to 35 °C and cooled gradually to 4 °C. The resulting 2-nt gapped oligonucleotide substrate was incubated with dideoxyCTP or dideoxyATP to create 1-nt gapped substrates with dideoxy terminated primer, and the templating nucleotide in the gap (X) was A, G or T.

Crystallization and data collection.

The pol β/DNA mixture with 1-nt gap as described above was crystallized by sitting-drop vapor diffusion. The drops were incubated at 18 °C and streak seeded after 1 day; crystals grew in approximately 2 to 4 days after seeding. The ternary complexes were obtained by soaking the binary complex crystals in cryo-protectant solution containing artificial mother liquor with 90 mM sodium acetate, 100 mM $MgCl_2$, and 2, 4, or 10 mM of dNTP analogues with 12% ethylene glycol, and then flash-frozen at 100 K in a nitrogen gas stream. Data were collected at 100 K on a CCD detector system mounted on a Rigaku MicroMax-007HF rotating anode generator. Data were integrated and reduced with HKL2000 software[18]. Structures were determined by molecular replacement with previously determined structures of pol β complexed with incoming dNTP analogues[5, 10–13, 19]. The crystal structures have similar lattices and are sufficiently isomorphous to determine the molecular replacement model using PHENIX[20]. Further refinement and model buildings were carried out using COOT. The figures were prepared in Chimera[21].

Chemical syntheses of the G analogues have been published[5, 10–13, 19], and chemical syntheses of the A, T, and C analogues are being reported separately.

Results

Overview of the crystal structures:

Using the panel of acidity modified dNTP analogues described by McKenna and associates, we obtained ternary complex crystal structures of human pol β. Each structure is compared with a known ternary complex reference structure (PDB code 2FMS) and also with new structures of normal incoming dATP and dCTP. The comparisons include many new pol β structures with incoming dATP, dTTP and dCTP analogues, respectively. A DNA duplex with a single-nucleotide gap is used in all structures, and a base pair complementary template base is in the gap. The template-primer terminus is the G:C base pair for incoming A and T analogues and T:A base pair for incoming C analogues, and the dCMP/dAMP primer terminus sugar is dideoxyribose, used to block the insertion reaction during crystal growth. The divalent metal ion in the nucleotide metal site is magnesium, and the catalytic metal site is occupied by sodium, as expected[5]. Statistics for the structures with these novel incoming dATP, dTTP and dCTP analogues are shown in Tables S1, S4 and S6, respectively. In two cases, dATP and dCTP, reference structures with normal incoming dNTPs also were obtained and these were found to be similar to the reference structure, 2FMS; the comparisons to be described are with superimposition of enzyme residues 10–335 or residues 155–260 (for the active site), as specified.

In all of these ternary complex structures, the enzyme is in the “closed conformation.” Nevertheless, in the structures of two of the dATP analogues (A-CF_2 and A-CCl_2), altered

crystal cell dimensions are observed even though the enzyme is in the closed conformation within the active site. In addition, with one of the dTTP analogue structures (T-CBr_2), the incoming nucleotide is not present and the binary enzyme-DNA complex is in the open conformation; this structure is not further discussed. Overall, the structures of the various ternary complexes with different incoming bases are similar to each other and the reference structures with a normal incoming dNTP. In each case, the incoming base is Watson-Crick base paired with the template base and the sugar and Pα are in similar positions. However, close inspection reveals subtle differences within the active sites that are relevant in explaining catalytic differences observed with these analogues[11–16].

To facilitate description of the structures, mapping of active site contact distances to be discussed is illustrated in Figure 1 and the description and significance is summarized in Table S2. Summaries of these contact distances are provided in the (Figures 3, 7, and 10) and (Tables S3, S5, and S7) for dATP, dTTP, and dCTP analogues, respectively. In these Figures and Tables, the contact is listed at the top and the incoming dNTPs are listed at the bottom, in the order of reference structure(s) on the left to dihalogen analogues on the right. In the figures, distances more than 0.2 Å ($1/10^{th}$ of the average resolution) longer or shorter than those in the reference structure are considered as significantly different compared with the reference structure and are color-coded in red or green, respectively. For example, in the incoming dATP structures (Fig.3), some interactions have shorter distances (coded green) than in the reference structure and other interactions have longer distances (coded in red); interactions that are similar to the reference structure are shown in black. In discussing the structures, special consideration is given to the interactions summarized in Figure 1 and Table S2, including four active site arginine side chains: R254, R283, R149 and R183.

dATP Analogues:

The active site of a new high-resolution structure (1.8 Å) with normal incoming dATP is illustrated in figure 2. Many substrate and enzyme interactions revealed in earlier structures of ternary complexes are seen in this new structure, including those in the reference structure 2FMS. The dATP structure shown in Figure 2 is used for orientation and reveals coordination of the nucleotide metal, the R183 guanidinium group N1 contact with a P_β oxygen, and the network of interactions around the Pγ oxygens. The structure has two water molecules bridging the guanidinium group N atoms of R149 to oxygens of Pγ. These water molecules have relatively low B-factors and stabilize Pγ. Other Pγ oxygen contacts illustrated in Figure 2 are with the main-chain at G189 and the side chain oxygen of S180, both stabilizing the Pγ oxygen, O3G. R254 has close contacts with both D256 and an oxygen in the primer terminus DNA backbone. In addition, an important stabilizing contact distal to the incoming dNTP binding pocket is noted in Figure 2, i.e., between R182 and E316.

An overlay (Fig. S2) of the ternary complex structures with five of the dATP analogues (CH_2, CHF, CHCL, CFCL, and CBr_2) reveals similarity to the reference structure and to each other (rmsd values 0.1 – 0.4 Å). In contrast, the structures with the dichloro- and difluoro-dATP analogues exhibit differences compared with the other analogues as follows: I) In both structures, a halogen makes contact with an R183 guanidinium N atom, and this

results in displacement of the enzyme backbone and a change in the location of R183; these differences are illustrated in an overlay of the active sites of the dATP and A-CF_2 structures (Fig. 4A). The figure illustrates that, in addition to R183, several residues are in different positions in the A-CF_2 structure, i.e., S180, R283, E316, and R182. Differences in interactions with Pγ oxygens and the position of Pγ are further illustrated in Figure 4B. A subtle backbone displacement radiates through the active site, and the displacement at R183 causes a conformational change in the entire C-terminal region of the enzyme distal to the active site. With the change at R182, its guanidinium group now fails to make contact with E316; this contact is completely lost in both dA-CF_2 and dA-CCl_2 structures (Fig. 4A), and loss of the stabilizing interaction is associated with distortion of the loop region near E316 (Fig. S3). This change is responsible for the altered cell dimensions noted above (Table S1). The overall result of this dATP dihalogen-mediated distortion is that the open to closed conformational transition of the enzyme is altered in the vicinity of E316 and the nearby loop (Fig.3; Fig. 5); and II) there are subtle differences, versus the reference, within the active sites of structures with the dATP dihalogen compounds. As summarized in Figure 3, these differences involve a “looser fit” for key interactions: the C3’ to Pα distance; the Y271 and primer base contact; and the N279 incoming base contact. In contrast to these dihalogen analogues, the structures of the dATP monofluoro- and monochloro-analogues are similar to the reference structure for these three interactions. In the structures with A-CFCl, A-CBr_2 and A-CCl_2 , the interactions between the oxygens of Pγ and R149 are different than those in the reference structure (Fig. 3). This is illustrated in Figure 4B from the structure of the A-CCl_2 analogue; the Pγ oxygens have rotated and the intervening water molecules seen in the dATP reference structure are not present. In the overlay in Figure 5, differences between the A-CCl_2 and A-CF_2 structures are illustrated. In the A-CCl_2 structure, the position of the catalytic magnesium has shifted (~1.2 Å) and the R254-D256 interaction is different in the A-CCl_2 structure, and a R149-oxygen distance (Fig. 3) is less in A-CCl_2 than in A-CF_2. As noted, the bridging water molecules seen in the reference structure are not present in the dihalogen structures.

Next, it is interesting to consider the structures with the A-CH_2 and A-$CHCH_3$ analogues, since these compounds are relatively more basic than the others. The active site regions are similar to the reference structure (Fig. 6); however, some of the interactions are tighter than in the reference structure. As shown in Figure 3, these are in the nascent base pair region around the primer sugar and Pα, between Y271 and the primer base, and in the R283 contact with the template strand (Tn-1) sugar oxygen (O4). Regarding the contact distances: In the A-CH_2 structure, the distance between the guanidinium group of R149 and PγO3G is shorter, as is the R254 and primer strand oxygen interaction; in the A-$CHCH_3$ structure, the distances between C3’ of the primer terminus and Pα-dNTP, OH-Y271 and O2-dNTP, and S180 and O3G-dNTP are shorter (Fig. 3). These differences are likely important in explaining the reduced enzymatic described in the accompanying article[14].

dTTP Analogues:

The crystallographic statistics for the dTTP analogue structures are shown in Table S4. The structures are similar to each other and to the 2FMS reference structure (Fig. S4), with rmsd ranging from ~0.2 to ~1.1 Å (Table S5).

Three monofluoro-analogue structures (the diastereomeric mixture and separated isomers) reveal stereoselective interaction between the R183 guanidinium N and fluoride atoms; in these cases, the fluoride atom of the R isomer interacts with the guanidinium N (Fig. 8A). For the diastereomeric mixture, the fluoride atom occupancy of the R isomer is illustrated, and there is no binding of the S isomer. The structures with the chloride analogues (Fig. 8B) also exhibit stereoselective interaction between the chloride atom and the guanidinium N of R183. In the case of the purified R isomer, the occupancy is 100%; for the purified S isomer, the chloride atom is pointing away from R183 and there is 100% occupancy (Fig.8B). For the T-CHCl mixture, the Cl atom R isomer occupancy is ~60 %, but the S isomer is not observed.

For the dihalogen dTTP analogues, the active site structure of the T-CF_2 analogue resembles the structure of the A-CF_2 analogue (c.f., Fig.3 and Fig. 7). Comparison of the T-CF_2, T-CFCl and T-CCl_2 structures reveals they are similar, but coordination of the Pγ oxygen O2G is different (Fig. 9). Interestingly, the structures for the dTTP difluoro- and dichloro-analogues do not exhibit the C-terminal region distortion observed with the dATP dihalogen analogues. Finally, differences are observed in the interactions at R254 (Fig. 7 and Table S5). In the case of the T-$CHCH_3$ analogue, R254 has rotated away from D256 and is in the position of an alternate rotamer observed previously in the D256E mutant enzyme[22]. Large changes in the positions of R254 also are seen with both monochloro-analogues (R and S) (Fig. 7 and Table S5).

dCTP Analogues

Crystallographic statistics for the dCTP reference and analogue structures are shown in Table S6. The resolution ranged from 1.8 to 2.02 Å, and virtually all of the contacts in the analogue structures are similar to the reference structure, with rmsd ranging from ~0.35 to 0.4 Å. Interaction distances in these structures are summarized in Figure 10 and Table S7, and the active site of a new dCTP reference structure is illustrated in Figure S6. Familiar active site contacts are observed in this reference structure, including contacts between a Pγ oxygen and the main-chain at G189 and side chain oxygen of S180. In the case of dCTP, the guanidinium group of R149 is too far away from the Pγ oxygens to interact directly and intervening water molecules are not observed (Fig. 10). In the analogue structures, the contacts among R183, R182 and E316 lack distortions and are similar to those in the reference structure with normal incoming dCTP. However, as illustrated in Figure 11, interactions in several of the analogue structures between R149 and the Pγ group are different compared with the reference structure. For these analogue structures, the R149 guanidinium group makes contact with PγO3G through water molecules, and the contact between the S180 side chain oxygen and the Pγ group oxygen is shorter in most cases. These differences in interactions are detailed in Figure 10. Alternate positions in R254 were not observed in these dCTP analogue structures, and the contact between the R183 guanidinium group and PβO2G is similar to the reference structures. For the C-CHF mixture of distereoisomers, both isomers are observed with the fluoride atom pointing toward R183 at 64% occupancy or pointing away from R183 at 36% occupancy. For the purified R isomer of C-CHF, the fluorine atom interacts with R183 and occupancy is 100%. In the cases of the

two individual C-CHCl structures, the S isomer chloride atom points away from R183, and the R isomer structure shows the chloride atom interacting with R183 at 61% occupancy.

Discussion:

Overview

We find that pol β ternary complex crystal structures with a range of incoming dNTP 5′-triphosphate group analogues are similar to the reference structures with oxygen between Pβ and Pγ of the 5′-triphosphate group. Therefore, large distortions in the active site do not explain the pol β kinetic differences described in the accompanying article[14]. Instead, unique stabilizing contacts and charge differences in the analogues account for the kinetic results. Subtle differences in active site interactions are revealed, especially with interactions around the Pγ oxygens (Fig. 2). The results of the pol β enzymatic rate constant measurements described in the accompanying article[14] illustrate that the k_{pol} values are generally higher with the more acidic analogues, and when the values are plotted as a function of leaving group pK_a (i.e., in the Brønsted plot format) there is an approximately linear relationship between the $\log_{10} k_{pol}$ versus the pK_a value for the leaving group. In addition, for the dihalogen compounds, the relationship between k_{pol} and pK_a is different from that of the reference dNTPs and other analogues; in the former case, there is a shift in the relationship toward lower pK_a values. These differences with the dihalogen analogues[14] correlate with structural differences in the C3′ to Pα distance, the PγO3G to R149 distance, the position of the R149 guanidinium group, and with the presence or absence of stabilizing water molecules between the Pγ oxygens and the R149 guanidinium group. In addition, a trend among the structures with the monohalogen analogues correlates with the kinetic results: In this case, when the halogen atom interacts with the nearby R183 guanidinium group NH_2, a higher k_{pol} is observed.

It is interesting to compare the active sites in the structures with the difluoro-analogues of dATP, dTTP and dCTP. These structures are generally similar to each other (Fig. 12), but more differences compared with the reference structure are seen with the dATP analogues. It is likely that the larger adenine base introduces a degree of strain resulting in greater sensitivity to the acidity differences in the panel of Pβ-Pγ substituted analogues. The positions of Pγ and coordination of its oxygens are different in the cases of the A-CF_2, A-CCl_2 and A-CFCl structures, and there is a difference in the position of the R149 side chain, and the structures of the difluoro- and dichloro-analogues differ from each other in the region of the respective Pγ oxygens. These differences are illustrated in Figure 4A and B with structures of the dATP analogues. The structures of dATP analogues A-CF_2 and A-CCl_2 are similar, but there are differences between them; for example, the nucleotide metal is in a different location in the A-CCL_2 structure and interactions at the respective Pγ oxygens are different (Fig. 5). These and other structure-activity correlations will be discussed below.

Structure-Activity Considerations

Incoming dATP Analogues:

Key architectural features of the active site with incoming dATP and two metal ions, magnesium and sodium, are summarized in Figures 1 and 2. Briefly, the template base and adenine base are stabilized opposite each other by hydrogen bonds and stacking interactions with the adjacent template (G) and primer bases (C), as well as by van der Waals' interactions with the side chains of R283 and Y271. The template and primer bases are also stabilized by virtue of side chain hydrogen bond interactions with R283, N279 and Y271, respectively. Surrounding the triphosphate group of incoming dATP, an extensive network of enzyme-substrate and metal-enzyme-substrate interactions is evident, and contacts near the Pβ-Pγ bridging atom are expected to be sensitive to charge effects mediated by the electron withdrawing and donating properties of the dNTP analogues. For example, the guanidinium groups of R183 and R149 are proximal to the Pβ and Pγ groups, respectively, and contacts between the side chain oxygen of S180 and backbone nitrogen of G189 are evident. The nucleotide metal makes contacts with oxygens of Pβ and Pγ and side chain oxygens of D190 and D192. Thus, there are many opportunities for charge effects in substitutions at the Pβ-Pγ bridging position to influence chemistry.

Comparisons between the kinetic results with the acidity-modified analogues and the crystal structures are interesting to consider[14]. First, with each type of incoming base, there is striking similarity in structures across a range of analogues exhibiting linear k_{pol} vs pK_{a4} relationships, and this allows us to conclude that the k_{pol} differences involve acidity effects rather than gross conformational differences in the active site. Thus, structures of the dATP reference compound and dATP analogues are similar, while there is a 10-fold difference in the k_{pol} values, with the most electrophilic analogue A-CF_2 exhibiting more activity than the least electrophilic analogue A-CH_2. Second, some of the k_{pol} values deviate from the line average describing the k_{pol} vs pK_{a4} relationships. For example, the k_{pol} value for A-CCl_2 falls significantly below the line average. This lower than expected k_{pol} value with A-CCl_2 correlates with displacement of the nucleotide metal by the Cl atom, as illustrated in Figure 5. Another example of deviation from the line average is seen with the most alkaline analogue, A-$CHCH_3$. The crystal structure with this compound reveals differences by virtue of closer contacts in the vicinity of the incoming nucleobase (i.e., C3'-Pα, P10-Y271, and Tn-1-R283) and in the R149 contact with PγO3G; these differences correlate with the lowest k_{pol} values among all of the dATP analogues. Third, the C-terminal distortion (Fig. S3) seen with A-CF_2 does not correlate with a k_{pol} difference, suggesting this C-terminal distortion is independent of adverse effects on k_{pol}. This information is useful because a similar C-terminal region distortion has been observed with a cancer associated variant (K289M) of pol β (Unpublished observations). Fourth, differences in the interactions of the Pγ oxygens and R149 in the dihalogen analogues are evident, compared with the reference dATP. This is illustrated in Figure 4B where the Pγ oxygens and R149 side chain are in different positions compared with the reference structure. All four dihalogen analogue structures reveal a contact between R183 and a Pβ-Pγ halogen atom, and the S180 and R149 side chains are displaced. These differences may contribute to the distinct grouping of the dihalogen analogues in the k_{pol} vs pKa relationship.

The Incoming dTTP Analogues:

Structures of the thymine analogues exhibit subtle differences in the vicinity of Pγ and the contact with R183 (Fig. 10). Assessment of the k_{pol} vs pK_{a4} relationship with incoming dTTP and the dTTP analogues reveals clear separation into two groups, each with the more acidic compounds exhibiting higher k_{pol} values than the less acidic compounds[14]. The k_{pol} vs pK_{a4} relationship is linear for each group, with the dihalogen analogues in one group and the remaining analogues in the other. As for the dATP analogues, the dTTP structural results allow us to conclude that the observed k_{pol} differences are due mainly to charge effects, rather than gross conformational differences. Only 2 of the analogues exhibit lower activity than expected from the line average. These two cases are with the S-isomers of the monochloro- and monofluro-analogues; the k_{pol} value for each of these is lower than that of the corresponding R-isomer, which exhibits interaction with R183. These results indicate the R-isomer halogen interaction with R183 correlates with a higher k_{pol}.

Further, the difference in k_{pol} value between T-CF_2 and T-CCl_2 is about 10-fold, and the main difference in the two structures is a bridging water molecule between R149 and PγO3G with T-CCl_2, but not with T-CF_2 (Fig. 12). This bridging water in T-CCl_2 could neutralize or block a charge effect that is beneficial for activity. Interestingly, the nucleotide metal displacement discussed above with A-CCl_2 is not seen with the T-CCl_2 analogue where the k_{pol} value falls squarely on the average line, while the A-CCl_2 k_{pol} value falls below the average line.

Another feature emerges from comparison of the structures of the most alkaline T-CCH_3 and A-CCH_3 compounds. The structure of the T-CCH_3 analogue does not reveal the tighter fit in region of O3', as seen with A-CCH_3, and the T-CCH_3 k_{pol} value falls on the average line instead of below it as with A-CCH_3. This comparison suggests that the tighter fit in the vicinity of O3', as seen with A-CCH_3, is a factor in its unusually low k_{pol} value. Finally, with the two most alkaline dTTP analogues, a much closer contact is seen between R149 and PγO3G, and this is associated with lower activity for these compounds.

The Incoming dCTP Analogues:

The structure with the dCTP reference (Fig. S6) is similar to the structures with the other reference dNTPs (c.f., Figs. 2, S4), except that R149 is too far away from PγO3G (5 Å) to interact directly, and a bridging water molecule(s) between the Pγ oxygens and R149 is not present. In contrast, all of the halogen containing dCTP analogues have bridging water contacts between PγO3G and R149, and this pattern of bridging water contacts with Pγ and R149 is associated with a halogen atom interaction with R183, in isomers where the halogen atom points toward the guanidinium group. These interactions with R149 are illustrated in Figure 12, and interactions with R183 are illustrated in Figure 11. It is interesting to note that all 6 analogues with this halogen/R183/and Pγ/water/R149 network of interactions exhibit higher k_{pol} values than dCTP, lacking these interactions[14]. Thus, the halogen interaction at R183 and stabilizing water interactions at R149 are associated with higher k_{pol} values; this conclusion is reinforced by comparisons of the R and S isomers, where the R183 interacting R- isomers have higher k_{pol} values. Finally, two of the dihalogen analogues, C-CCL_2 and C-Br_2, have lower k_{pol} values than dCTP. In these two cases, PγO3G has a

bridging water interaction to R149 and the structures are virtually identical to those of the more active analogues, suggesting that acidity effects in these dichloro- and dibromo-analogues account for the weaker activity observed.

Molecular Insights into Complex Brønsted Plots:

DNA polymerases employ a stepwise mechanism to bind dNTPs (Scheme 1). In this scheme, dNTP binds to E-DNA$_n$ (K_1) to form an initial complex. This is followed by conformational adjustments to yield the activated ternary substrate complex (E*-DNA$_n$-dNTP). In its simplest arrangement, these adjustments are depicted as a single step generating a unique pre-catalytic ternary substrate complex (K_2) that undergoes chemistry (k_3). Since the forward nucleotide insertion reaction predominates, the pyrophosphorolysis (reverse reaction) is negligible[23]. Importantly, changes in the measured insertion rate (k_{pol}) can occur even though the magnitude of the chemical step in not altered[24, 25]. This is because k_{pol} is dependent on k_3 (chemistry) and the concentration of activated ternary complex (E*-DNA$_n$-dNTP) (i.e., $k_{pol} = k_3$*[E*-DNA$_n$-dNTP]). Accordingly, k_{pol} is sensitive to changes in the concentration of pre-catalytic ternary substrate complex, as well as to the magnitude of chemistry. A change in the activated ternary substrate complex concentration could occur due to altered enzyme equilibria or altered substrate conformations in the ternary complex (Scheme 1:, K*). This would in turn alter the measured value of k_{pol}. Rather than the simple linear diagram[14], perhaps a more encompassing representation of the nucleotide insertion chemistry involves a “generalized” three-dimensional reaction coordinate diagram as proposed previously[26] ; Florian, 2005 #102}. This third dimension could in principle reflect the impact of alternate pre-chemistry complexes that would in turn impact the concentration of productive ternary complexes and/or the chemical step.

If the β-γ modification of the incoming nucleotide only affected the barrier for the chemical reaction, then the observed Brønsted slopes should be independent of the identity of the incoming nucleotide. As reported in the accompanying work[14], and shown previously[15], the Brønsted plot slopes with acidity modified dNTPs are dependent on the identity of the nascent base pair. This indicates that precise base pair H-bonding and stacking interactions contribute to the free energy relationships. Although the binding affinities of the analogues are similar to the natural nucleotide, they are not identical suggesting altered binding (K_1 and/or K_2). Nucleotide binding is initiated through triphosphate/Mg^{2+} interactions in the open polymerase conformation[27]. In this state, the Watson-Crick edge of the correct base, but not an incorrect base, of the incoming nucleotide is within H-bonding distance to the templating base, although the geometry is not optimal and base stacking is poor. Thus, triphosphate/Mg^{2+} initiates nucleotide binding. Subsequently, when the polymerase closes upon the nascent base pair, base stacking with the enzyme and primer terminus, as well as Watson-Crick H-bonding interactions in the nascent base pair, are optimized. Triphosphate modifications could alter kinetic steps at or after initial binding, but before chemistry (k_3). Thus, chemistry may be unchanged, but an altered k_{pol} would be a reflection of an altered concentration of productive ternary complex; this may occur even though there is not a change in the rate determining step of the overall insertion reaction. Recently published LFER plots with the K289M pol β mutant are consistent with this interpretation[16]. This mutant exhibits Brønsted slopes with modified dNTPs that are not as dependent on pK_{a4}

(i.e., slopes are less negative than wild-type enzyme). It was suggested that the methionine mutation at Lys289 decreased the stability of the closed form. Likewise, it is not to be unexpected that modification at the bridging position between Pβ and Pγ may also influence interconversion or stability of alternate enzyme forms.

Our observation that subtle active site structural differences are observed that depend on the identity of the incoming nucleotide and the triphosphate modification, suggest that the reaction coordinate could be modified perhaps to include a third dimension[26, 28], as discussed above. In addition to the heterogeneity observed in the transition state for the chemical step, the reaction coordinate could also include varying activation barriers *prior* to the chemical step to account for alternate forms of the pre-catalytic ternary complex. This conformational heterogeneity, leading to alternate forms of pre-catalytic complexes as evident from the structures described above and the omit maps shown in Figure S8, is likely responsible of the deviations from the LFER lines seen with the various analogues. These structural differences are summarized in detail above. Taken together, the results indicate that the Brønsted slopes are sensitive to the protonation state of Pγ of the incoming nucleotide, stacking and H-bonding interactions for the nascent base pair, and dynamic equilibria between alternate forms of the pre-catalytic ternary substrate complex.

After binding DNA to apoenzyme (not indicated), the nucleotide triphosphate binds forming an initial ternary complex (K_1). The polymerase/substrate complex (circle) undergoes rapid conformational adjustments that lead to a productive ternary substrate complex (K_2, square). Nucleotide insertion (k_3) leads to a post-chemistry product ternary complex with extended DNA (DNA_{+1}) and PP_i. Enzyme conformational changes may occur that alter substrates (S*) that deter chemistry (i.e., form inactive complexes; red square). Alternate conformations (n) of the DNA, dNTP, and/or metal binding site could exist. Similar steps after nucleotide insertion, indicated in gray, are also indicated. Indeed, altering the bridging oxygen in PP_i alters post-chemistry conformational equilibria and the rate-determining step for the reverse reaction[23].

Implications for probe/inhibitor development

The structures with the purified monofluoro- and monochloro-diastereoisomers across the panel of analogues reveal that the R-isomer halogen atom interacts with R183, whereas the S-isomer halogen atom points away from R183. The R-isomers have higher k_{pol} and lower K_d values than the corresponding S-isomers, such that the halogen interaction with R183 is associated with higher enzyme efficiency (k_{pol}/Kd) values. In the cases of the C and T analogues, it is interesting that the halogen-R183 distance is shorter in the dihalogens than in the monohalogens (i.e., 3 to 3.1 Å vs. 3.3 to 3.5 Å), suggesting stronger binding. The presence of at least one bridging water molecule between R149 and Pγ oxygens is associated with higher k_{pol} values in most cases, suggesting that it may be useful to examine analogues with modifications at the Pγ oxygens that mimic the water interaction. The tighter fit (i.e., closer interaction between R149 and PγO3G) seen with the CH_2 analogues is associated with lower k_{pol} and efficiency values, and the possibility of reducing enzymatic activity using more alkaline compounds seems evident. This is the case with the A-$CHCH_3$

analogue, for example, where there is an unusually tight fit around the nucleobase and between R149 and PγO3G and this is associated with unusually low activity.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

Use of the Advanced Photon Source was supported by the US Department of Energy, Office of Science.

Funding Sources

This research was supported by the Division of Intramural Research of the NIH, National Institute of Environmental Health Sciences [project numbers Z01 ES050158 and ES050159], and was in association with NIH grant 1U19CA105010.

ABBREVIATIONS

dNTP	deoxy nucleotide triphosphate
dATP	deoxy adenosine triphosphate
dCTP	deoxy cytidine triphosphate
dTTP	deoxy thymidine triphosphate
LFER	linear free energy relationship

References

[1]. Bessman MJ, Lehman IR, Simms ES, and Kornberg A (1958) Enzymatic synthesis of deoxyribonucleic acid. II. General properties of the reaction, J Biol Chem 233, 171–177. [PubMed: 13563463]

[2]. Bebenek K, and Kunkel TA (2004) Functions of DNA polymerases, Adv Protein Chem 69, 137–165. [PubMed: 15588842]

[3]. Steitz TA (1999) DNA polymerases: structural diversity and common mechanisms, J Biol Chem 274, 17395–17398. [PubMed: 10364165]

[4]. Beese LS, and Steitz TA (1991) Structural basis for the 3'−5' exonuclease activity of Escherichia coli DNA polymerase I: a two metal ion mechanism, EMBO J 10, 25–33. [PubMed: 1989886]

[5]. Batra VK, Beard WA, Shock DD, Krahn JM, Pedersen LC, and Wilson SH (2006) Magnesium induced assembly of a complete DNA polymerase catalytic complex, Structure 14, 757–766. [PubMed: 16615916]

[6]. Sawaya MR, Prasad P, Wilson SH, Kraut J, and Pelletier H (1997) Crystal structures of human DNA polymerase β complexed with gapped and nicked DNA: Evidence for an induced fit mechanism, Biochemistry 36, 11205–11215. [PubMed: 9287163]

[7]. Freudenthal BD, Beard WA, Shock DD, and Wilson SH (2013) Observing a DNA polymerase choose right from wrong, Cell 154, 157–168. [PubMed: 23827680]

[8]. Perera L, Beard WA, Pedersen LG, and Wilson SH (2017) Hiding in Plain Sight: The Bimetallic Magnesium Covalent Bond in Enzyme Active Sites, Inorg Chem 56, 313–320. [PubMed: 27976880]

[9]. Lin P, Pedersen LC, Batra VK, Beard WA, Wilson SH, and Pedersen LG (2006) Energy analysis of chemistry for correct insertion by DNA polymerase beta, Proc Natl Acad Sci U S A 103, 13294–13299. [PubMed: 16938895]

[10]. Batra VK, Beard WA, Shock DD, Pedersen LC, and Wilson SH (2008) Structures of DNA polymerase β with active site mismatches suggest a transient abasic site intermediate during misincorporation, Mol. Cell 30, 315–324. [PubMed: 18471977]

[11]. McKenna CE, Kashemirov BA, Upton TG, Batra VK, Goodman MF, Pedersen LC, Beard WA, and Wilson SH (2007) (R)-beta,gamma-fluoromethylene-dGTP-DNA ternary complex with DNA polymerase beta, Journal of the American Chemical Society 129, 15412-+. [PubMed: 18031037]

[12]. Sucato CA, Upton TG, Kashemirov BA, Batra VK, Martinek V, Xiang Y, Beard WA, Pedersen LC, Wilson SH, McKenna CE, Florian J, Warshel A, and Goodman MF (2007) Modifying the beta,gamma leaving-group bridging oxygen alters nucleotide incorporation efficiency, fidelity, and the catalytic mechanism of DNA polymerase beta, Biochemistry 46, 461–471. [PubMed: 17209556]

[13]. Sucato CA, Upton TG, Kashemirov BA, Osuna J, Oertell K, Beard WA, Wilson SH, Florian J, Warshel A, McKenna CE, and Goodman MF (2008) DNA polymerase beta fidelity: Halomethylene-modified leaving groups in pre-steady-state kinetic analysis reveal differences at the chemical transition state, Biochemistry 47, 870–879. [PubMed: 18161950]

[14]. Oertell K, Kashemirov BA, Sweasy JA, Batra VK, Beard WA, Wilson SH, McKenna CE, and Goodman MF (2018) Probing DNA polymerase chemistry using acidity modified dNTP analogues. , Biochemistry.

[15]. Oertell K, Chamberlain BT, Wu Y, Ferri E, Kashemirov BA, Beard WA, Wilson SH, McKenna CE, and Goodman MF (2014) Transition State in DNA Polymerase beta Catalysis: Rate-Limiting Chemistry Altered by Base-Pair Configuration, Biochemistry 53, 1842–1848. [PubMed: 24580380]

[16]. Alnajjar KS, Garcia-Barboza B, Negahbani A, Nakhjiri M, Kashemirov B, McKenna C, Goodman MF, and Sweasy JB (2017) A Change in the Rate-Determining Step of Polymerization by the K289M DNA Polymerase beta Cancer-Associated Variant, Biochemistry 56, 2096–2105. [PubMed: 28326765]

[17]. Beard WA, and Wilson SH (1995) Purification and domain-mapping of mammalian DNA polymerase β, Methods Enzymol. 262, 98–107. [PubMed: 8594388]

[18]. Otwinowski Z, and Minor W (1997) Processsing of X-ray diffraction data collected in oscillation mode, Methods Enzymol. 276, 307–326.

[19]. Batra VK, Beard WA, Hou EW, Pedersen LC, Prasad R, and Wilson SH (2010) Mutagenic conformation of 8-oxo-7,8-dihydro-2 ‘-dGTP in the confines of a DNA polymerase active site, Nat Struct Mol Biol 17, 889–890. [PubMed: 20526335]

[20]. Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung L-W, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, and Zwart PH (2010) PHENIX: A comprehensive Python-based system for macromolecular structure solution, Acta Crystallogr. D 66, 213–221. [PubMed: 20124702]

[21]. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, and Ferrin TE (2004) UCSF Chimera—A visualization system for exploratory research and analysis, J. Comput. Chem 25, 1605–1612. [PubMed: 15264254]

[22]. Batra VK, Perera L, Lin P, Shock DD, Beard WA, Pedersen LC, Pedersen LG, and Wilson SH (2013) Amino acid substitution in the active site of DNA polymerase β explains the energy barrier of the nucleotidyl transfer reaction, J. Am. Chem. Soc 135, 8078–8088. [PubMed: 23647366]

[23]. Shock DD, Freudenthal BD, Beard WA, and Wilson SH (2017) Modulating the DNA polymerase beta reaction equilibrium to dissect the reverse reaction, Nat Chem Biol 13, 1074–1080. [PubMed: 28759020]

[24]. Beard WA, Shock DD, Batra VK, Prasad R, and Wilson SH (2014) Substrate-induced DNA polymerase β activation, J. Biol. Chem 289, 31411–31422. [PubMed: 25261471]

[25]. Schlick T, Arora K, Beard WA, and Wilson SH (2012) Perspective: pre-chemistry conformational changes in DNA polymerase mechanisms, Theor Chem Acc 131, 1287. [PubMed: 23459563]

[26]. Tsai YC, and Johnson KA (2006) A new paradigm for DNA polymerase specificity, Biochemistry 45, 9675–9687. [PubMed: 16893169]

[27]. Freudenthal BD, Beard WA, and Wilson SH (2012) Structures of dNTP intermediate states during DNA polymerase active site sssembly, Structure 20, 1829–1837. [PubMed: 22959623]

[28]. Florian J, Goodman MF, and Warshel A (2005) Computer simulations of protein functions: searching for the molecular origin of the replication fidelity of DNA polymerases, Proc Natl Acad Sci U S A 102, 6819–6824. [PubMed: 15863620]

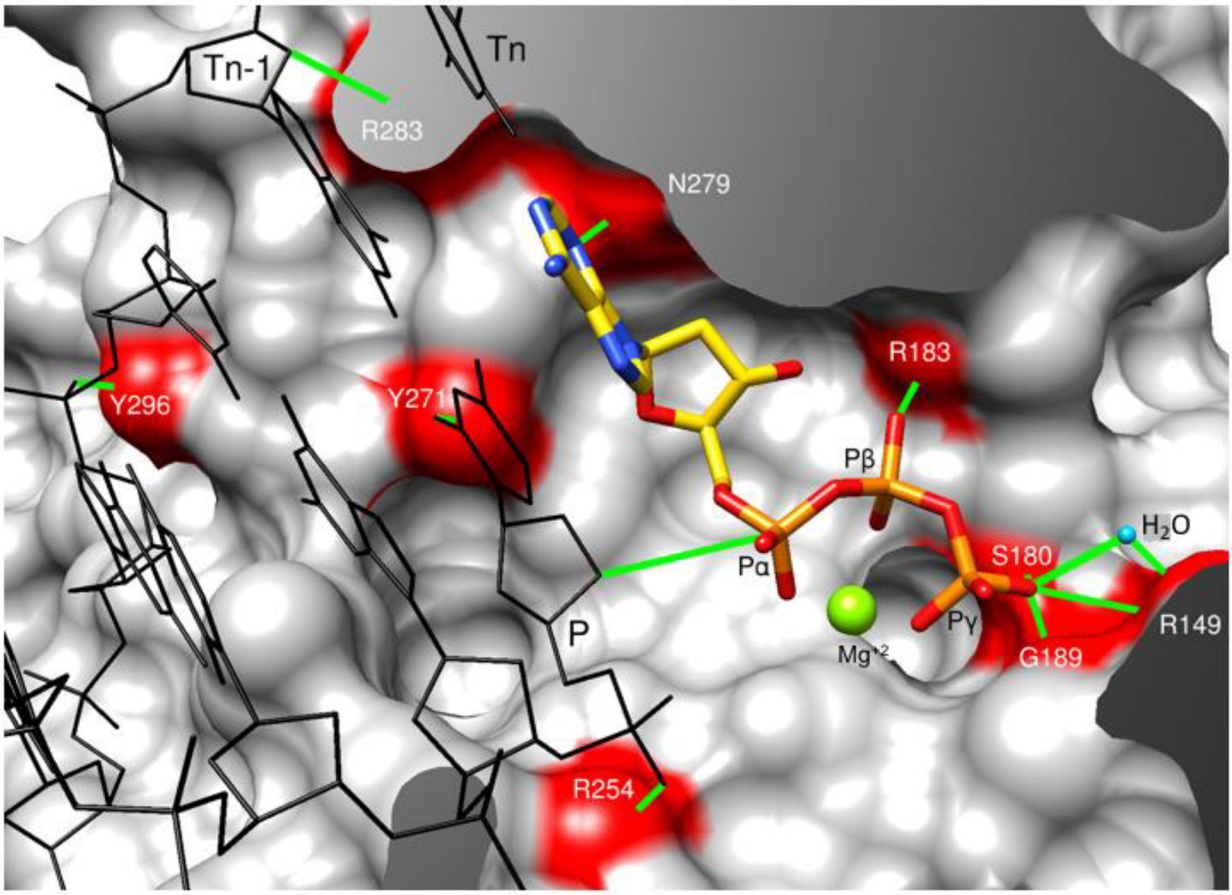


Figure 1:
A surface image of the DNA polymerase β active site illustrating many of the enzyme-substrate interactions (in green) to be discussed.

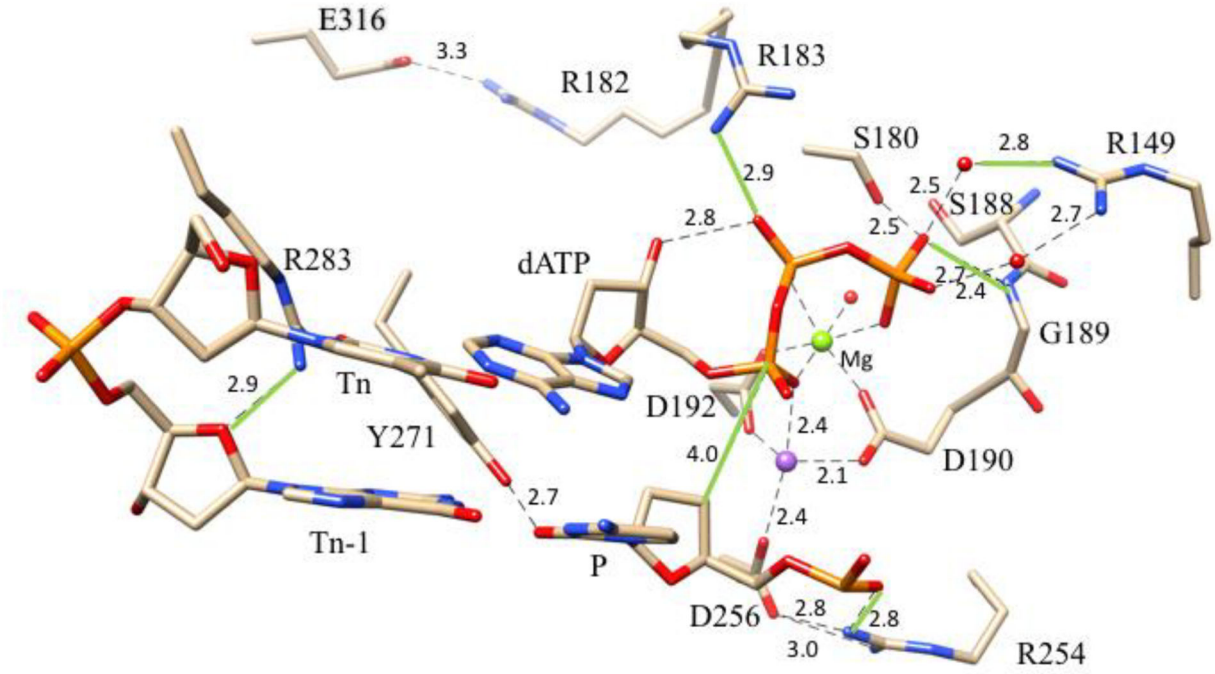


Figure 2: The active site of the DNA polymerase β reference structure (1.8 Å) with a dideoxy primer and incoming dATP.

The nucleotide metal, magnesium (green sphere), is shown, and the catalytic metal site is occupied by sodium (purple sphere). The catalytic and nucleotide metals were assigned as magnesium and sodium, respectively, based on the average coordination distances (shown in Figure S2) and the number of ligands. Coordinating water molecules are shown (small red spheres). Coordination distances (Å) are shown. Key active site amino acid residues are labeled, and the template (Tn) and primer (P) nucleotides are labeled. Two residues distal to the active site are also labeled, R182 and E316.

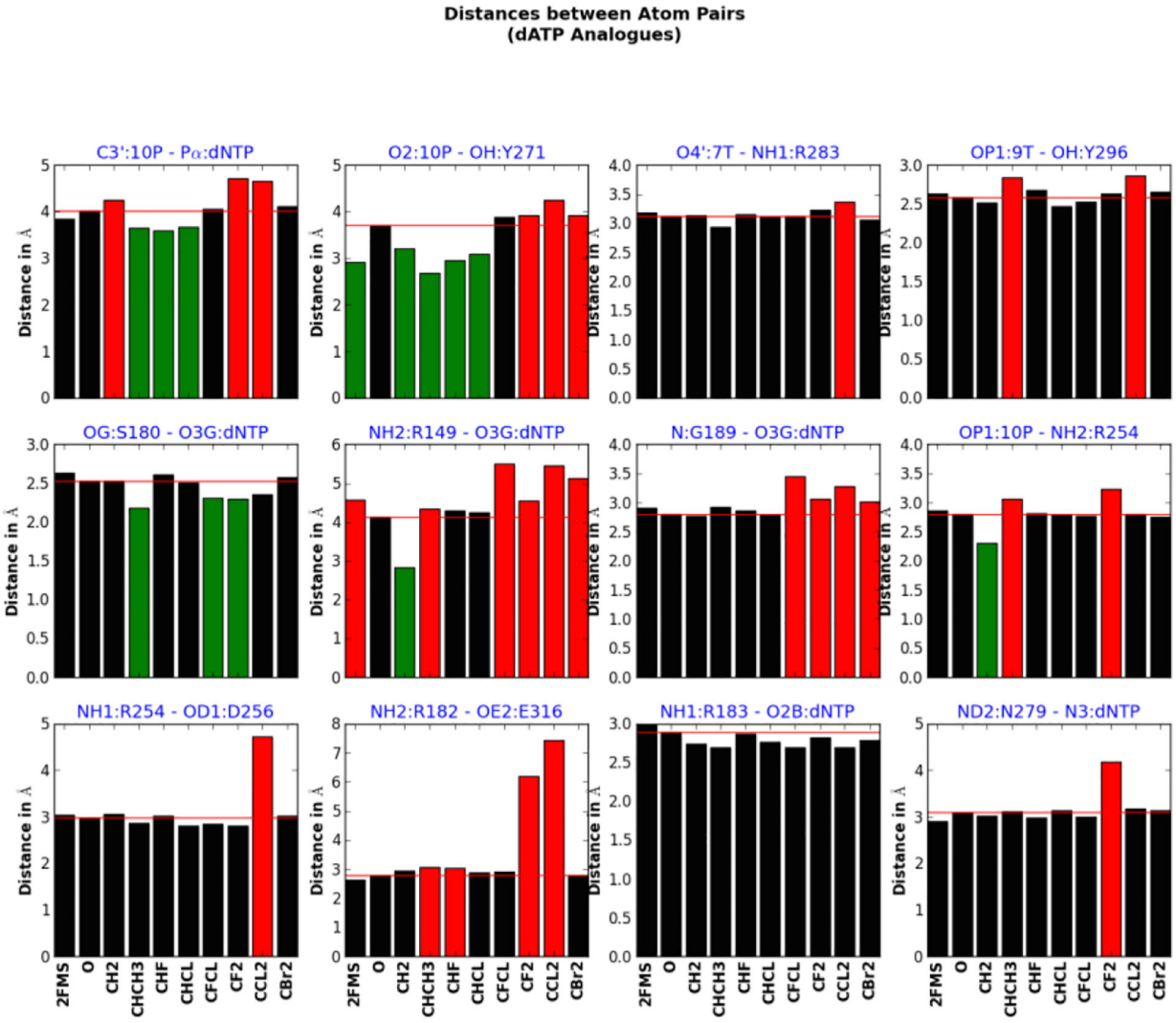


Figure 3: Bar diagram of main inter-atomic distances for dATP analogues

The distances that are similar to that of reference structure (dATP) are colored black, distances that are at least 0.2 Å longer, or shorter than that of reference structure are colored red and green respectively. The red horizontal line shows the distance in Å for the structure with the parent compound, with O in the β,γ-bridging position. The descriptions and significance of these measured distances are summarized in Table S2.

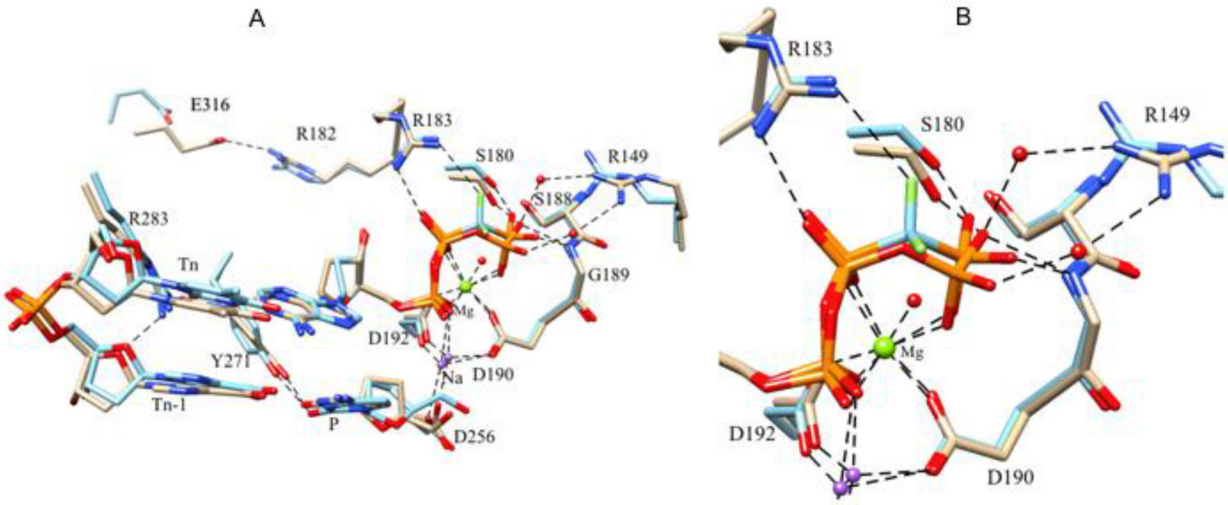


Figure 4: Overlay of active sites of DNA polymerase β with dideoxy primers with incoming dATP (light gray) and difluoro dATP analogue (cyan).

(A) The overlay was made with the 155–260 region of the enzyme. B) Magnification of the overlay in panel A illustrating the contacts around the gamma phosphate group. The CF_2 analogue F atoms are shown in light green.

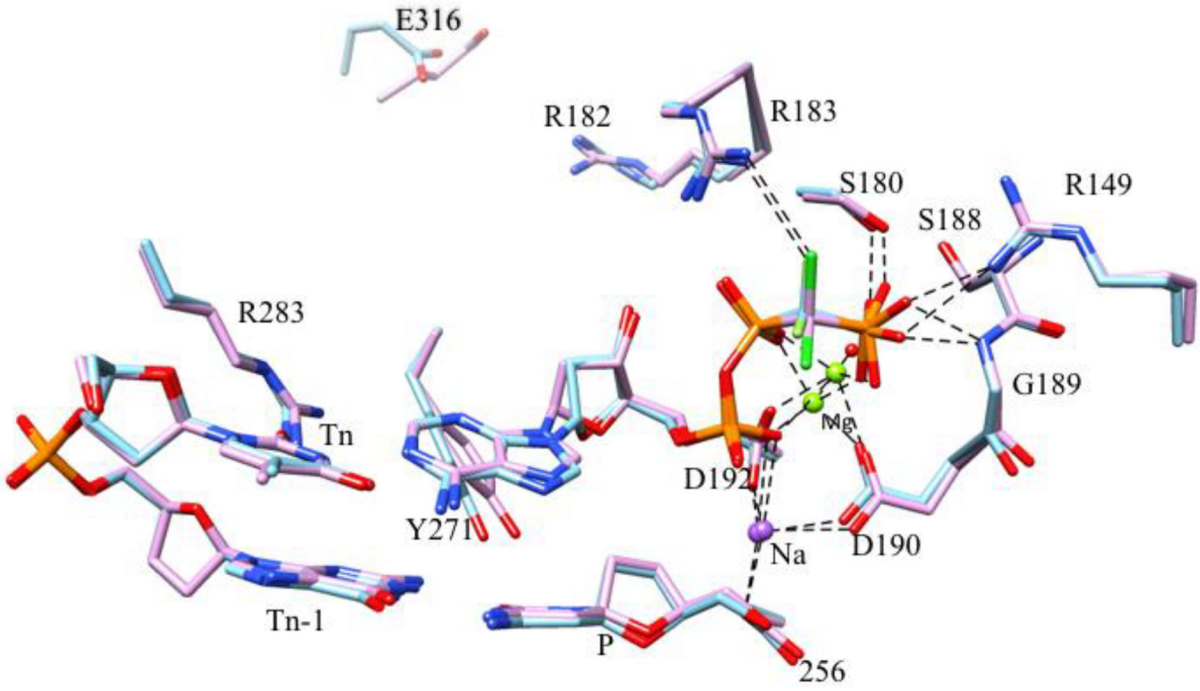


Figure 5: Overlay of the active sites of DNA polymerase β with dideoxy primers (p) and incoming dichloro (pink) and difluoro dATP (cyan) analogues.

The CCl_2 analogue Cl atoms are in dark green and the CF_2 analogue F atoms are in light green in the background. E316 no longer makes a contact with R182 in these structures, and R283 does not make a contact with the Tn-1 sugar ring O4 oxygen, as is seen in the reference structure. The nucleotide magnesium atoms are in a different location in these structures, and contacts with Pγ oxygens are different.

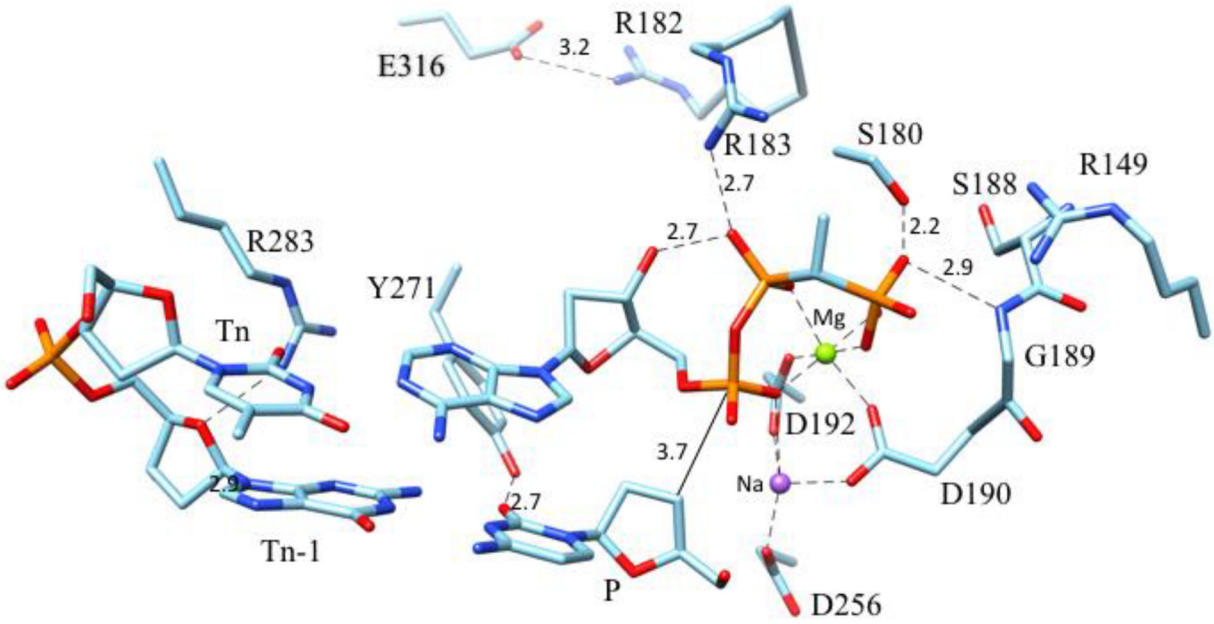


Figure 6: The active site of DNA polymerase β with the incoming A-CHCH$_3$ analogue of dATP.
The structure is similar to the reference dATP structure, but many contacts are shorter than in the reference structure (see Fig.3), including the primer sugar C3′ and Pα distance (3.7 Å), and in addition, the Pγ oxygens are too far away from R149 to make significant contacts.

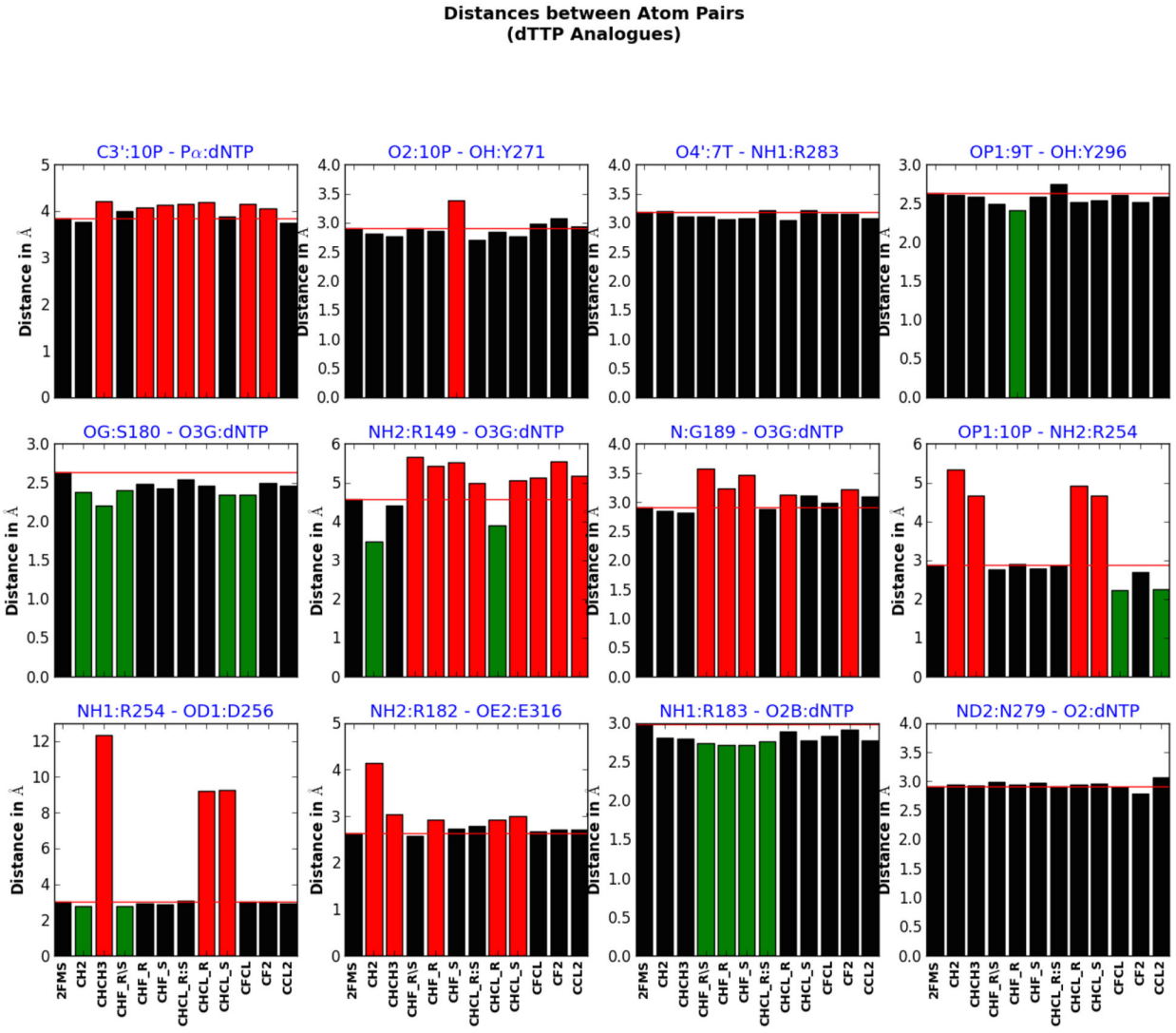


Figure 7: Bar diagram of main inter-atomic distances for dTTP analogues

The distances that are similar to that of reference structure (2FMS) are colored black, distances that are at least 0.2 Å longer, or shorter than that of reference structure are colored red and green respectively. The red horizontal line shows the distance in Å for the reference structure (2FMS). The descriptions and significance of these measured distances are summarized in Table S2.

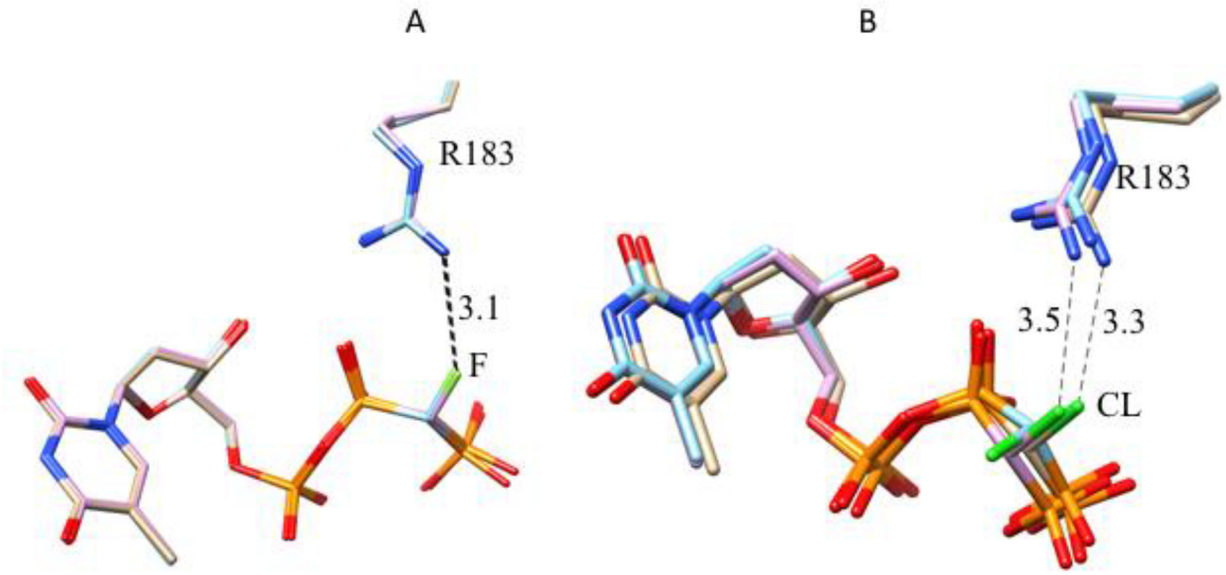


Figure 8: Comparisons of F and Cl monohalogen analogue structures of dTTP.

A) Overlay of R-isomer (pink, slow on HPLC), S-isomer (cyan, fast on HPLC), and diastereoisomer mixture (gold) of CHF analogues of dTTP in the active site of DNA polymerase β. B) Overlay of the R-isomer (pink, slow on HPLC), S-isomer (cyan, fast on HPLC), and diastereoisomer mixture (gold) of T-CHCL analogues of dTTP in the active site of pol β. These overlays were made with the 155–260 region of the enzyme. The distances (Å) for the halogen and R183 contact are shown.

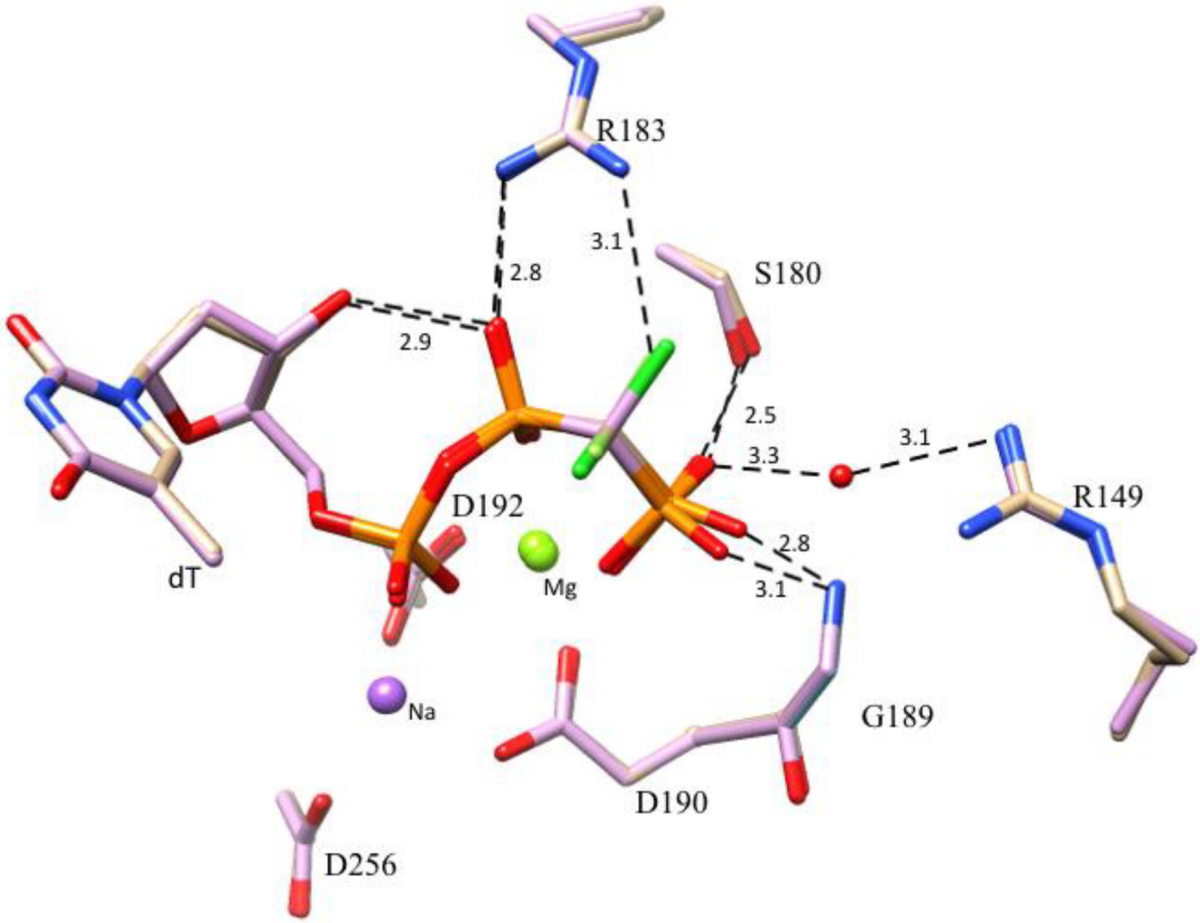


Figure 9: Overlay of active site structures of T-CF_2 (gold) and T-CCL_2 (pink) analogues.

A significant difference in the interaction with R149 is illustrated. A water molecule (red sphere) connects the guanidine group of R149 to the Pγ O3G of the T-CCl_2 analogue structure, but not in the case of the T-CF_2 structure. The two main-chain contacts at G189 are different in the two structures with 3.1 Å in the structure of T-CF_2. The catalytic and nucleotide metals were assigned as magnesium and sodium respectively based on the average coordination distances (shown in Figure S5) and the number of ligands.

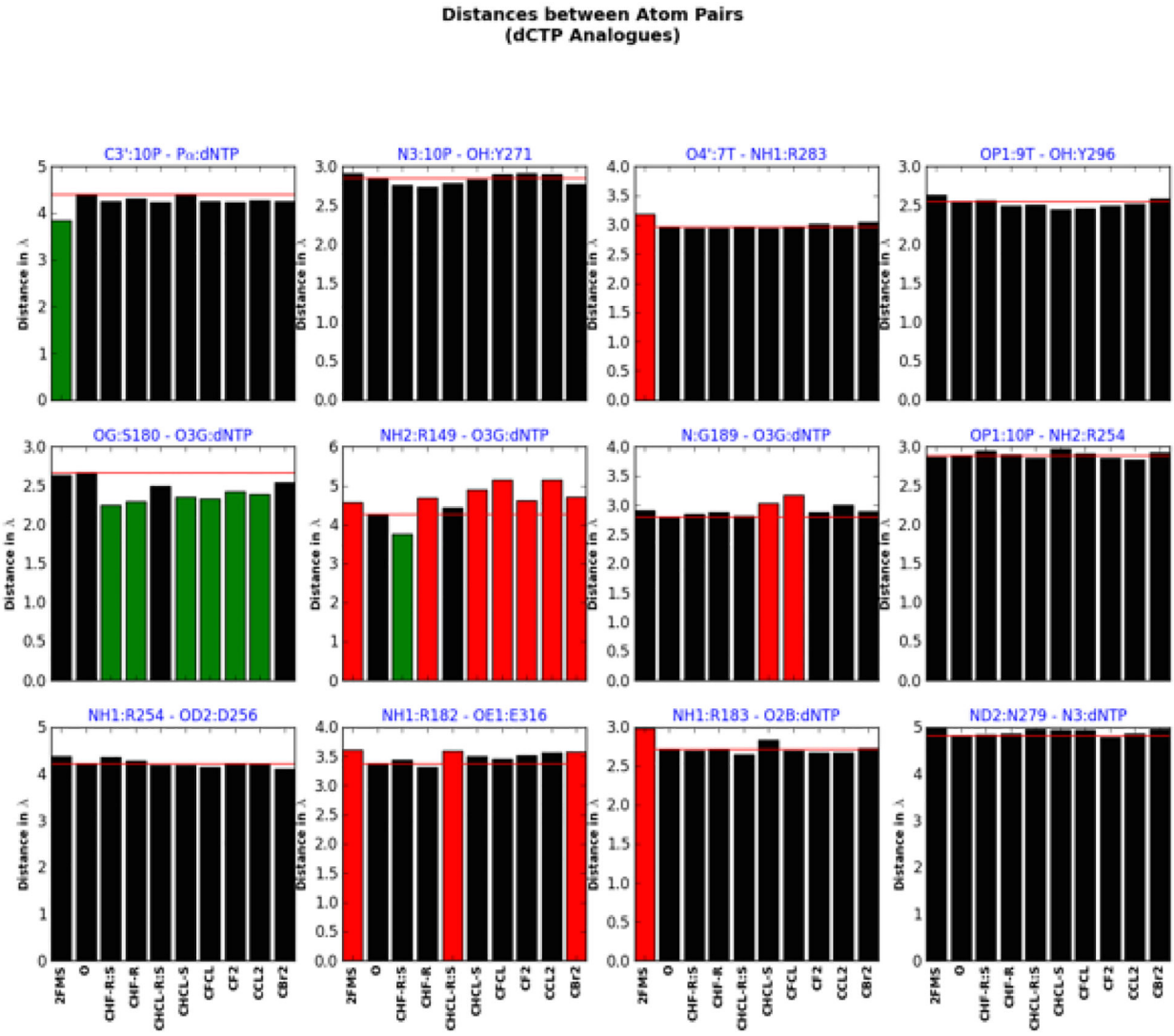


Figure 10: Bar diagram of main inter-atomic distances for dCTP analogues

The distances that are similar to that of reference structure (dCTP) are colored black, distances that are at least 0.2 Å longer, or shorter than that of reference structure are colored red and green respectively. The red horizontal line shows the distance in Å for the reference structure (2FMS). The descriptions and significance of these measured distances are summarized in Table S2.

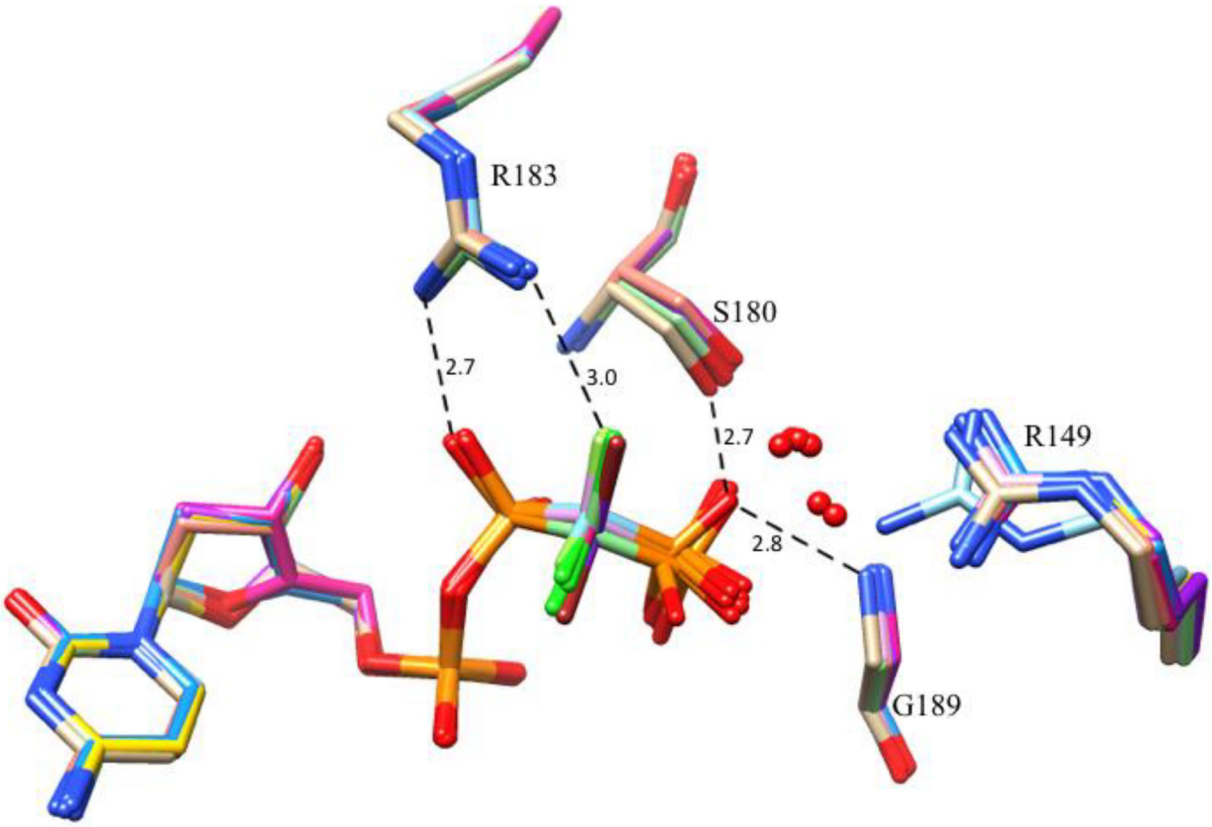


Figure 11:
Overlay of active sites of DNA polymerase β with incoming dCTP analogues.

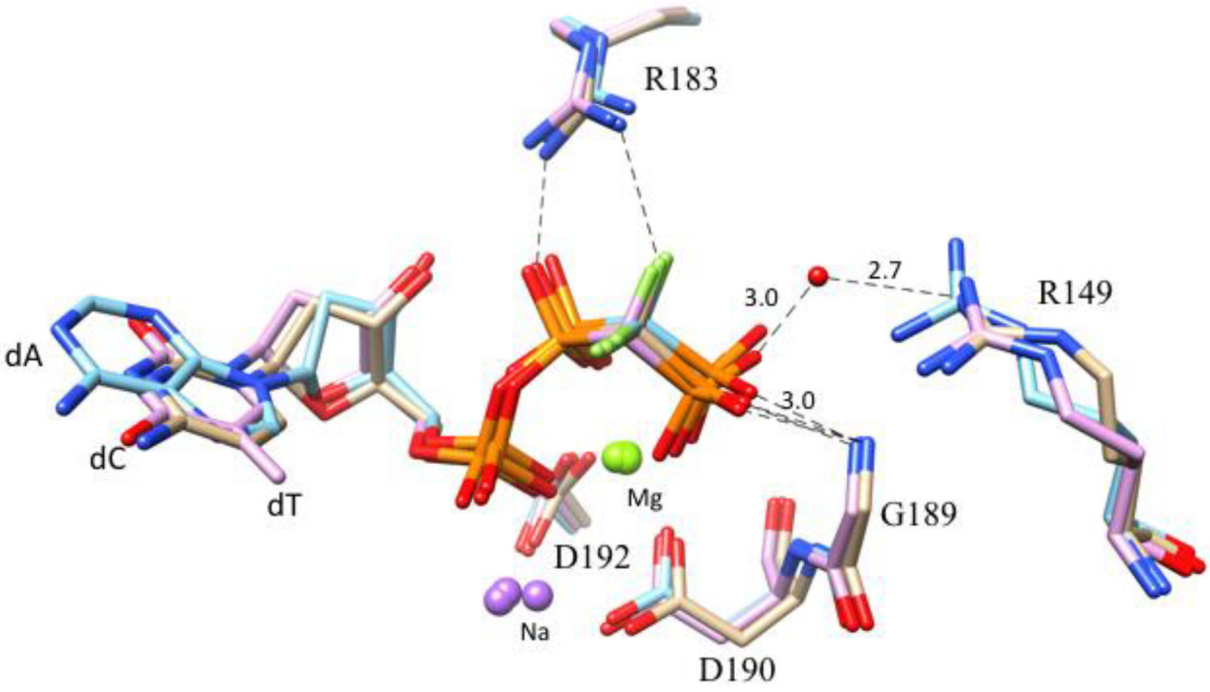
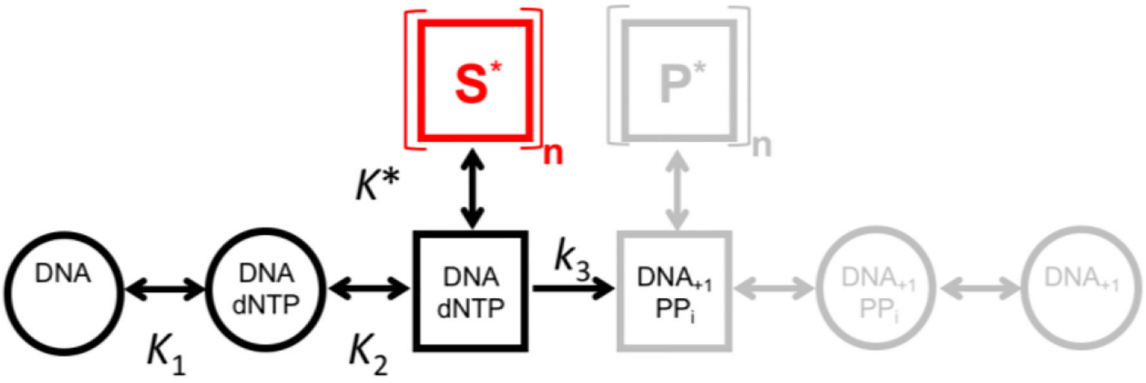


Figure 12: Overlay of active site structures of three CF_2 analogues with bases dA (cyan), dC (gold) and dT (pink).

The different positions of R149 are illustrated along with the distinct intervening water molecule and distances for the guanidine nitrogen and Pγ O3G contact in the C-CF_2 structure. The main-chain contacts with G189 and the contacts with R183 are similar in the three structures. The overlay was made with the 155–260 region of the enzyme.



Scheme 1:
General DNA Polymerase Reaction Pathway