The Dissertation Committee for Zhinan Jin Certifies that this is the approved version of the following dissertation:

The Molecular Basis of Nucleotide Recognition for T7 DNA Polymerase

Committee:

Kenneth A. Johnson, Supervisor

Dean R. Appling

Kevin Dalby

Rick Russell

Whitney Yin

The Molecular Basis of Nucleotide Recognition for T7 DNA Polymerase

by

Zhinan Jin, B.S.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

August, 2008

Dedication

To my wife, Min Wu

Acknowledgements

I would like to thank my advisor, Dr. Kenneth Johnson, for all his support throughout my graduate education. I stand on the shoulder of a giant, who makes me see further and reach higher. He has taught me much more than enzyme kinetics and always encouraged me to be a better scholar. I would also like to thank Dr. JoAnn Hunter Johnson for her constant support in the lab and in my life. Without her help and encouragement, I would not be the researcher and the person I am today.

I would like to thank my committee members for their precious time and valuable guide on my projects: Drs. Dean Appling, Kevin Dalby, Rick Russell, and Whitney Yin.

I am very proud to be a member of the Johnson Lab and would like to thank all the current and past members who make this lab a premier place for research. My special thanks to Dr. Yuchih Tsai, who taught me the most of the techniques I used throughout this dissertation. I enjoyed learning many valuable things from Dr. John Brandis through his fantastic presentations and our many conversations. I would like to thank other current members: Matthew Kellinger, Jennifer Prowell, and Patricia Estep, and past members: Vi Dougherty, Jeremiah Hanes, Harold Lee, Scott Auerbach, Louise Wang, Jeff Bartron, and Jarle Lillemoen.

Finally, I am thankful for the support of my family and my beloved wife, Min Wu. I am endlessly grateful for the sacrifices she has made for our family.

The Molecular Basis of Nucleotide Recognition for T7 DNA Polymerase

Publication No._____

Zhinan Jin, PhD

The University of Texas at Austin, 2008

Supervisor: Kenneth A. Johnson

DNA replication demands extraordinary specificity and efficiency of catalysis from a DNA polymerase. Previous studies on several DNA polymerases suggested that a rate-limiting conformational change preceding chemistry accounts for the high specificity following the induced fit mechanism. However, the identity of this rate-limiting conformational change and how it contributes to the fidelity is still under debate. An important study of T7 DNA polymerase performed by Tsai and Johnson using a conformationally sensitive fluorophore (CSF) characterized a conformational change directly and presented a new paradigm for nucleotide selectivity. This thesis describes work to further characterize the underlying molecular basis regulating the conformational change by a combination of site-directed mutagenesis, transient kinetics and crystallography.

One flexible segment (gly-ala-gly) within the fingers domain was mutated to (ala-ala-ala). The kinetic analysis on this mutant showed that the mutations decreased the forward rate of the conformational change reported by the fluorophore about 1200-fold but there was no significant change on the reverse rate. The data suggested that the movement of the fingers domain is not a rigid body motion but may be complex due to the movements of various helices

within the fingers domain. Quantification of the kinetics of incorporation of correct and incorrect base pairs showed the decrease of fidelity mainly was from the decreased forward rate during correct nucleotide incorporation.

The roles of three active site residues, K522, H506, and R518, which form polar interactions with α-, β- and γ-phosphates of the incoming nucleotide respectively, in conformational change and catalysis were also characterized. All the mutants showed a slower conformational change than the wild type enzyme. After this conformational change, there was a rate limiting step with a rate comparable to k_{pol} measured by quench-flow experiments. Correct nucleotide binding caused an increase in fluorescence, suggesting that the conformational change of the fingers domain delivers incoming nucleotide to a misaligned status even for a correct nucleotide with each of the mutants. The data suggested that active site residues play important roles in maintaining a fast conformational change and an accurate alignment of the active site during correct nucleotide incorporation.

Yellow crystals of CSF-labeled T7 DNA polymerase with DNA and correct nucleotide (closed complex), incorrect nucleotide (misaligned complex) or no nucleotide (open complex) were grown to good size and diffracted to 3 Å during X-ray data collection. The structures of these complexes are still under refinement.

Table of Contents

List of Tables

List of Figures

List of Illustrations

Chapter 1: Overview of DNA Polymerases and T7 DNA Polymerase

1.1 INTRODUCTION

The goal of enzymology is to understand how an enzyme achieves its specificity and efficiency in terms of the basic chemistry, kinetics, thermodynamics and the structures of the enzyme. All the information gathered from these studies can be combined and analyzed to explicitly define the mechanism of an enzyme (1). DNA polymerases stand out as an ideal model system to study enzyme specificity and efficiency, since they are among the enzymes with highest specificity and efficiency. The tremendous task of replication of genomic DNA demands polymerases with high fidelity and a fast rate of catalysis. As an example, T7 DNA polymerase can incorporate 300 nucleotides per second and only makes one error per 10^8-10^9 incorporated bases (2). Moreover, studies of DNA polymerases will improve our understanding of diseases such as viral infection and cancer and perhaps illuminate the basis of aging (3,4), and possibly help to develop medical treatments. For example, understanding the mechanisms governing the specificity of HIV reverse transcriptase and human DNA polymerase γ can guide the development of nucleoside analogs to treat AIDS with higher effectiveness and lower toxicity (5).

1.2 CLASSIFICATION OF DNA POLYMERASES AND STRUCTURAL COMPARISON

Based on amino acid sequence similarity, Ito et al. classified DNA polymerases into A, B, C and X families (6). In their classification, families A, B, and C were grouped based on the sequence homologies to *E. coli* DNA polymerases I, II, and III, respectively. Family A contains replicative polymerases such as T7 DNA polymerase and the mitochondrial DNA polymerase γ, and repair enzymes such as *E. coli* DNA polymerase I and *Thermus aquaticus* (Taq) DNA polymerase. Family B includes *E. coli* DNA polymerase II, human DNA polymerase α, T4 DNA polymerase and RB69 DNA polymerase. These DNA polymerases are mainly replicative enzymes with high fidelity. Family C includes major bacterial replicative DNA polymerases such

as *E. coli* DNA polymerase III α subunit and *Bacillus subtilis* DNA polymerase III. Eukaryotic DNA polymerase β is not homologous to the polymerases in the A, B, and C families and thus the enzymes from rat and human are grouped into the X family. Polymerase β mainly is involved in base excision in eukaryotic cells while DNA polymerase I is responsible for this function in prokaryotic cells. Several newly discovered DNA polymerases such as eukaryotic polymerases η, κ, ι and Rev1 along with prokaryotic *E. coli* Pol IV (DinB) and Pol V (UmuC), are assigned to a new family called the Y-family by Filee *et., al.* (7). The enzymes in this family are involved in bypassing stalled DNA replication from DNA damage and lesion using translesion synthesis (TLS). They all have lower fidelity for nucleotide insertion which is probably related to their ability to bypass errors at the lesion sites (8). Finally, the reverse transcriptase family (RT family) includes polymerases from retroviruses and eukaryotic telomerases. Among these, the HIV reverse transcriptase has been extensively studied because of its medical importance.

Since the first crystal structure of a polymerase, the Klenow fragment of *E. coli* DNA polymerase I, was reported in 1985 (9), crystal structures from almost all DNA polymerase families are now available (10-29). Although different DNA polymerases from different families fulfill their differential roles during replication, repair and error bypassing and retain their own unique properties, they share many common features including a two-metal-ion mechanism to catalyze the phosphoryl transfer reaction (30-33). The crystal structures of all DNA polymerases resemble a human right hand with “palm”, “fingers” and “thumb” domain (Figure 1.1) (10,14,34). The palm domain is responsible for the catalysis of the phosphoryl transfer reaction with three catalytically essential carboxylates. The structure of this domain is highly conserved among Pol I , RB69 and HIV RT but not Pol β. The fingers domain is proposed to be involved in binding of the templating base and the incoming nucleotide, so it can recognize the identity of the base pair and thereby govern the fidelity of the DNA polymerase. The structures of the fingers domain are different between enzymes from different families. However, a functionally conserved α-helix

that binds the incoming nucleotide is located at a similar position within the fingers domain of different DNA polymerases; in HIV RT, an anti-parallel β-ribbon with the same function sits at this position instead of an alpha-helix. The thumb domain mainly is involved in the binding and translocation of the duplex DNA thus playing an important role in replication processivity. The structures of the thumb domain are different among different families but they mainly contain parallel and anti-parallel α-helices and at least one helix makes important contacts with the minor groove of the DNA duplex.

The available structures of DNA polymerases also revealed a large conformational change of the enzyme following nucleotide binding (22,35). When the polymerases are bound with either DNA or nucleotide alone, the active site is more open than it is in the ternary complex of the enzyme with DNA and a correct nucleotide. The major change is associated with a large movement of the fingers domain while the changes of the palm and thumb domains are not noticeable. The open-to-closed conformational change supports the induced-fit mechanism of enzyme catalysis in that the enzyme initially opens to take substrate from solution and then closes the active site for the best alignment to allow fast chemistry to occur.

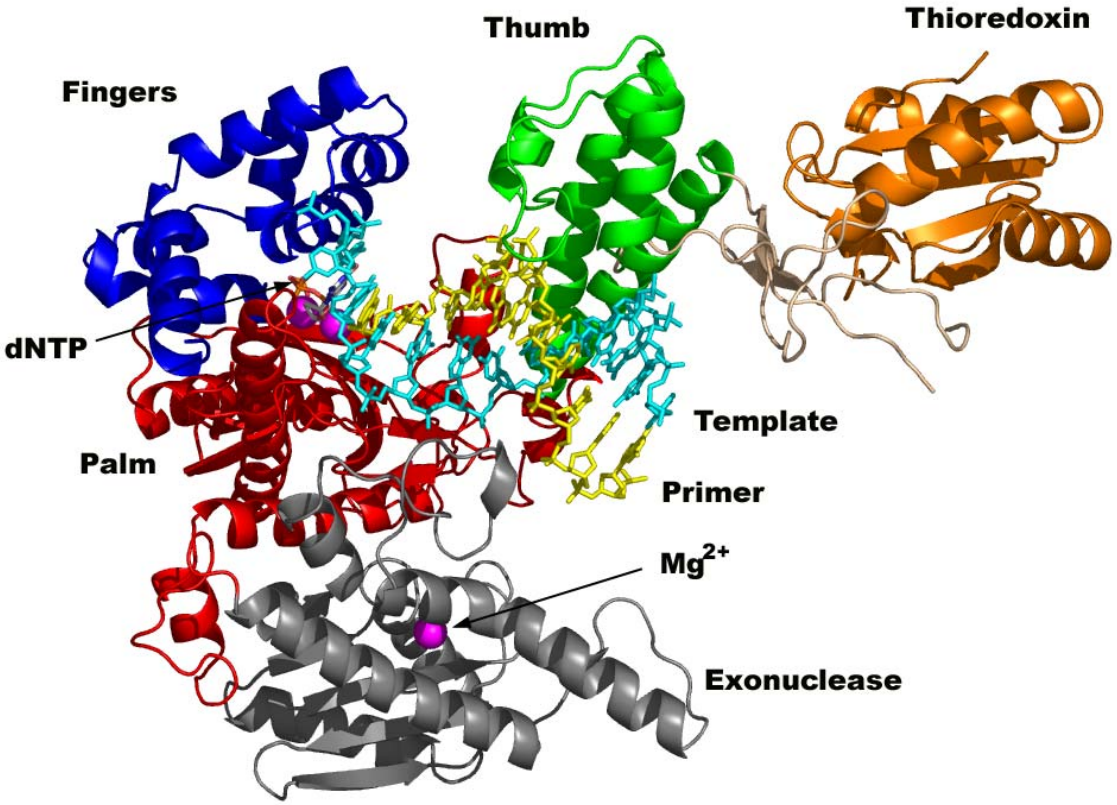


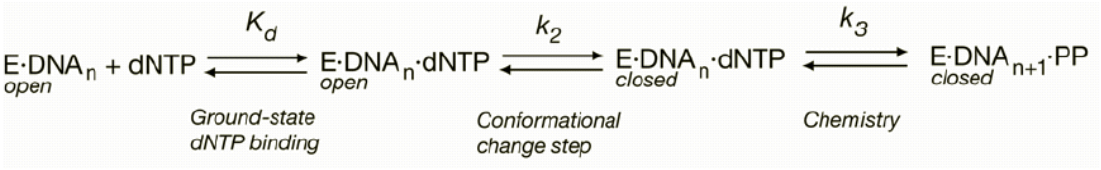
Figure 1.1: The structure of T7 DNA polymerase ternary complex.

The structure shows the "closed" complex resulting from the rotation of the fingers domain (blue) to bring the catalytic residues into contact with the incoming dNTP (CPK colors). The palm subdomain is shown in red with two metal ions in magenta showing the polymerase site. A single metal ion at the exonuclease site is also shown in magenta. The thumb subdomain is in green. A loop in wheat color binds one *E. coli* thioredoxin molecule (orange color). The DNA template strand is shown in cyan, and the primer strand in yellow. Drawn from PDB structure file 1T7P.

1.3 DNA POLYMERASE NUCLEOTIDE RECOGNITION MECHANISM

DNA polymerases catalyze a phosphoryl transfer reaction via a two-metal-ion mechanism. During this reaction, the 3'-OH group of the primer terminus is activated by one metal ion and attacks the α-phosphate of the incoming nucleotide, resulting in the transfer of the dNMP to the primer terminus and the release of pyrophosphate. Since the chemistry of incorporating a correct or an incorrect nucleotide is the same, DNA polymerases must have evolved some mechanism to account for the high fidelity of DNA replication other than just chemistry.

Transient kinetic methods revolutionized the field of DNA polymerases in that more complete and realistic kinetic models can be derived (36-38). Based on previous kinetic analyses of T7 DNA polymerase (2,39) and related studies on the Klenow fragment of *E. coli* Pol I (40,41), the Johnson lab proposed a model for nucleotide recognition through a three step mechanism (37). Nucleotide binds to the open enzyme state; a matched nucleotide induces a conformational change which closes all the catalytic amino acids around the substrate and optimally aligns the 3'-OH group of the primer with the α-phosphate of the incoming nucleotide; then the fast phosphoryl transfer reaction proceeds (Scheme 1). This model has been confirmed to varying degrees in kinetic and structural studies on Klenow, Taq DNA polymerase, Pol β and HIV RT. Thus further kinetic and structural studies on T7 DNA polymerase can be a guide to other related polymerases.



Scheme 1.1: Pathway of nucleotide incorporation

In this model, nucleotide selectivity is controlled to various degrees by the microscopic rate constants of each step. Measurement of the rate of incorporation at various nucleotide concentrations in a single turnover experiment can define two important kinetic parameters, the apparent K_d for ground-state binding of nucleotide and the maximum rate of catalysis, k_{pol}. Table 1.1 lists the average fidelity contributions of these two parameters for several polymerases(39,42-45). The contribution from ground-state binding ($K_{d,incorrect}/K_{d,correct}$) only ranges from a factor of 170 to 300. Thus the fidelity contributed by most polymerases at ground-state binding is similar. However, when we inspect the fidelity contribution of the rate of incorporation ($k_{pol,correct}/k_{pol,incorrect}$), it ranges from a factor of 50 in HIV RT to 5000 for T7 DNA polymerase. This 100-fold difference between a low and high fidelity polymerase implies that the steps defining the rate of incorporation contribute the most to the overall accuracy for a polymerase. According to this model, the maximum rate of incorporation, k_{pol}, can be a function of a rate-limiting conformational change step (k_2), a rate-limiting chemistry step (k_3), a combination of both steps, or possibly additional steps in the pathway. Thus, identifying the rate-limiting step and establishing the structural elements controlling the rate of reaction will explain the origin of the nucleotide selectivity.

Table 1.1: Fidelity contributions of ground-state binding and polymerization

Polymerase	$K_{d,incorrect}/K_{d,correct}$	$k_{pol,correct}/k_{pol,incorrect}$	Overall Fidelity
T7 DNA polymerase	300	5,000	1,500,000
Klenow	170	2,300	391,000
Pol γ DNA polymerase	200	1,500	300,000
Pol β	290	610	176,900
HIV reverse transcriptase	260	50	13,000

The existence of the closed conformational intermediate was supported by pulse chase experiments observing the transient appearance of a tightly bound dNTP state preceding chemistry (2,41,42,46) . A rate-limiting conformational change preceding a fast chemistry step was also suggested by the observation of a small thio-elemental effect during incorporation of a correct nucleotide (39,41). The thio-elemental effect was defined by the ratio of the rate of incorporating a normal dNTP to the rate of incorporating a corresponding α-thio-substituted dNTP (with a non-bridging *Sp*-oxygen of the α-phosphate replaced by a sulfur group) ($k_{pol,dNTP}/k_{pol,dNTP\text{-}\alpha\text{-}S}$). Since the sulfur group decreases the chemistry rate due to its decreased electronegativity compared to oxygen, a large elemental effect is expected when chemistry is rate-limiting while a small thio-elemental effect may suggest that steps other than chemistry may be rate-limiting. However, rigorous analysis of the magnitude of the thio-elemental effect suggested a smaller intrinsic elemental effect during phosphoryl transfer reaction, and thus, observing a small elemental effect does not necessarily mean steps other than chemistry are rate-limiting (47,48). Therefore more direct methods were required to identify and characterize the conformational change step preceding chemistry.

Studies on Pol β using a 2-aminopurine (2-AP) in the template strand near the templating base revealed that a fluorescence change, corresponding to a conformational change preceding chemistry, was not rate-limiting (49-51). Subsequent studies on DNA polymerase I using 2-AP also implied a rapid conformational change occurring before chemistry (52).Using a FRET pair,

an acceptor fluorophore on the fingers domain and a donor fluorophore on the DNA, the study on the Klenow fragment of the Taq DNA polymerase also suggested that the large closure of the fingers domain is much faster than chemistry (53). These studies suggested that the kinetically defined conformational change reported by 2-AP fluorescence change or FRET signal is not rate-limiting and it might be corresponding to the structurally defined open-to-closed large conformational change. Furthermore, the contribution of the conformational changes to the nucleotide selectivity in terms of specificity was questioned, and Showalter and Tsai proposed a simplified model to attribute specificity solely to the rate-limiting chemistry for both correct and incorrect nucleotide incorporation based on their study on Pol β (47). However, the 2-AP signal is more likely sensing the translocation of the DNA during nucleotide incorporation, so the ambiguity in interpreting 2-AP data makes the identity of the rate-limiting conformational change still under debate. Some other data suggested that the binding of the second metal ion (54), rearrangement of the fingers domain (53), and rotation of one or more active site residues before chemistry (55) might be rate-limiting. Our recent studies of T7 DNA polymerase with a fluorescence tag on the fingers domain gave more insights of the specificity contribution of the conformational change as I will review in the following section.

1.4 T7 DNA POLYMERASE: FUNCTION, STRUCTURE AND KINETICS

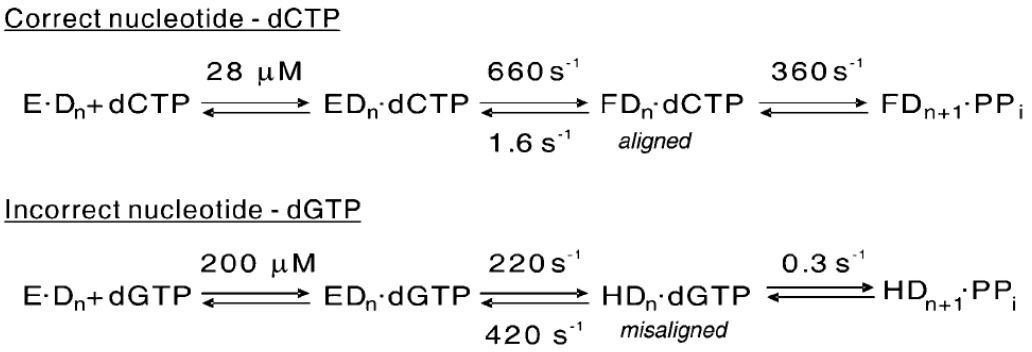
With the previously well-defined minimal pathway for nucleotide recognition (2,39) and its high resolution crystal structures (14,56-59), T7 DNA polymerase stands out as an ideal model to study the high fidelity of DNA replication. T7 DNA polymerase replicates the bacteriophage T7 genome. It is a single polypeptide encoded by the gene 5 of the phage, with a molecular weight of 80 KDa (60). During genome replication, T7 DNA polymerase forms a complex with *E. coli* thioredoxin (12 KDa) in one-to-one stoichiometry (61-63). The binding of these two proteins is very tight with a K_d of 5 nM due to fast association and slow dissociation (64,65). Thioredoxin does not have any catalytic activity for DNA replication but it improves DNA

binding and processivity of T7 DNA polymerase (66,67). Similar to the Klenow fragment of DNA polymerase I, T7 DNA polymerase has polymerase activity and 3'-5' exonuclease activity. With thioredoxin, the polymerase active site incorporates 300 nucleotides per second, and only makes one error out of one million bases. When the polymerase does make an error, it stalls and transfers the primer strand to its exonuclease active site where the misincorporated nucleotide can be removed. This exonuclease activity accounts for an extra 200-fold in fidelity, making an overall fidelity of one error in 10^8-10^9 bases(39,68). Such high fidelity of DNA replication implies that this DNA polymerase provides a good model to understand the mechanisms governing nucleotide selectivity.

Figure 1.1 shows an overview of the crystal structure of the ternary complex of T7 DNA polymerase with DNA and an incoming nucleotide (E·D·dNTP). This structure provides essential information for us to directly probe the roles of the individual structural elements of the enzyme by a combination of site-directed mutagenesis and rigorous kinetic analyses. In the structure, T7 DNA polymerase resembles a human right hand with fingers, palm and thumb domains, which is a common feature of most polymerases. The 5' end of the DNA template strand does not run through the cleft formed between the fingers and the thumb, but bends at 90 degrees away from the active site. In response to dNTP binding, there is about 45 degree rotation of the fingers sub-domain, which has also been called the recognition domain, shown in blue in Figure 1.1. This brings all the catalytic residues in the recognition domain into close contact with the incoming nucleotide leading to fast catalysis. This rotation of the recognition domain was revealed by a comparison of this structure (E·D·dNTP, a closed complex) with a structure of T7 DNA polymerase alone (open complex, unpublished structure kindly provided by Dr. Tom Ellenberger). It has been suggested that the rate-limiting conformational change preceding chemistry from the previous studies of T7 DNA polymerase may correspond to this rotation of the recognition domain.

By labeling the fingers domain of T7 DNA polymerase with a conformationally-sensitive fluorophore, 7-diethylamino-3-((((2- maleimidyl)ethyl)amino)carbonyl)-coumarin (MDCC), Tsai and Johnson measured the kinetics and equilibrium of nucleotide binding, nucleotide-induced conformational change and chemistry for correct and incorrect nucleotides, which are summarized in Scheme 1.2 (69). First of all, the fluorophore somehow showed different fluorescence states when the enzyme-DNA duplex (the primer of the DNA was dideoxy terminated to prevent chemistry) was bound with correct and incorrect nucleotides, suggesting different structures were formed for correct versus incorrect nucleotides. This observation implied that the conformational change reported by MDCC allows the enzyme to recognize nucleotides and delivers the correct nucleotide to a good alignment for fast catalysis but delivers the mismatched nucleotides to a misaligned state. Therefore, the fluorescence data defined three distinct structures for the enzyme when bound with different substrates: open, closed and misaligned. Secondly, the conformational change is fast and partially rate-limiting in the forward reaction upon binding a correct nucleotide and its slow reverse rate traps the correct nucleotide in an aligned enzyme state leading to fast catalysis. However, the binding of an incorrect nucleotide is unfavorable due to a fast reverse rate, and the chemistry is very slow, presumably due to the misalignment of the catalytic residues. When the specificities in terms of k_{cat}/K_m for correct and incorrect nucleotides were examined, it turned out that whenever the reverse rate of the conformational change is much slower than the chemistry step, the rate of chemistry is cancelled from the expression for $k_{cat}/K_{m,}$ and therefore the specificity for correct nucleotide is solely dictated by the ground-state binding K_d and the forward rate of the conformational change. In contrast, the specificity for incorrect nucleotide incorporation is still defined by the nucleotide binding and the slow chemistry step. This study clearly described how the conformational change regulates nucleotide selectivity. For a long time, the argument that only the rate-limiting step along the reaction pathway dictates the specificity misled many researchers in the wrong direction

to understand the fidelity of DNA polymerases by arguing whether the conformational change or the chemistry is the rate-limiting step. The application of these new findings in enzymology in general was summarized in a recent review by Johnson (70).



Scheme 1.2: Pathways of correct and incorrect nucleotide incorporation.

1.5 PROJECT SUMMARY

The new study of T7 DNA polymerase in our lab described how the conformational change step determines the nucleotide selectivity; however, questions remain to be answered about this conformational change.

In the previous study done by Tsai and Johnson, for labeling MDCC at the desired position, residue 514, on the T7 DNA polymerase, a cysteine replaced E514 and eight cysteines on the surface were mutated to alanine or serine. However, two wild-type cysteine residues (C483, C622) remained in addition to the engineered 514C. Therefore, it is critical to examine whether only 514C was labeled with MDCC in order to more accurately interpret the fluorescence data. In chapter 2, we showed that the 514C was the only site labeled with MDCC.

We tried to understand the identity of the conformational change reported by the fluorescence from MDCC. The concept of "open-to-closed" movement of the fingers domain might give us an impression that the closure of the fingers domain is a rigid body movement of the whole fingers domain around a hinge. However, a careful inspection of the open and closed

crystal structures of T7 DNA polymerase shows not only the "open-to-closed" large change of the fingers domain, but also small changes in the packing states of the helices within the fingers domain. For example, the gap from E514 to the O2 helix at the back side of the fingers domain is larger in the closed state than in the open state. Thus the conformational change of the fingers domain is mechanistically very complex. Many questions remain. What is the relationship between the large motion of the fingers domain and the small changes in the helix packing? Do they occur simultaneously or sequentially? Based on the labeling position of MDCC on the fingers domain and its environmental sensitivity, MDCC may sense the small changes in structure within the fingers domain. To better understand the conformational change reported by MDCC, the relative motions of the helices within the fingers domain were examined by site-directed mutagenesis and transient kinetics and the results are shown in chapter 3.

It should be noted that in the study done by Tsai and Johnson, the MDCC tag is on Cys^{514}, which is just three residues away from the catalytically important Arg^{518}; thus, changes in fluorescence may reflect the positioning of this residue, and other catalytic residues nearby(69). This signal will provide a means to quantify the effects of mutations on these catalytic residues on the conformational change and catalysis, thus supplying more mechanistic information regarding the roles of these residues during nucleotide binding and incorporation than before. Chapter 4 describes the characterization of the mutants of three catalytic residues on the recognition domain, R518, H506, and K522, and their roles during nucleotide binding and incorporation are discussed.

As reported by the MDCC fluorescence, T7 DNA polymerase exists in three distinct conformational states during nucleotide binding: open, closed and misaligned. However, the precise structures of these three states are unknown, and how MDCC senses the different states is still a mystery. Crystal structures of these complexes may yield insights into the mechanisms. In chapter 5, I described the crystallization of the complexes of MDCC labeled T7 DNA polymerase

with DNA and correct nucleotide (closed), incorrect nucleotide (misaligned) and no nucleotide (open).

In summary, these studies will provide definitive evidence for the origin of the high fidelity observed in DNA replication. Further studies on T7 DNA polymerase will continue to provide fundamental insights into polymerase mechanisms applicable to all DNA polymerases.

Chapter 2: Site-specific Labeling of T7 DNA Polymerase with a Conformationally-sensitive Fluorophore

2.1 INTRODUCTION

Previous kinetic analysis of several DNA polymerases implied the existence of a rate-limiting conformational change between nucleotide binding and the phosphoryl transfer reaction (2,41,71). The structures of DNA polymerases also revealed a large open-to-closed conformational change induced by binding of a correct nucleotide (35). According to the induced-fit mechanism, this open-to-closed transition may contribute to the large specificity observed in DNA polymerases. Since specificity is a kinetic term, direct measurement of the kinetics of this conformational change during nucleotide incorporation was needed.

An environmentally sensitive fluorophore, 7-diethylamino-3-((((2-maleimidyl) ethyl)amino)carbonyl)coumarin (MDCC), has been successfully used to characterize the conformational changes of a phosphate binding protein from *E. coli* (72,73). In our approach to understand the conformational changes of the fingers domain of T7 DNA polymerase, we sought to attach this fluorophore specifically on the fingers domain at residue 514. MDCC only can react with a thiol group, therefore the E514C mutation was introduced in this enzyme. Since T7 DNA polymerase has 10 native cysteines which also may react with MDCC, a mutagenesis and specific labeling strategy was developed. In this strategy, solvent accessible cysteines were mutated to serine or alanine by site-directed mutagenesis, while two buried cysteines were retained. Our goal was to keep the DNA replication activity of the mutated protein, while labeling with MDCC only at residue 514C. In this study, tryptic digestion, reverse phase HPLC and mass spectrometry were utilized to determine the sites of labeling with MDCC on the protein. After several rounds of mutagenesis, specific labeling by MDCC at residue 514C was achieved for an 8-cys-light mutant

of T7 DNA polymerase and the activity of this mutant was comparable to the wild-type enzyme (69).

2.2 MATERIAL AND METHODS

The MDCC-labeled mutant enzymes used in this chapter were prepared by Yu-Chih Tsai following the protocol described below. I then performed the tryptic digestion, HPLC and mass spectrometry analyses to establish the sites of MDCC labeling in these mutants.

Expression, purification, and MDCC-labeling of the T7 DNA polymerase mutants

The plasmids encoding the cys-light mutants of the T7 gp5 were transformed into the competent *E. coli* strain BL21-Gold(DE3)pLysS (Stratagene). Colonies carrying the expression vector were selected on LB agar plates with 100 μg/ml of ampicillin at 37°C. A single bacterial colony was inoculated into 10 ml of NZY media (10 g of NZ amine, 5 g of NaCl, 5 g of bacto-yeast extract, and 2 g of $MgSO_4$-$7H_2O$ dissolved in 1 L of ddH_2O) containing 50 μg/ml of ampicillin and 12.5 μg/ml of chloramphenicol. The culture was incubated overnight with shaking at 37°C. Afterward, the culture was transferred into 1 L of the same media plus 50 μg/ml of ampicillin and grown until OD_{600} reached 0.3~0.4. The expression of the recombinant T7 DNA polymerase was induced by adding 0.4 mM IPTG and 12.5 μg/ml of chloramphenicol. The culture was grown for 3 more hours at 37°C, and cells were harvested by centrifugation at 5,000 x g for 15 minutes.

Typically, cell pellets prepared from 6 L of culture were washed briefly with rinse buffer (50 mM Tris-HCl, pH 8, 25 mM EDTA, 150 mM NaCl, 1 mM β-mercaptoethanol, and 0.1 mM DTT). After a freeze-thaw cycle, pellets were resuspended with lysis buffer (same as rinse buffer except with 2.5 mM of EDTA, 10 mM of phenylmethylsulfonyl fluoride, and 0.3 mg/ml of lysozyme). The cell suspension was stirred for 15 min at room temperature followed by sonication on ice for 30 s. For the remaining steps, all solutions were kept on ice or at 4°C refrigerator. Following sonication, sodium deoxycholate was added to a final concentration of

0.1%, and Polyethyleneimine (PEI), pH 8, was added slowly to a final concentration of 0.5% (w/v) with stirring. The solution was stirred for an additional 15 min after the addition of PEI. Next, NaCl was added to a final concentration of 0.5 M and stirred for 30 min. The cell lysate was cleared by centrifugation at 16,000 x g for 20 min and then fractionated by 35-70% ammonium sulfate precipitation. The pellets were resuspended and diluted with buffer without salt (50 mM Tris-HCl, pH 7.5, 0.1 mM EDTA, 10% glycerol, and 5 mM DTT) until the conductivity of the solution was equal to that of the low salt DEAE running buffer (50 mM Tris-HCl, pH 7.5, 0.1 mM EDTA, 100 mM NaCl, 10% glycerol, and 5 mM DTT). The diluted protein extract was loaded onto a 200 ml DEAE-sepharose column at a 2 ml/min flow rate using the Pharmacia fast protein liquid chromatography (FPLC) system. The column was washed with the low salt buffer until the UV absorbance returned to the base line. Bound proteins were eluted with a 1.6 L 100-400 mM NaCl gradient. Fractions containing T7 DNA polymerase were identified with SDS-PAGE, pooled, and diluted with buffer containing 40 mM Tris-HCl, pH 7.5, 0.1 mM EDTA, 10% glycerol, and 1 mM DTT. Diluted protein was further purified by a 12 ml single-stranded DNA cellulose (ssDNA-cellulose) column with a 100 to 350 mM NaCl gradient in buffer containing 40 mM Tris-HCl, pH 7.5, 0.1 mM EDTA, 10% glycerol, and 1 mM DTT. Fractions containing the T7 DNA polymerase were pooled and concentrated with an ultrafiltration cell using YM-10 membrane (Amicon).

For MDCC Labeling of the T7 DNA Polymerase, a stock solution (10 mM) of MDCC (Molecular Probes) was prepared by dissolving solid MDCC with dimethyl sulfoxide (DMSO) and stored at -80 °C prior to use. Protein labeling was carried out in the labeling buffer [40 mM N-2-hydroxyethylpiperazine-*N*'-2-ethanesulfonic acid (HEPES) at pH 8, 0.1 mM ethylenediaminetetraacetic acid (EDTA), 10% glycerol, 50 mM NaCl, and 1 mM Tris(2-carboxyethyl)-phosphine (TCEP) hydrochloride]. The labeling reaction was initiated by adding 20-fold excess of MDCC over protein at 4 °C and incubated with constant mixing overnight. The

reaction was quenched by adding an excess amount of dithiothreitol (DTT) solution. Labeled mutant protein was purified through a ssDNA-cellulose column to remove excess MDCC. Eluted T7 DNA polymerase was dialyzed against the T7 storage buffer (40 mM Tris-HCl, pH 7.5, 0.1 mM EDTA, 50 mM NaCl, 50% Glycerol and 1 mM DTT) and stored at -80 °C.

Tryptic digestion of T7 DNA polymerase cys-light mutants

The MDCC labeled proteins, stored in the final dialysis buffer, was mixed with trypsin (sequencing grade modified trypsin, Promega) at a ratio of 20:1 (w/w), and the mixture was then incubated overnight at 37ºC. The resulting digested peptides were separated by reverse phase HPLC to resolve fluorescently labeled peptides.

Reverse phase HPLC

To separate the peptides, we used POROS R2 perfusion column (PerSeptive Biosystems). The column was equilibrated with Buffer A (0.1% TFA (trifluoro acetic acid), 2% acetonitrile, and ddH_2O) using an ÄKTA HPLC instrument (Amersham Pharmacia Biotech). Then 100 □l of the sample was loaded onto the column through a sample loop. The column was washed with buffer A and then the peptides were eluted with a gradient of acetonitrile, buffer A to buffer B (0.08% TFA, 80% acetonitrile, and ddH_2O), at a flow rate of 1.2 ml/min. UV absorbance at 220 nm was measured to detect the peptides peaks. To detect MDCC, absorbance was also measured at 425 nm. The eluant was collected in 1.2 ml fractions. The fractions corresponding to the absorbance peaks at 425 nm were picked for mass spectrometry analysis. To concentrate these samples, the samples were frozen in liquid nitrogen and a Savant SpeedVac Concentrator (Forma Scientific) was used to dry the samples.

Mass spectrometry and tandem mass spectrometry (MS/MS) analysis

The freeze-dried samples were dissolved in 10 µl of a solution of 50% Acetonitrile, 50% H_2O, and 1% TFA. The matrix solution was made of α-cyano-4-hydroxycinnamic acid saturated

in a solution of 70% Acetonitrile, 30% H_2O, 0.1% TFA, and 5 mM $(NH_4)_2HPO_4$. The dissolved samples were mixed with the matrix solution at 1:1 ratio (v/v), and then 0.5 µl of the mixture was spotted onto a MALDI stainless steel target. The mass spectra were obtained by an ABI 4700 Proteomics Analyzer MALDI-TOF/TOF instrument (Applied Biosystem). To verify the identities of the ions in the mass spectra, the ABI 4700 uses high energy collision induced dissociation (CID) to fragment selected ions to produce MS/MS spectra allowing assignment of peptide sequences.

2.3 RESULTS

Our strategy effectively led us to obtain an 8-cys-light 514C mutant of which MDCC was specifically labeled at 514C. The HPLC profiles of peptide fragments of several mutant forms of T7 DNA polymerase after tryptic digestion are shown in Figure 2.1. Four cysteines exposed on the surface of the protein were first mutated in the 4-cys-light E514C mutant. The HPLC profile of this mutant protein showed five absorbance peaks at 425 nm. Mass spectra analysis of these peaks identified that MDCC was labeled at C451, C313, and 514C. Then three more cysteines including C451 were mutated to generate the 7-cys-light E514C mutant. In its HPLC profile of tryptic digestion, the absorbance peak associated with C451 disappeared, and MDCC was found labeled only at C313 and 514C. Finally, an 8-cys-light E514C mutant was generated by mutating the cysteine at 313 to a serine. In this mutant, unique labeling of MDCC at 514C was observed and the detailed evidence is shown in the following paragraphs.

The 8-cys-light mutant contains three cysteines (C483, 514C, C622) which could potentially be modified by MDCC due to the reaction of a thiol with the maleimide. To test this hypothesis, we modified the enzyme with MDCC, and applied tryptic digestion, reverse phase HPLC and mass spectrometry to identify which cysteines were modified. Our data show that only 514C on the surface of the enzyme was modified. The other two cysteines appeared to be buried in the enzyme and less accessible to MDCC under our labeling condition.

The MDCC modified 8-cys-light mutant was digested by trypsin under the condition of complete digestion. The digested sample was loaded onto a reverse phase column. To detect peptides, absorbance at 220 nm was measured. The number of peaks at 220 nm was less than expected, suggesting that each peak may contain a mixture of several different peptides. For absorbance at 425 nm, there were only two peaks at fraction 12 and 13 (Fig. 2.2), which should contain MDCC. These two samples were concentrated by freeze-drying and were examined by mass spectral analysis.

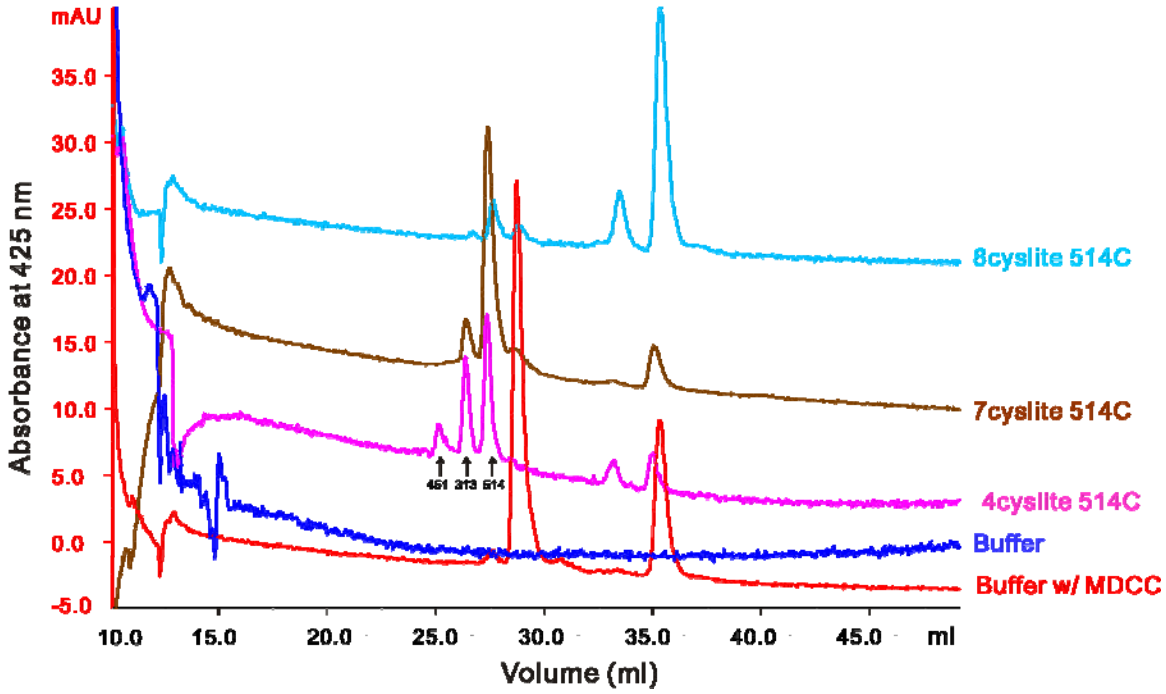


Figure 2.1: HPLC profiles of T7 DNA polymerase cys-light mutants after tryptic digestion

T7 DNA polymerase 4-cys-light (magenta), 7-cys-light (brown), and 8-cys-light (light blue) mutants were digested by trypsin and then the samples were analyzed with HPLC by monitoring light absorbance at 425 nm. As controls, T7 DNA polymerase final dialysis buffer and the buffer with free MDCC were also analyzed. Three black arrows point out the peaks which contain MDCC-labeled peptides; mass spectrometry analysis identified sites of labeling as C451, C313 and 514C.

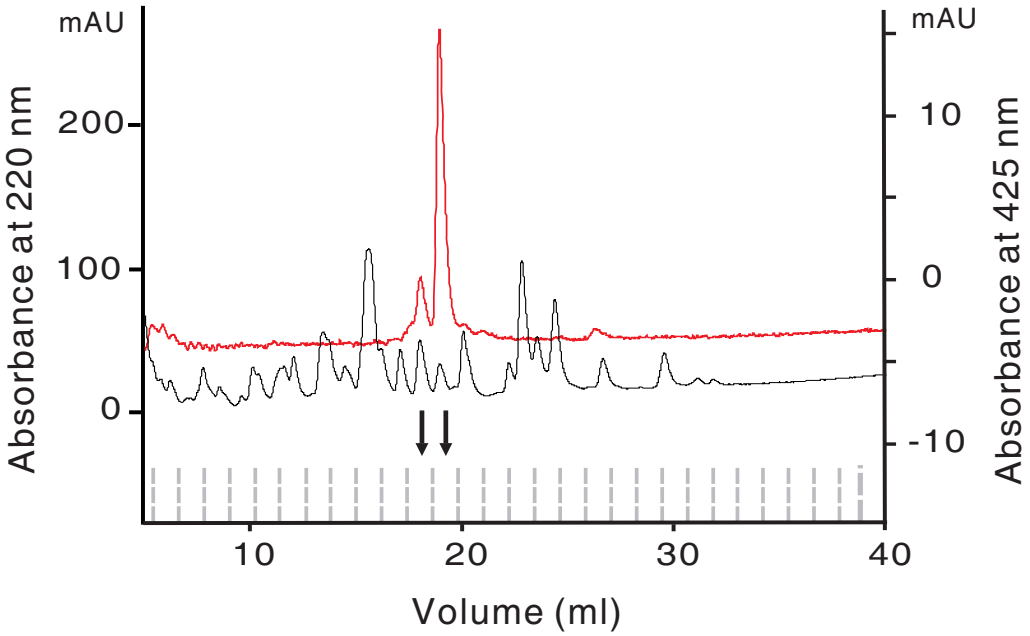


Figure 2.2: HPLC profile of the 8-cys-light mutant after tryptic digestion

In this chromatogram, absorbance was measured at 220 nm (black trace) and 425 nm (red trace). The broken line indicates the concentration of buffer B gradient from 0 to 100%. Fraction 12 and fraction 13 (indicated by the arrows) were selected for MS analysis.

To identify which peptides were modified by MDCC, we searched the MS spectrum of fraction 12 (data not shown) and fraction 13 (Fig. 2.3) for possibly modified peptides (Table 2.1). In both spectra, only one ion (m/z 1469.8) was identified to match the mass of the MDCC modified peptide, ^{509}NQIAACLPTR^{518}, which contains a 514C. Since no ions were identified to match the masses of the other two peptides containing cysteine with MDCC modification, our data strongly suggested that C483 and C622 were not modified by MDCC under our labeling condition.

Table 2.1: Tryptic digested polypeptide fragments containing cysteines

peptide fragment	cysteine position	calculated mass (Da)[a]	calculated mass with MDCC (Da)
^{483}CLAHFMAR^{490}	483	948.5	1331.6
^{509}NQIAACLPTR^{518}	514	1086.6	1469.7
^{605}SPHAALNTLLQSAG ALICK^{623}	622	1908.0	2291.2

[a] These values were calculated using the monoisotopic MH^+ masses; the monoisotopic mass of MDCC is 383.15 Da.

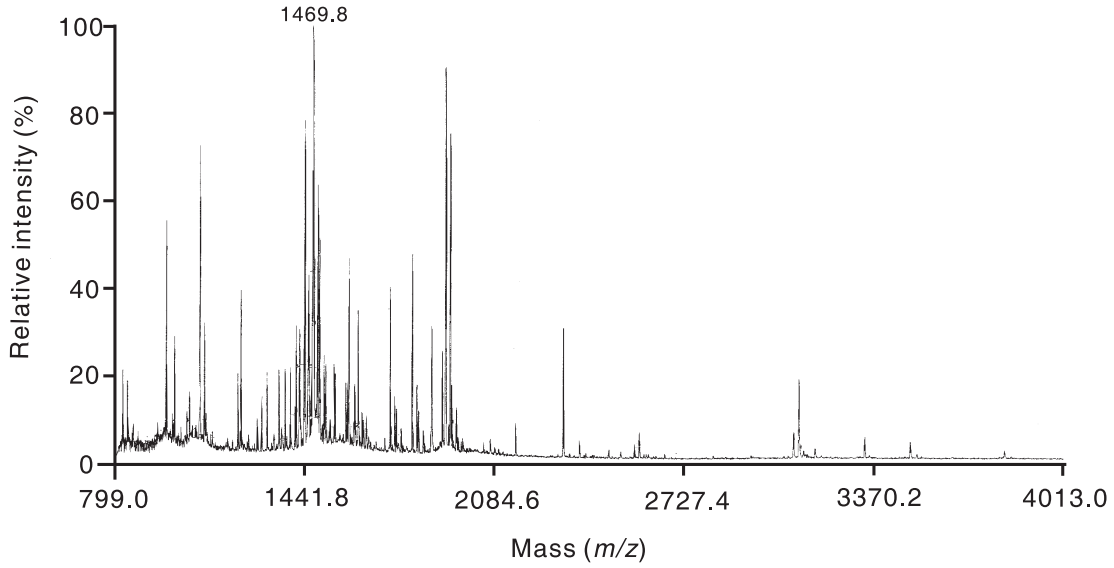


Figure 2.3: Mass spectrometry analysis of the sample in fraction 13

The mass of one ion with m/z of 1469.8 matches the calculated mass of the peptide ^{509}NQIAACLPTR^{518} with one molecule of MDCC (MH^+ of 1469.7).

To verify the identity of the ion with *m/z* of 1469.8, this precursor ion was subjected to MS/MS analysis. Since the monoisotopic mass of MDCC is 383.15 Da, the y ions from y5 to y9 and the b ions from b6 to b9 of ^{509}NQIAACLPTR^{518} should be shifted by a mass of +383.15 Da when the peptide is modified by MDCC at 514C. With this correction of the masses, all the y ions and most of the b ions of the MDCC modified ^{509}NQIAACLPTR^{518} could be assigned to the peaks in the MS/MS spectrum (Fig. 2.4). The b9 ion could not be assigned in the spectrum at first, however, an ion with *m/z* of 1052.6 was proposed to be an adduct of b9 ion with a fragmented MDCC which lose a part with mass of 244.1 Da. The existence of an ion with *m/z* of 244.1 supported this hypothesis, thus finally we could assign the peak of 1052.6 to the b9 ion with a shifted mass of +140.0 Da. This MS/MS spectral analysis unambiguously confirmed the identity of the MDCC modified ^{509}NQIAACLPTR^{518} and showed that the modification is at 514C. In conclusion, our results indicate that the MDCC is specifically labeled at 514C under our labeling condition.

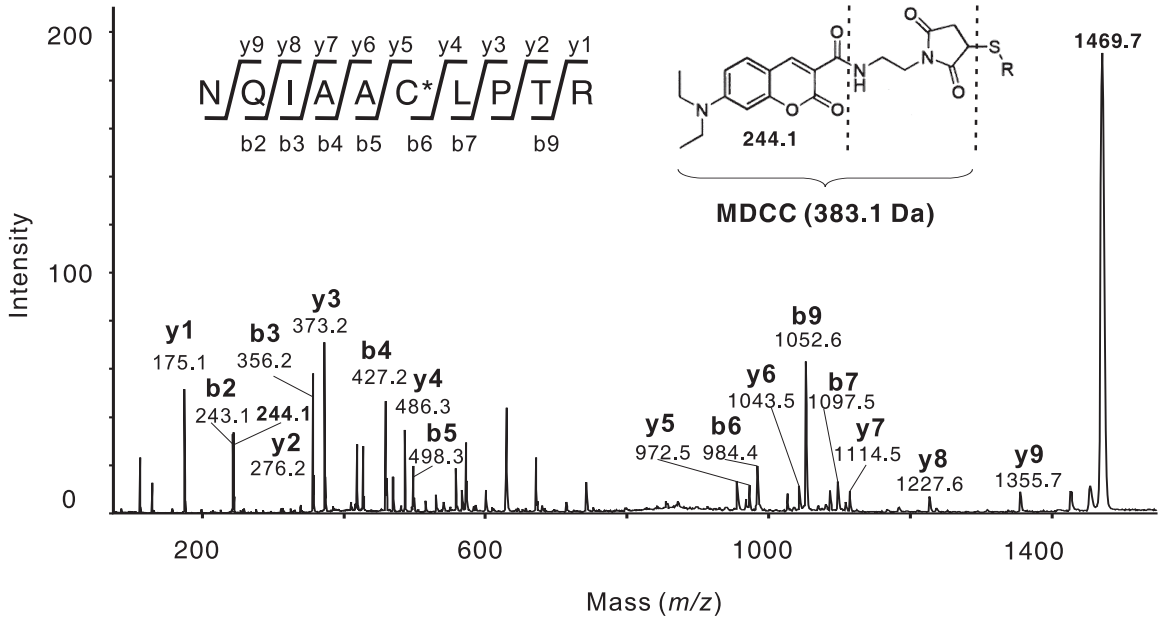


Figure 2.4: MS/MS spectrum of the precursor ion with m/z of 1469.8

The MS/MS spectrum of the precursor ion with m/z of 1469.8 is shown at the bottom and the assignment of the peaks to b ions and y ions calculated from peptide ^{509}NQIAACLPTR^{518}, which contains MDCC modification at Cys-514, is shown at top-left. At top-right is the structure of MDCC, and the hypothesized fragmentation site of MDCC is indicated by a broken line and the calculated mass of the ion generated from the left portion of MDCC is 244.1 Da.

2.4 DISCUSSION

Specific labeling of MDCC at 514C of T7 DNA polymerase was successfully fulfilled by our strategy presented here. This strategy can also be applied for labeling other enzymes with MDCC, or with other modifications on proteins. For example, to study conformational changes of HIV reverse transcriptase during nucleotide binding and catalysis, specific labeling of HIV RT with MDCC following this strategy is underway in our lab.

Our methods can be utilized as a guide to make mutations on a protein to remove unwanted labeling of MDCC. As shown in Figure 2.1, removal of the initial four cysteines was not enough as C451 was inevitably labeled by MDCC as identified by HPLC-MS analysis. Mutating this cysteine to a serine led to disappearance of one absorbance peak at 425 nm which has the peptide with MDCC labeling at C451. These results demonstrated the usefulness of our methods in that we can identify peptides containing cysteine with MDCC modification. By process of elimination, we can then establish the minimal mutations to remove native cysteine residues that would otherwise interfere with site specific labeling, while retaining enzyme activity.

The drawback of our methods might be that sometimes we can not tell whether a cysteine is labeled by MDCC due to the limitation of mass spectrometry. The better ionization ability of a molecule, the better signal it will obtain in mass spectrometry. Therefore it is possible that some peptides with MDCC label can not be identified in our methods if the peptides can not be efficiently ionized. However, this might not happen to our protein since we detected the non-specific labeling of MDCC in the 4-cys-light and the 7-cys-light mutants, while only labeling at 514C was identified after complete search for labeling positions of MDCC for the 8-cys-light mutant. These data strongly suggest that unique labeling of MDCC at 514C was achieved.

Chapter 3: Characterization of T7 DNA Polymerase Hinge Mutants

3.1 INTRODUCTION

Enzymes are dynamic. There are several different modes of motion and flexibility observed in enzymes such as molecular tumbling, “breathing”, rotation of amino acid side chains, shear and hinge motions of domains, flexible motions of loops, and thermally driven vibrations of atoms (1). However, how the dynamics of enzymes at different levels and on different time scales are correlated with catalysis is largely unknown. In the induced-fit theory of enzyme catalysis, substrates induce conformational changes in the enzyme to make the enzyme fit its favored substrate best in its active site for fast catalysis (74). It has been reasoned that the substrate-induced conformational changes should be large since they allow different substrates to access the open state of the enzyme and then enclose the good substrate to the closed state for fast catalysis. Therefore these kinds of conformational changes are important for specificity and catalysis for an enzyme.

DNA replication demands extraordinary specificity and efficiency of catalysis for DNA polymerases. DNA polymerases have to recognize and incorporate one correct nucleotide out of the pool of four nucleotides with similar properties during replication. It has been proposed that DNA polymerases follow the induced-fit mechanism for their catalysis and the existence of an open to closed conformational change of DNA polymerases was supported by kinetic and structural studies (2,22,35,41,70).

For the past two decades there has been considerable debate regarding the fidelity contribution of the conformational change step. The debate was mainly focused on whether the conformational change is rate-limiting and whether it contributes to specificity. If it is not rate-limiting but reaches equilibrium before chemistry, then the rate of the conformational change will not contribute to the specificity of an enzyme (1,75). Although earlier studies suggested that there

is a rate-limiting conformational change step preceding chemistry (2,40,41,52), kinetic studies on DNA polymerase β using the fluorescence signal from 2-AP and Taq DNA polymerase I using FRET suggested that the open-to-closed large conformational change is not rate-limiting , thus not contributing to the specificity of the enzymes (47,49,50,53). However, studies on T7 DNA polymerase in our lab using an MDCC fluorescent tag revealed that the forward rate of the conformational change preceding chemistry reported by MDCC was partially rate-limiting (69). The more important discovery from this study was that when the reverse rate of the conformational change is much slower than the chemistry step, only the forward rate of the conformational change and the ground-state nucleotide binding K_d dictates the specificity of correct nucleotide incorporation. This study demonstrated the import role of a conformational change in specificity and the proposed "new paradigm" for enzyme specificity may also apply to other enzymes.

The inconsistencies regarding whether the conformational change preceding chemistry is rate-limiting for various DNA polymerases might suggest that the conformational changes characterized by different methods are not the same, although the authors claimed that their observed signals provided a measurement of the open-to-closed conformational change. It is possible that there are indeed at least two conformational steps preceding chemistry (69). One step might correspond to the fast large open-to-closed movement of the fingers domain, and another step could be a critical, rate-limiting step corresponding to the subtle changes of the catalytic residues. However, more studies are needed to identify the origin of fluorescence signals designed to measure conformational changes.

In this study, the conformational change reported by the MDCC fluorescence was characterized by introducing several mutations within the fingers domain of T7 DNA polymerase. Figure 3.1 shows the structure of the exonuclease-free T7 DNA polymerase. After the examination of the structure of the recognition domain, we found there were two gly-ala-gly

sequences flanking the O1 helix. We speculated that these two gly-ala-gly loops may serve as hinges to allow the movement of the fingers domain due to the flexibility of the glycines. Here, we presented the study of these hinges by site-directed mutagenesis and transient kinetics (36,38,76). The mutations of these glycine residues to alanine decreased the forward rate of the conformational change dramatically. Our data suggested that the conformational change reported by MDCC corresponds to the large rotation of three catalytically important helices and it contributes significantly to the specificity of correct nucleotide incorporation.

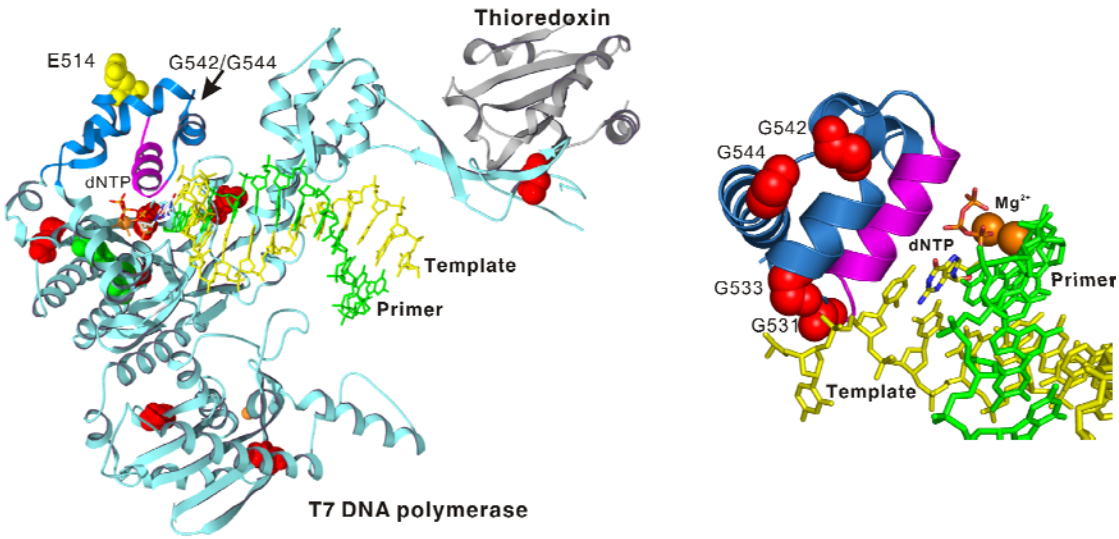


Figure 3.1: Fluorescently labeled G542A/G544A mutant

Left, it shows the exonuclease-free T7 DNA polymerase (D5A, E7A) with cysteine mutations in red: C20S, C88A, C275A, C313A, C451S, C660A, C688A, C703A. A single cysteine residue, E514C, was introduced for site-specific fluorescent labeling. Two glycine mutations, G542A and G544A, which are on one flexible loop were introduced for this study. Right, close view of the recognition domain. T7 DNA polymerase recognition domain was bound with DNA (template in yellow, primer in green), incoming nucleotide (in CPK color) and two magnesium metal ions (orange). G542, G544, G531 and G533 are shown in red.

3.2 MATERIALS AND METHODS

Construction, expression, purification, and fluorescently labeling of the hinge mutants

The double hinge mutant (G531A-G533A-G542A-G544A) and the single hinge mutant (G542A-G544A) were constructed by PCR method from a plasmid, pG5X-T7-8Cys-light-514C exo-, which encodes T7 DNA polymerase cys-light mutant (C20S-C88A-C275A-C313A-C451S-C660A-C688A-C703A-E514C) as described by Tsai and Johnson (69). For construction of the double hinge mutant, the PCR segments with the mutations were amplified by using DNA primer pairs: T7P-43F (GCT CGC CCG GGG ATC CGG) and 531G-A-rev (CAA CAA TCT GTC CAA TCT TCT CAT CAG CAG CAG CAT AGA GGA ACC CAT AGA TGA); 544G-A-for (TGA TGA GAA GAT TGG ACA GAT TGT TGC TGC TGC TAA AGA GCG CGG TAA GGA ACT C) and T7P-2435R (TAG AGC CAA GCT TGC ATG CC). Then the complete coding sequence for the double hinge mutant was created by joining the two PCR segments by another round of PCR with T7P-43F and T7P-2435R primer pairs. The plasmid and the PCR product were both digested by *Bam*HI and *Hind*III, and then were ligated into the final construct. For construction of the single hinge mutant, the only difference from the double hinge mutant is that the first PCR segment (no mutations) was synthesized from primers: T7P-43F (GCT CGC CCG GGG ATC CGG) and 531-544-joint-rev (CAA CAA TCT GTC CAA TCT TCT CAT CAC C). The final constructs of both mutants were confirmed by DNA sequencing.

The hinge mutants were expressed, purified, and fluorescently labeled as described in chapter 2 with some modifications. After addition of 0.4 mM IPTG and 12.5 μg/ml of chloramphenicol to induce protein expression, the culture was grown for overnight at 14 °C to get more soluble proteins. Another change was that 0.5 M NaCl was added to the cell lysate before addition of 0.5% PEI to increase protein yield.

DNA substrates for kinetic studies

DNA oligomers were obtained from Integrated DNA Technologies (IDT, Inc.) and purified by using 15% polyacrylamide/7 M urea denaturing gel electrophoresis. A 27-mer primer (5'GCC TCG CAG CCG TCC AAC CAA CTC AAC3') and a 45-mer template (5'GGA CGG CAT TGG ATC GAG GTT GAG TTG GTT GGA CGG CTG CGA GGC3') were adopted from a previous study to preserve comparison between all our relative studies.

There were two versions of the 27-mer primer. For nucleotide binding studies, a ddCMP-terminated 27-mer primer was used to prevent incorporation of the next incoming nucleotide, dCTP. The dCMP-terminated primer was used for single nucleotide incorporation experiments and it was 5'-^{32}P-labeled using T4 polynucleotide kinase according to the manufacturer's instructions (Invitrogen). The labeling reaction was terminated by heating at 95 ºC for 5 min and excess ^{32}P-labeled nucleotide was removed by using a Biospin 6 column (Bio-Rad).

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

Chemical quench experiments

Single nucleotide incorporation assays were performed at 20 ºC in T7 reaction buffer containing 40 mM Tris (pH 7.5), 1mM EDTA, 50 mM NaCl, 1 mM DTT, 12.5 mM $MgCl_2$. For dCTP and dCTPαS incorporation assays, an RQF-3 rapid-quench-flow apparatus (KinTek Corp.) was used. One sample loop was loaded with preformed enzyme-DNA complex (600 nM enzyme, 12μM thioredoxin and 200 nM 5'-^{32}P-labeled 27-mer/45-mer DNA duplex) in T7 reaction buffer without $MgCl_2$. Another loop was loaded with nucleotide and 25 mM $MgCl_2$. The reaction was started by rapidly mixing the two reactants and then the reaction was quenched by mixing with 0.5 M EDTA after time intervals ranging from several milliseconds to several seconds.

Nucleotide misincorporation assays were done manually without the quench flow instrument since the reactions took longer time. An enzyme-DNA complex was manually mixed with nucleotide and $MgCl_2$ to start the reaction. Aliquots were pipetted from the reaction mixture at designated time intervals and immediately mixed with 0.5 M EDTA to stop the reactions.

Pyrophosphorolysis assays were set up as described for nucleotide misincorporation assays, except that pyrophosphate was added instead of nucleotide.

The processivity assay was set up as described for dCTP incorporation assay, except that in one sample loop dCTP, dTTP and dGTP (100 μM each) were added instead of only dCTP allowing the sequential incorporation of four nucleotides.

The products from the assays described above were resolved on a 15 % denaturing polyacrylamide sequencing gel and then the dried gel was exposed to a phosphor screen. The screen was scanned by a Storm 860 scanner (GE Healthcare). The amount of product formation at each time interval was analyzed by the ImageQuant software (GE Healthcare).

Fluorescence emission spectra

Emission spectra were recorded using a PTI fluorometer (PTI, Inc.). The excitation wave length was set at 425 nm. The fluorescence emission spectrum of a 199 μl enzyme (200 nM), thioredoxin (4 μM) and DNA (300nM, ddCMP terminated 27-mer was used) solution in T7 reaction buffer was recorded first. The emission spectrum ranged from 440 nm to 540 nm. Then 1 μl of 100 mM dCTP (correct) or dGTP (incorrect) was added to the solution and the emission spectrum was recorded again. Spectra were normalized by dividing all the emission data to the peak of the emission spectrum of the enzyme-DNA complex.

Stopped-flow experiments

All experiments were performed at 20 ºC in T7 reaction buffer using a SF2004 stopped-flow apparatus (KinTek Corp.). The excitation wave length was set at 425 nm and a photomultiplier with a 450 nm cutoff high pass filter was used to monitor the fluorescence. For

nucleotide binding stopped-flow assays, one syringe was loaded with the preformed enzyme-DNA complex (400 nM enzyme, 8 mM thioredoxin, and 600 nM DNA with ddCMP terminated 27-mer primer) in T7 reaction buffer with $MgCl_2$ and another syringe was loaded with the next incoming nucleotide at various concentrations in T7 reaction buffer with $MgCl_2$. Aliquots of these two syringes were rapidly mixed and the fluorescence emission from the reaction was recorded.

For nucleotide release assays, one syringe was loaded with preformed enzyme-DNA-dCTP complex (0.8 μM enzyme, 1.6 μM DNA and 1 μM dCTP) in T7 reaction buffer with $MgCl_2$ and another syringe was loaded with unlabeled enzyme-DNA complex (3 μM wild type T7 DNA polymerase and 6 μM DNA with dCMP terminated 27-mer) in T7 reaction buffer with $MgCl_2$.

Equilibrium titration experiments

Nucleotide binding titration assays were carried out at 20 ºC in T7 reaction buffer with 12.5 mM $MgCl_2$ using a KinTek TMX Titration Module. The excitation wave length was set at 425 nm and a photomultiplier with a 450 nm cutoff high pass filter was used to monitor fluorescence. A 20 μl solution of 250 μM dCTP in T7 reaction buffer was slowly injected into and rapidly mixed with 280 μl of preincubated enzyme-DNA complex (200 nM enzyme, 4 μM thioredoxin and 300 nM DNA with ddCMP terminated 27-mer primer) in the reaction cuvette. During the injection which lasted 180 seconds, the fluorescence emission from the reaction was recorded continuously and was corrected for the small dilution.

Data fitting and analysis

The data were analyzed by nonlinear regression using GraFit 5 software (Erithacus Software Limited). Each time-course of single nucleotide incorporation at various nucleotide concentrations was fit to a single exponential equation, $y = A \cdot \exp(-k_{obs}t) + C$, where A is the amplitude, k_{obs} is the observed rate, and C is the endpoint. The concentration dependence of the rate of incorporation was fit to a hyperbola, $k_{obs} = k_{pol} \cdot S_0/(K_d+S_0)$, where k_{obs} is observed rate, k_{pol}

is maximum incorporation rate, S_0 is variable starting nucleotide concentration, and K_d is apparent ground-state binding constant. The pyrophosphorolysis data were analyzed similarly. Each time course of fluorescence change from the stopped-flow experiments was also fit to a single exponential to extract the rate and amplitude of the change. The data from nucleotide releasing were fit to a single exponential. The data from the titration experiment were fit to a hyperbolic equation, $F = F_0 + \Delta F \cdot S_0/(K_d+S_0)$, where F_0 is the starting fluorescence, ΔF is the fluorescence change.

3.3 RESULTS

Generation of the hinge mutants

To characterize the conformational change reported by MDCC during nucleotide binding and incorporation, all the mutants were constructed based on a cys-light T7 DNA polymerase first made by Tsai and Johnson (69). In this cys-light mutant, all of the surface cysteines were replaced by alanines or serines, and a cysteine substituted for a glutamate at 514 to react with MDCC, thus only specific labeling of MDCC at 514C was observed. The double mutation D5A, E7A completely inactivated the exonuclease activity, so the studies at the polymerase active site were not affected by exonuclease activity. The characterization of this cys-light mutant showed nearly full activity with only modest changes in the kinetic parameters compared to the wild type (exo^-) T7 DNA polymerase.

In this study, a double hinge mutant (G531A-G533A-G542A-G544A) was first made by mutating four glycines within the hinges to alanines. Nucleotide incorporation assay showed that this mutant lost all of its catalytic ability (data not shown). This might be caused by misfolding of the protein or blockage of the motion of the fingers domain. However, inspection of the structure suggested that the two mutations, G531A and G533A, might sterically interfere the binding of the DNA template. Since this mutant was not active, further characterization was not performed. We made a single hinge mutant (G542A-G544A), which was then labeled with MDCC at 514C.

Nucleotide incorporation burst assays of this mutant before and after labeling showed that it was active and that the fluorescent labeling did not change the burst rate of this mutant significantly (data not shown). The remaining studies were performed using this G542A-G544A mutant which we referred to as the "hinge mutant".

Correct nucleotide incorporation and thio-elemental effect

To measure the rate of incorporation of a correct nucleotide, single turnover quench flow experiments were performed by mixing a preformed enzyme-DNA complex with a solution containing the correct nucleotide, dCTP, at various concentrations. The time course of product formation at each dCTP concentration was fit to a single exponential to obtain the rate of nucleotide incorporation. The dCTP concentration dependence of the rate of incorporation was fit to a hyperbola, yielding a maximum nucleotide incorporation rate, k_{pol} = 0.52 ± 0.02 s^{-1}, and an apparent nucleotide dissociation constant, K_d = 1.7± 0.2 μM (Figure 3.2A). Compared to the previously reported values of k_{pol} = 234 ± 9.4 s^{-1}, and K_d = 24 ± 3.1 μM for the cys-light T7 DNA polymerase, the hinge mutant showed about a 450-fold decrease in nucleotide incorporation rate and about a 14-fold increase in the apparent affinity of nucleotide binding.

Since the mutations at residues 542 and 544 are remote from the active site and not directly involved in catalysis, we proposed that these mutations mainly slow the conformational change step. To examine whether the chemistry step was rate-limiting, we studied the thio-elemental effect of nucleotide incorporation. The thio-elemental effect was defined by the rate of incorporating a normal dNTP divided by the rate of incorporating a corresponding α-thio-substituted dNTP (with a *pro*-Sp non-bridging oxygen of the α-phosphate replaced by a sulfur) ($k_{pol,dNTP}/k_{pol,dNTP\alpha S}$). Since the sulfur decreases the rate of phosphoryl transfer due to its decreased electronegativity compared to oxygen, a large thio-elemental effect will be observed if the chemistry step is rate-limiting (2,39,41). We studied the dCTPαS concentration dependence of the rate of nucleotide incorporation using single turnover quench flow method (Figure 3.2B). The

data were fit to a hyperbola, defining a k_{pol} = 0.60 ± 0.04 s^{-1}, and a K_d = 2.4 ± 0.4 μM. Since $k_{pol,dCTP}$ / $k_{pol,dCTP\alpha S}$ = 0.52/0.6 ≈ 0.9, almost no thio-elemental effect was observed. This suggested that the rate-limiting step is not the chemistry step but rather a conformational change step upstream in the pathway.

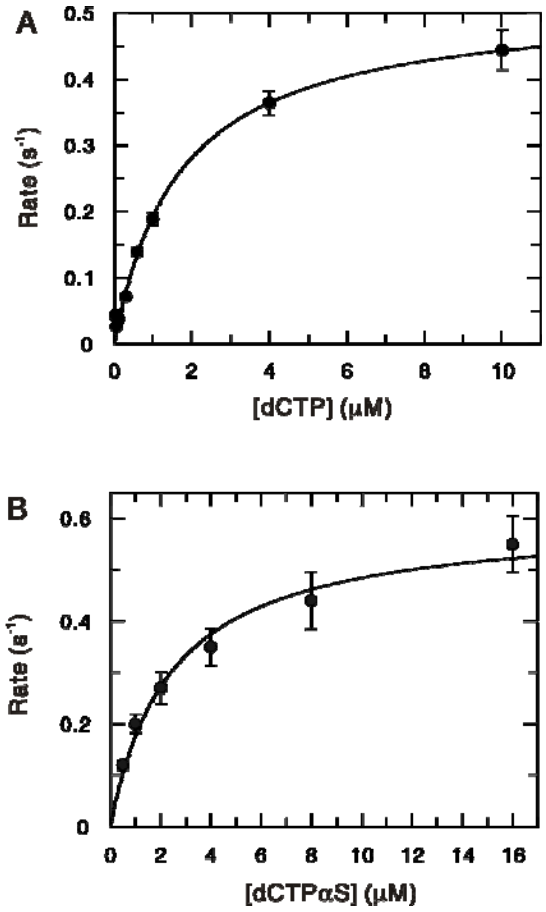


Figure 3.2: dCTP and dCTPαS incorporation kinetics

(A) A preformed enzyme-DNA complex (600 nM enzyme and 200 nM DNA) was rapidly mixed (1:1) with dCTP solution at various concentrations to start the experiments. The time course of product formation at each dCTP concentration was fitted to a single exponential to obtain the incorporation rate. The dCTP concentration dependence of incorporation rates was fitted to a hyperbola, defining an apparent K_d = 1.7 ± 0.2 μM, and a maximal incorporation rate k_{pol} = 0.52 ± 0.02 s^{-1}. (B) The dCTPαS concentration dependence of nucleotide incorporation rates defined an apparent K_d = 2.4 ± 0.4 μM, and a maximal incorporation rate k_{pol} = 0.60 ± 0.04 s^{-1}.

Kinetics and equilibrium of correct nucleotide binding

To test whether the conformational change was the rate-limiting step as suggested by the thio-elemental effect study, the conformational change was characterized by several fluorescence measurements. Fluorescence emission scans were performed to examine the fluorescence emission of the MDCC labeled hinge mutant with different substrates (Figure 3.3A). In these experiments, a DNA with dideoxy-terminated primer was used to prevent the chemistry step, so only the binding of the next incoming nucleotide would be observed. Upon binding a correct nucleotide, dCTP, to the enzyme-DNA complex, the intensity of the fluorescence decreased about 30%. In contrast, after binding of an incorrect nucleotide, dGTP, an 18% increase of the fluorescence intensity was observed. The fluorescence emission spectra of this hinge mutant were similar to those of the cys-light T7 DNA polymerase in various nucleotide-binding states: fluorescence decreased upon binding a correct nucleotide but increased upon binding an incorrect nucleotide. This suggested that the fluorescence signal that reported the conformational change for the cys-light T7 DNA polymerase was not distorted significantly by the hinge mutations.

The kinetics of the conformational change reported by MDCC were measured by stopped-flow methods. The enzyme-DNA (with dideoxy-terminated primer) complex was rapidly mixed with a solution containing the incoming correct nucleotide, dCTP, at various concentrations in the stopped-flow instrument and the fluorescence changes following mixing were recorded. Figures 3.3B and 3.4A show the time courses of fluorescence changes upon binding dCTP at various concentrations. Each curve was fit to a single exponential by nonlinear regression to obtain the amplitude and rate of the fluorescence change. The dCTP concentration dependence of the amplitude was fit to a hyperbola, defining an apparent dissociation constant, $K_{d,app} = 0.52 \pm 0.04$ μM (Figure 3.4B). This apparent dissociation constant is an estimate of a net equilibrium dissociation constant, $K_{d,net} = K_{d,1}/(1+K_2)$, where $K_{d,1}$ is the ground-state binding

constant and K_2 is the equilibrium constant of the second step according to a two-step binding mechanism:

$$E \cdot D_{dd} + dCTP \overset{K_d}{\rightleftarrows} E \cdot D_{dd} \cdot dCTP \underset{k_{-2}}{\overset{k_2}{\rightleftarrows}} F \cdot D_{dd} \cdot dCTP$$

Scheme 3.1: A two-step nucleotide binding mechanism

According to this pathway, the conformational change rate should show dCTP concentration dependence, following the function of $k_{obs} = k_2 \cdot S_0/(K_d+S_0)+k_{-2}$. However, the rates from the data fitting fluctuated over a small range, and seemed to have a more complex relationship with the nucleotide concentration (Figure 3.4C). One explanation could be that the k_2 is much smaller than k_{-2} thus k_{obs} is dominated by k_{-2} and not sensitive to concentration. The smooth line shown in Figure 3.4C was the data fitting result using an equation with two hyperbolas, $k_{obs} = k_0+\Delta k_{max1}*S/(K_{d1}+S)+ \Delta k_{max2}*S/(K_{d2}+S)$, and this would imply two different binding sites. However, we don't have more data to support this possibility and this phenomenon could be an artifact from the single exponential fitting of the fluorescence change that was contaminated with a steady-state downward drift.

To estimate the reverse rate of the conformational change, we measured the rate of releasing bound dCTP from the preformed enzyme-DNA-dCTP tertiary complex (Figure 3.3C). In the stopped-flow instrument, a solution of enzyme-DNA-dCTP complex was rapidly mixed with a solution containing excess enzyme-DNA complex (preformed between an unlabeled wild type (exo$^-$) T7 DNA polymerase and a DNA with deoxy-terminated primer) which can incorporate dCTP at about 300 s^{-1} so that all of the free dCTP could be removed rapidly from solution. Since this is the reverse reaction of dCTP binding, a fluorescence increase was observed. The time course of the increasing fluorescence was fit to a single exponential, yielding a rate of 1.08 ± 0.01 s^{-1}, which is close to the k_{-2} of the cys-light T7 DNA polymerase (1.6 s^{-1}). Therefore, if we assigned the rate of nucleotide incorporation of 0.5 s^{-1} to the rate-limiting conformational

change (k_2), then the observed rates of dCTP binding from the stopped-flow experiments were dominated by the reverse rate of conformational change (k_{-2} = 1.1 s^{-1}). Based on all these data, the reaction pathway for this hinge mutant was proposed as shown in scheme 3.2.

$$E \cdot D_n + dCTP \underset{}{\overset{1.7\,\mu M}{\longleftrightarrow}} E \cdot D_n \cdot dCTP \underset{1.1 s^{-1}}{\overset{0.5 s^{-1}}{\longleftrightarrow}} F \cdot D_n \cdot dCTP \longleftrightarrow F \cdot D_{n+1} \cdot PP_i$$

Scheme 3.2: Pathway of correct nucleotide binding of the hinge mutant

As an independent check, the net equilibrium dissociation constant was measured by a titration experiment (Figure 3.3D). A solution of the enzyme-DNA complex was constantly stirred with addition of a solution of dCTP at a slow constant rate and the fluorescence intensity was recorded continuously. The dCTP concentration dependence of the fluorescence change from this titration experiment was fit to a hyperbola, defining the net equilibrium dissociation constant, $K_{d,net}$ = 1.12 ± 0.01 μM. According to the mechanism shown in scheme 3.2, the calculated $K_{d,net}$ = $K_{d,1}/(1+K_2)$ = 1.7/(1+0.5/1.1) ≈ 1.2 μM, and thus it agrees very well with the results of the titration experiment.

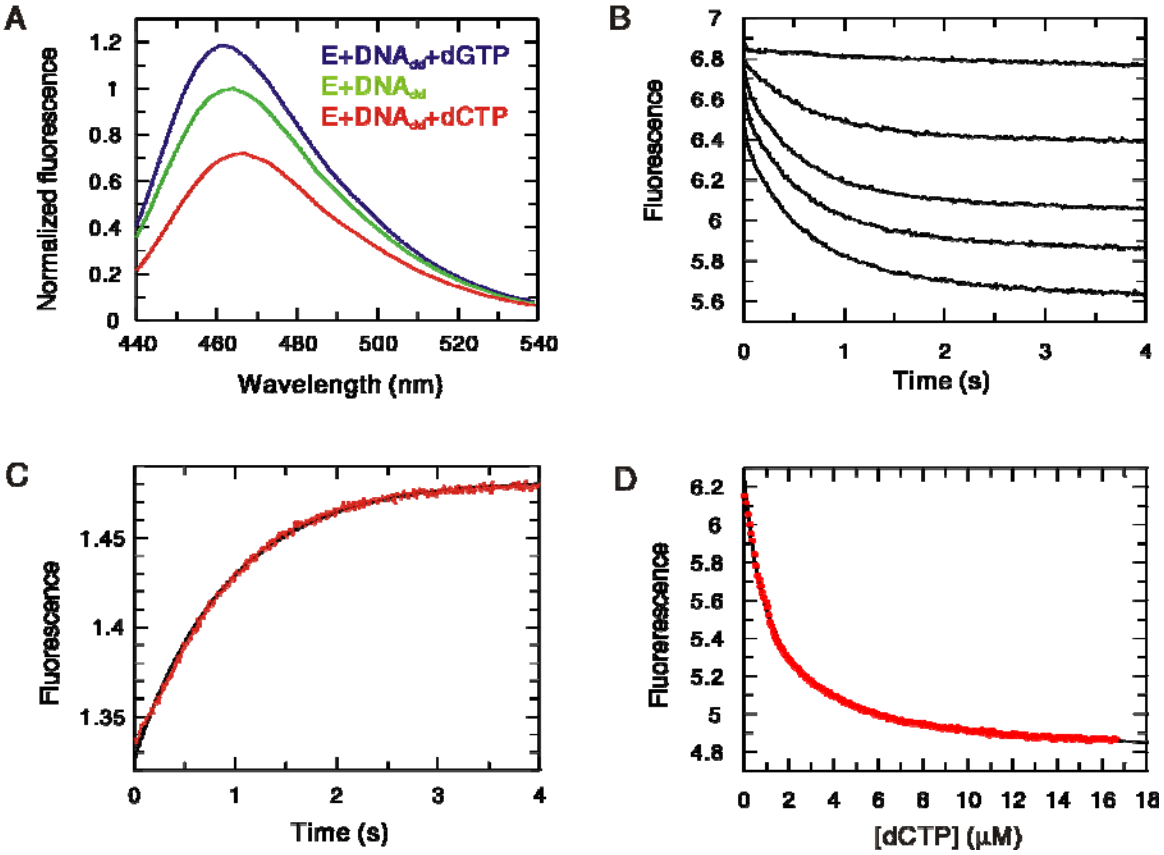


Figure 3.3: Kinetics and equilibrium of fluorescence changes observed from nucleotide binding

Figure 3.3: Kinetics and equilibrium of fluorescence changes observed from nucleotide binding. (A) Fluorescence emission spectra of CSF labeled mutant enzyme at different substrate-bound states are shown. (B) Stopped-flow fluorescence changes of the mutant enzyme induced by binding of dCTP at various concentrations. Preformed enzyme-DNA complex (200 nM enzyme and 600 nM DNA_{ddC}) was rapidly mixed (1:1) with dCTP solutions (0, 1, 4, 8, and 16 μM, from top to bottom) and the time courses of fluorescence change were recorded. (C) The time course of fluorescence increase induced by release of dCTP from Enzyme-DNA-dCTP ternary complex. The data were fit to a single exponential with a rate of 1.08 ± 0.01 s^{-1}. (D) Equilibrium titration for dCTP binding to enzyme-DNA_{ddC} duplex. A 280 μl of preformed 200 nM enzyme and 300 nM DNA_{ddC} solution was titrated with 20 μl of 250□ μM dCTP. The concentration dependence of the fluorescence decrease was fitted to a hyperbola, yielding a dissociation constant of 1.12 ± 0.01μM. This constant is $K_{d(net)} = K_d/(1+K_2)$.

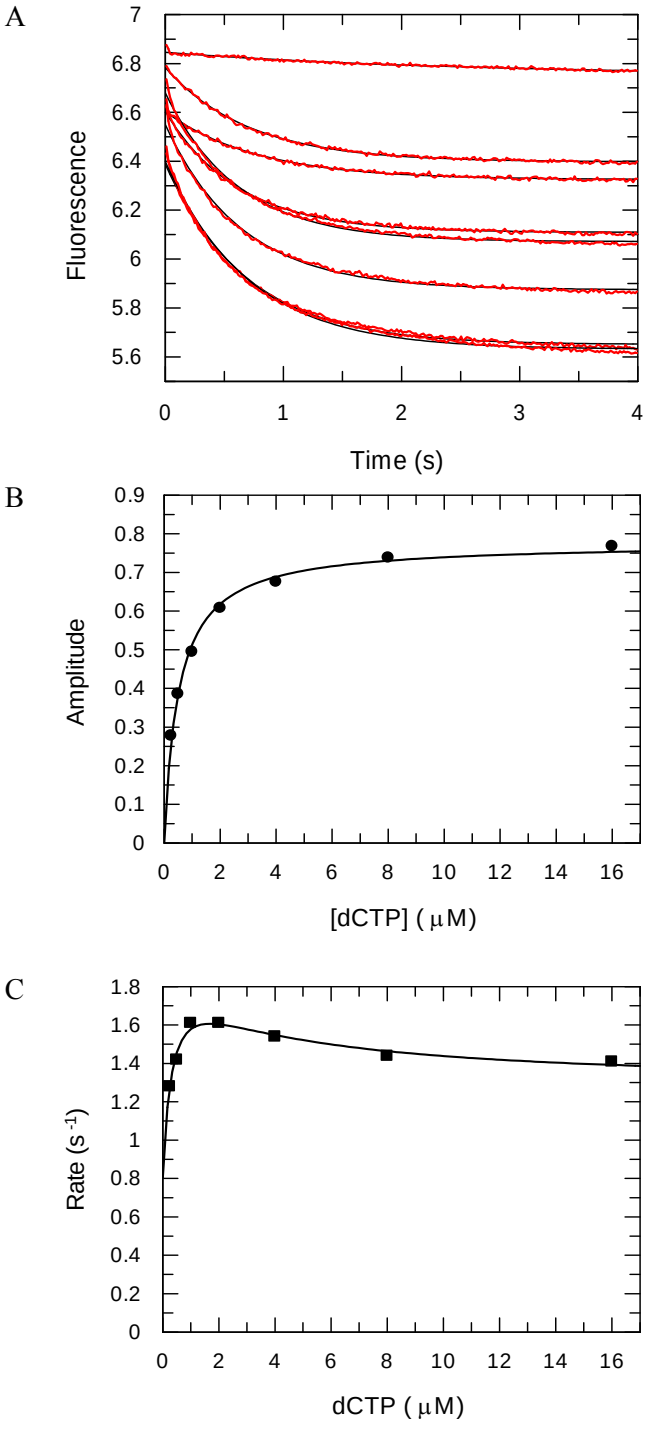


Figure 3.4: Kinetics of fluorescence changes observed from dCTP binding

Figure 3.4: Kinetics of fluorescence changes observed from dCTP binding.

(A) Fluorescence changes of the mutant enzyme induced by binding of dCTP at various concentrations (from top to bottom, 0, 0.25, 0.5, 1, 2, 4, 8, and 16 μM dCTP after mixing, respectively). Each fluorescence trace was fit to a single exponential to derive the rate and the amplitude. (B) The dCTP concentration dependence of the amplitude was fit to a hyperbola, yielding a $K_d = 0.52 \pm 0.04$ μM. (C) The dCTP concentration dependence of the rate.

Kinetics and equilibrium of dCTPαS binding

The kinetics of the conformational change induced by the dCTPαS binding was also studied by stopped-flow methods. Figure 3.5A shows the recorded fluorescence transients after mixing the enzyme-DNA complex with dCTPαS at various concentrations. Each fluorescence transient was fit to a single exponential to obtain amplitude and rate of the fluorescence change corresponding to the conformational change. Similar to the binding of dCTP, the amplitude increased with increasing dCTPαS concentration, and the concentration dependence of the amplitude was fitted to a hyperbola, defining a $K_{d,net}$ = 1.6 ± 0.2 μM (Figure 3.5B). The observed burst rate of the fluorescence transient appeared to decrease hyperbolically with dCTPαS concentration (Figure 3.5C). This phenomenon appeared to be opposite to the simple two-step binding mechanism, but we did not have other evidence to propose a more complex mechanism. This odd phenomenon could be caused by a faster reverse rate of the conformational change as discussed above. To test this hypothesis, we measured the rate of dCTPaS release from the enzyme-DNA-dCTPaS ternary complex by stopped-flow methods (Figure 3.6A). The time course of fluorescence increase corresponding to dCTPαS release was fit to an exponential, defining an off rate of 1.36 ± 0.01 s^{-1}. Based on the quench flow and stopped-flow data, the pathway of dCTPaS incorporation was proposed in scheme 3.3.

$$E \cdot D_n + dCTP \xleftrightarrow{2.4\,\mu M} E \cdot D_n \cdot dCTP \underset{1.4\,s^{-1}}{\overset{0.6\,s^{-1}}{\rightleftharpoons}} F \cdot D_n \cdot dCTP \longleftrightarrow F \cdot D_{n+1} \cdot PP_i$$

Scheme 3.3: Pathway of dCTPaS binding of the hinge mutant

As an independent check, the net equilibrium dissociation constant was measured by a titration experiment under the some condition for dCTP (Figure 3.6B). The dCTPαS concentration dependence of the fluorescence change from this titration experiment was fit to a hyperbola, defining the net equilibrium dissociation constant, $K_{d,net}$ = 1.71 ± 0.04 μM. According

to the mechanism shown in scheme 3.3, the calculated $K_{d,net} = K_{d,1}/(1+K_2) = 2.4/(1+0.6/1.4) \approx$ 1.68 μM, which agrees very well with the results of the titration experiment and the K_d of 1.6 μM from the concentration dependence of the amplitude of the stopped-flow experiment of dCTPαS binding.

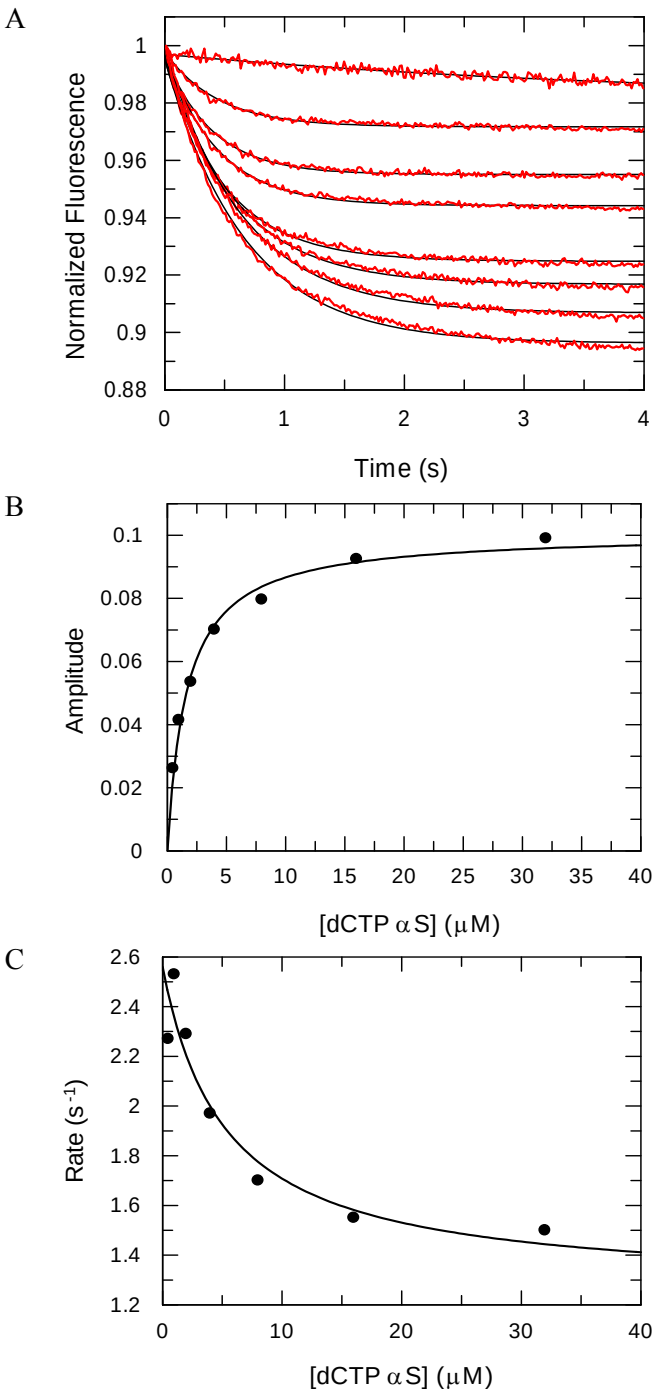


Figure 3.5: Kinetics of fluorescence changes observed from dCTPαS binding

Figure 3.5: Kinetics of fluorescence changes observed from dCTPαS binding.

(A) Stopped-flow fluorescence changes of the mutant enzyme induced by binding of dCTPαS at various concentrations (from top to bottom, 0, 0.5, 1, 2, 4, 8, 16, and 32 μM dCTPαS after mixing, respectively). Each fluorescence trace was fit to a single exponential to derive the rate and the amplitude. (B) The dCTPαS concentration dependence of the amplitude was fit to a hyperbola, yielding a $K_d = 1.6 \pm 0.2$ μM. (C) The dCTPαS concentration dependence of the rate.

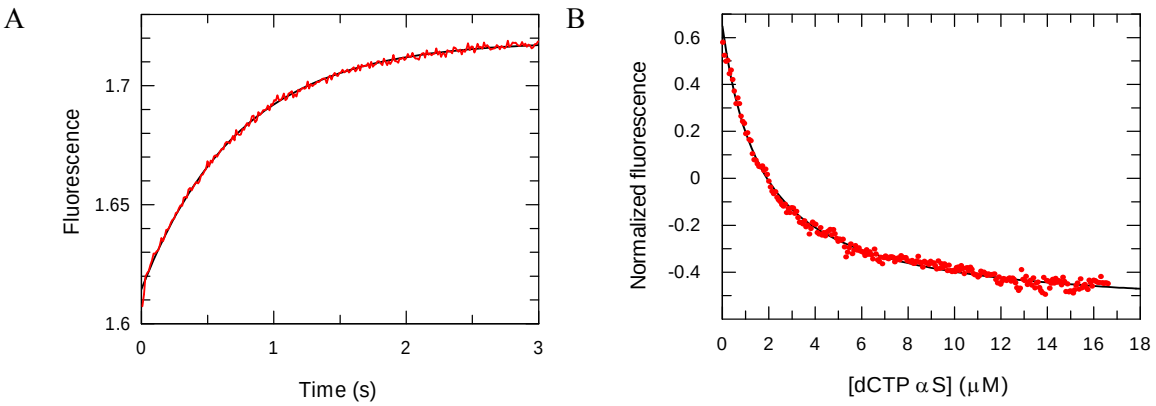


Figure 3.6: dCTPαS off rate measurement and titration experiment

(A) Time course of fluorescence increase induced by release of dCTPαS from Enzyme-DNA_{ddC}-dCTPαS ternary complex. The data were fit to a single exponential with a rate of 1.36 ± 0.01 s^{-1}. (B) Equilibrium titration for dCTPαS binding to enzyme-DNA_{ddC} duplex. A 280 µl of preformed 200 nM enzyme and 300 nM DNA_{ddC} solution was titrated with 20 µl of 250 µM dCTPαS. The concentration dependence of the fluorescence decrease was fitted to a hyperbola, yielding a dissociation constant of 1.71 ± 0.04 µM. This constant is $K_{d(net)} = K_d/(1+K_2)$.

Misincorporation and discrimination

Misincorporation of dATP, dGTP, or dTTP opposite a dG on the template was examined under single turnover conditions using the hinge mutant, as described in materials and methods. The time course of misincorporation at each nucleotide concentration followed a single exponential function. The nucleotide concentration dependence of the misincorporation rate was fit to a hyperbola to derive the ground-state binding K_d and the maximum incorporation rate, k_{pol}. A summary of the results for the K_d, k_{pol}, specificity constant and discrimination are reported in Table 3.1.

Table 3.1: Fidelity quantification of the hinge mutant

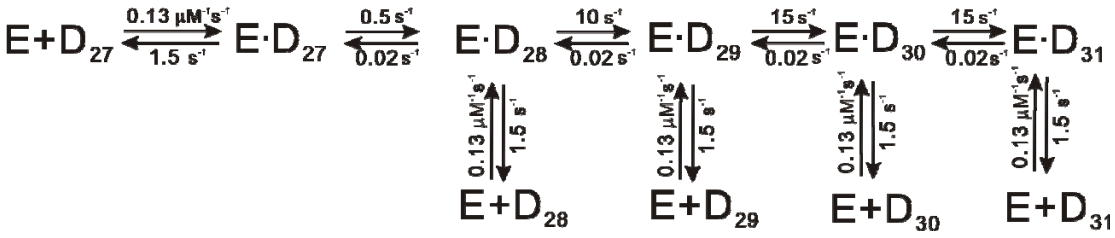
dNTP	k_{pol} (s^{-1})	K_d (μM)	k_{pol}/K_d ($\mu M^{-1}s^{-1}$)	Discrimination
dCTP	0.52	1.7	0.31	1
dATP	0.13	600	0.00022	1,400
dTTP	0.19	2,800	0.000068	4,600
dGTP	0.079	1,900	0.000042	7,400

Pyrophosphorylysis and processivity assays

To measure the rate-limiting step of the reverse reaction of nucleotide incorporation, the pyrophosphorolysis reaction was studied. As the reverse of nucleotide incorporation, the pyrophosphorolysis reaction catalyzes reattachment of pyrophosphate to a dNMP at the primer end and then releases a dNTP, so the primer is one nucleotide shorter after a single turnover. In this experiment, an enzyme-DNA complex was rapidly mixed with a solution containing pyrophosphate at various concentrations to start the reactions. Each time course of product formation was fit to a single exponential to derive the rate of pyrophosphorolysis at each pyrophosphate concentration. A fit of the concentration dependence of the observed rate to a

hyperbola yielded an apparent K_d = 3.8 ± 7.2 mM and a maximum pyrophosphorolysis rate of 0.02 ± 0.02 s^{-1} (Figure 3.7A).

As reported previously, wild type T7 DNA polymerase has a high processivity of 1500 during DNA replication. For the hinge mutant, we predicted it would have a much lower processivity due to the slower conformational change. A processivity assay was performed in a quench flow instrument by rapidly mixing the enzyme-DNA complex with a solution containing three dNTPs and excess non-radiolabeled DNA and then the reaction was quenched with 0.5 M EDTA at various time intervals. Three dNTPs were used to elongate the 27-mer primer only to a 31-mer based on the sequence on the template, thus simplifying the kinetic analysis. Excess non-radiolabeled DNA was added to trap the elongated products that dissociated from the enzyme during reaction to ensure the observed product formation was from a single binding event. Figure 3.7B shows the data analysis of this experiment. The time courses of the intermediates formation and disappearance were simulated using the KinTek Explorer program and the resulting rate constants are shown in Scheme 3.4. Different from the processivity assay of the wild type T7 DNA polymerase, the levels of the intermediates stayed low, so the appearance and disappearance of the successive intermediates could not be quantified accurately. Only the burst and steady state formation of the 31-mer was observed. The data were fit to a mechanism in which the rate of the first incorporation (0.5 s^{-1}) is much slower than the rates (10 s^{-1} or 15 s^{-1}) of the subsequent incorporations, and the DNA dissociation rate (1.5 s^{-1}) from the enzyme-DNA-dNTP complex is much faster than the DNA off rate (0.2 s^{-1}) from enzyme-DNA after pyrophosphate release. The faster ensuing steps for nucleotide incorporation can account for the absence of the rise-and-disappearance observation of the intermediates. The nucleotide binding induced fast DNA release can explain why the burst amplitude is so low and why intermediate bands did not disappear completely in our assay.



Scheme 3.4: Processive polymerization of T7 DNA polymerase hinge mutant

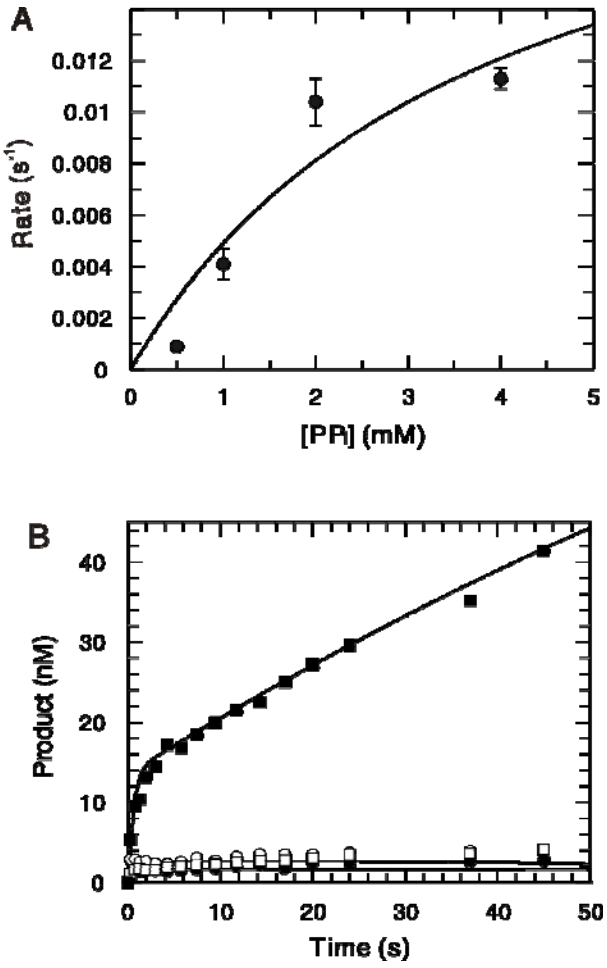


Figure 3.7: Kinetics of pyrophosphorolysis and processive polymerization

Figure 3.7: Kinetics of pyrophosphorolysis and processive polymerization.

(A) Single turnover pyrophosphorolysis kinetics of the mutant enzyme. A solution containing 600 nM enzyme and 200 nM 27/45-mer DNA was mixed with pyrophosphate at various concentrations to start the reaction. Each time course of product formation (26-mer) was fitted to a single exponential to obtain the pyrophosphorolysis rate at each pyrophosphate concentration. The PP_i concentration dependence of the rate was fit to a hyperbola, yielding a Kd of 3.8±7.2 mM and a maximum pyrophosphorolysis rate equal to 0.024 ± 0.026 s^{-1}. (B) Processive polymerization of 27/45-mer to 31/45-mer. The experiment was conducted by mixing the mutant polymerase (600 nM) and 5'-^{32}P-labeled 27/45-mer (200nM) with dCTP, dTTP, and dGTP (100 mM) in magnesium buffer. The reactions were quenched with 0.5 M EDTA and analyzed by sequencing gel electrophoresis. The data points (○), 28-mer; (●), 29-mer; (□), 30-mer; (■), 31-mer were obtained by quantitating the intermediate products from a scan of a phosphoscreen. The solid lines are curves obtained from computer simulation using a mechanism shown in Scheme 3. The rate constants for formation of each intermediate species are: 28-mer at 0.5 s^{-1}, 29-mer at 10 s^{-1}, 30-mer at 15 s^{-1}, 31-mer at 15 s^{-1}.

3.4 DISCUSSION

Mechanism of the hinge mutant

Our data supported the conclusion that the hinge mutations slow down the conformational change as reported by MDCC, which becomes the rate-limiting step for a single nucleotide incorporation. The small thio effect observed for correct nucleotide incorporation suggested that a step preceding chemistry limits the rate of reaction at 0.5 s^{-1}. This rate could be assigned to the rate of the conformational change reported by MDCC. The direct measurement of the rate of this conformational change step was done by stopped flow experiments. We found that the concentration dependence of the rate of correct nucleotide binding did not follow an expected pattern as defined by $k_{obs} = k_2 * S_0/(K_d + S_0) + k_{-2}$ which is derived from a two-step binding mechanism. However, by inspecting this equation, we notice that when k_{-2} is much faster than k_2, the observed rate k_{obs} will be dominated by k_{-2} and insensitive to nucleotide concentration. The reverse rate of the conformational change (k_{-2}) is 1.1 s^{-1} by measurement. If we assign the rate of 0.5 s^{-1} to the forward rate of the conformational change (k_2), these two rates can explain the fluctuation of the rate of the conformational change within a small range observed in the stopped flow experiments, including experimental error. Combined with the data measured from the quench flow experiments, we assigned the kinetic parameters to each step in a mechanism shown in Scheme 3.2. The calculation of the net equilibrium dissociation constant from these parameters, $K_{d,net} \approx 1.2$ μM, is comparable with the $K_{d,net} = 1.1$ μM from the titration experiment. This confirmed that there are no missing steps along the pathway of nucleotide binding. Therefore, all the data are consistent with the mechanism shown in Scheme 3.2, in which the conformational change reported by MDCC is rate-limiting. Other rate constants along the reaction pathway such as the chemistry step and pyrophosphate release are not defined by the data presented here.

The identity of the conformational change reported by MDCC

The conformational change characterized in the hinge mutant may be similar to that one reported for the cys-light T7 DNA polymerase (69). The profiles of the fluorescence emission of the enzymes bound with a correct or incorrect nucleotide are similar. They both showed a decrease in fluorescence upon binding a correct nucleotide, and an increase upon binding an incorrect nucleotide. Therefore, the hinge mutations did not change the direction of this conformational change but did decrease the rate. Structural data suggests that the conformational change is not as a rigid body as the generally accepted open-to-closed fingers domain movement alluded to; rather, the structural change from open to closed states involves movements of several helices within the fingers domain and mainly the rotation of three helices.

Figure 3.8 depicts the structural comparison of the fingers domain between the open and the closed complexes of T7 DNA polymerase. Most changes in comparing the two states are from the colored helices within the fingers domain. The O2 helix only rotates along its long axis slightly and the N helix tilted a small angle outwards from the plane of the image. The significant changes are in the O and O1 helices, which flip down together. As the result of all these changes, the back side of the fingers domain appears more solvent accessible in the closed complex than in the open complex as judged by the distance between the E514 (colored in magenta and where the MDCC is located) and the O2 helix. This observation might explain the origin of the fluorescence change since the MDCC tag sits on a loop of which the environment can be changed between the open and closed states of the enzyme. The two glycines at 542 and 544 are shown as spheres in the pictures and they connect the O1 helix with the O2 helix. Obviously, mutations on these two glycines should exert an effect on the large motion of the O and O1 helices. Then the slow conformational change as reported by MDCC for the hinge mutant could be strongly correlated with the large motion of the O and O1 helices. Since these two glycines are remote from the active site, they should not affect the subtle changes of the active site residues, which have been

proposed to be possible rate-limiting steps. Thus our data did not provide evidence for other more subtle rate-limiting steps preceding chemistry such as rotations of side chains or rearrangement of the active site residues. The original conception of the open-to-closed conformational change gave the impression of a rigid body motion of the whole fingers domain. We believe that this motion is more complex and involves the movements of several secondary structures within the fingers domain while the movement of the O and the O1 helices is large compared to others.

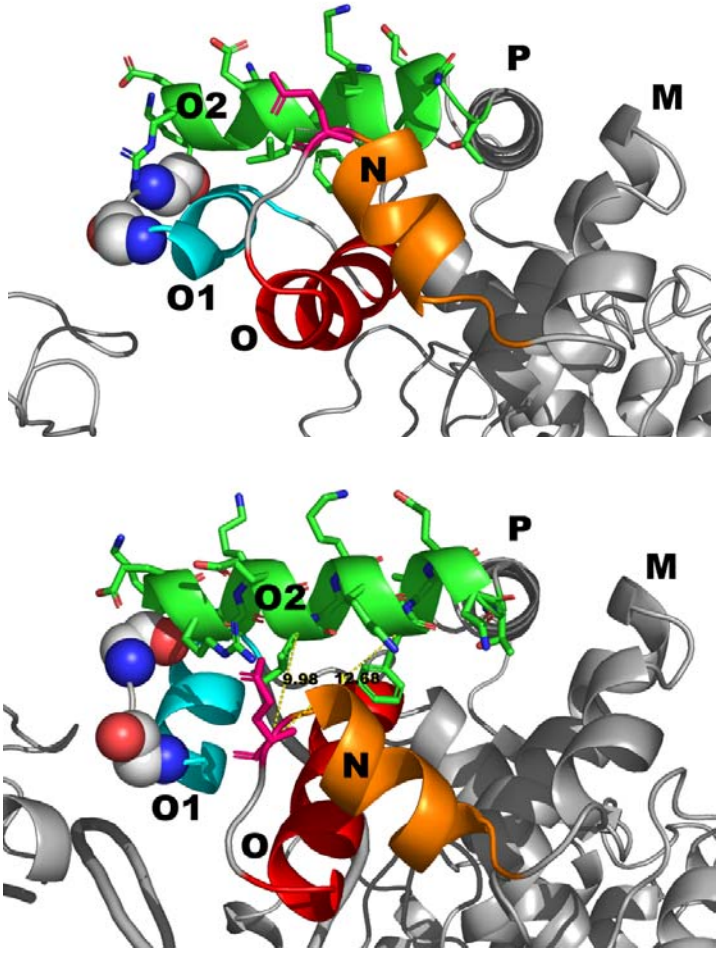


Figure 3.8: Comparison of the fingers domain of T7 DNA polymerase in the open and closed states

(Top) The back side of the fingers domain was drawn from an open complex (unpublished structure kindly provided by Dr. Tom Ellenberger) by using Pymol software. In this open complex, the helices are somewhat disordered and ill-defined; therefore, for the purpose of this illustration, segments of peptide in the N, O, O1, O2 helix were manually set to display as helix structures although they don't conform to standard alpha-helical geometry. (Bottom) The same view of the fingers domain drawn from the ternary complex of T7 DNA polymerase in its closed state (PDB file 1T7P).

Properties of the conformational change preceding chemistry

Besides the roles of the conformational change reported in the Tsai and Johnson paper (69), more features were revealed by the new data from the hinge mutant. First, the mutations of the hinge decreased the forward rate of the conformational change dramatically for the correct nucleotide and the decrease of the fidelity of the mutant is mainly from this effect. Although the rate of the conformational change for the incorrect nucleotides was also decreased (data not shown), it is still faster than the chemistry rate of the misincorporation, and thus no large effect on the specificity constant for misincorporation was observed. This result brought up a speculation that the fingers domain might have appeared later during evolution and it was evolved only to recognize and deliver correct nucleotide to the active site; originally, most DNA polymerases might be error prone enzymes which generate more mutations leading to the diversities of species.

The data from the processivity assay showed that the first incorporation of a correct nucleotide is much slower than the following incorporations. If the first incorporation is limited by the conformational change preceding chemistry, then this conformational change might have several stages and some earlier stage may be rate-limiting. Our data suggest that during processive nucleotide incorporation, after the first correct nucleotide is incorporated, the fingers domain does not fully open to the original resting state but partially open to bypass the slowest stage, thus allowing fast incorporation of the next correct nucleotides.

In summary, our data suggested that the conformational change reported by MDCC is a large and complex movement of the fingers domain, and it is closely related to enzyme catalysis. T7 DNA polymerase uses the induced-fit mechanism to regulate its specificity and fidelity. Evolutionally, the fingers domain might be a late add-on for recognizing and delivering the correct nucleotide to ensure the high fidelity of DNA replication.

Chapter 4: The Roles of the Catalytic Residues on the Recognition Domain during Nucleotide Binding and Incorporation

4.1 INTRODUCTION

To understand the molecular basis governing the high fidelity of T7 DNA polymerase, we have examined the roles of the residues around the active site of this enzyme by site-directed mutagenesis and transient kinetic analysis. In particular, the roles of three residues, K522, H506, and R518, on the fingers domain shown in Figure 4.1 will be reported in this chapter.

Using the MDCC tag on the fingers domain, the kinetics of the conformational change has been directly measured for the T7 cys-light mutant. Based on these new measurements, the pathway of nucleotide incorporation with two binding steps preceding chemistry was defined for correct and incorrect nucleotides (Scheme 4.1) (69).

Correct nucleotide - dCTP

$$E{\cdot}D_n + dCTP \underset{}{\overset{28\ \mu M}{\rightleftharpoons}} ED_n{\cdot}dCTP \underset{1.6\ s^{-1}}{\overset{660\ s^{-1}}{\rightleftharpoons}} FD_n{\cdot}dCTP \ (\textit{aligned}) \overset{360\ s^{-1}}{\rightleftharpoons} FD_{n+1}{\cdot}PP_i$$

Incorrect nucleotide - dGTP

$$E{\cdot}D_n + dGTP \underset{}{\overset{200\ \mu M}{\rightleftharpoons}} ED_n{\cdot}dGTP \underset{420\ s^{-1}}{\overset{220\ s^{-1}}{\rightleftharpoons}} HD_n{\cdot}dGTP \ (\textit{misaligned}) \overset{0.3\ s^{-1}}{\rightleftharpoons} HD_{n+1}{\cdot}PP_i$$

Scheme 4.1: Pathway of nucleotide incorporation of T7 DNA polymerase

Based on these models, we can study the molecular basis underlying nucleotide selectivity in a more detailed manner than before. For example, if a given mutation of a

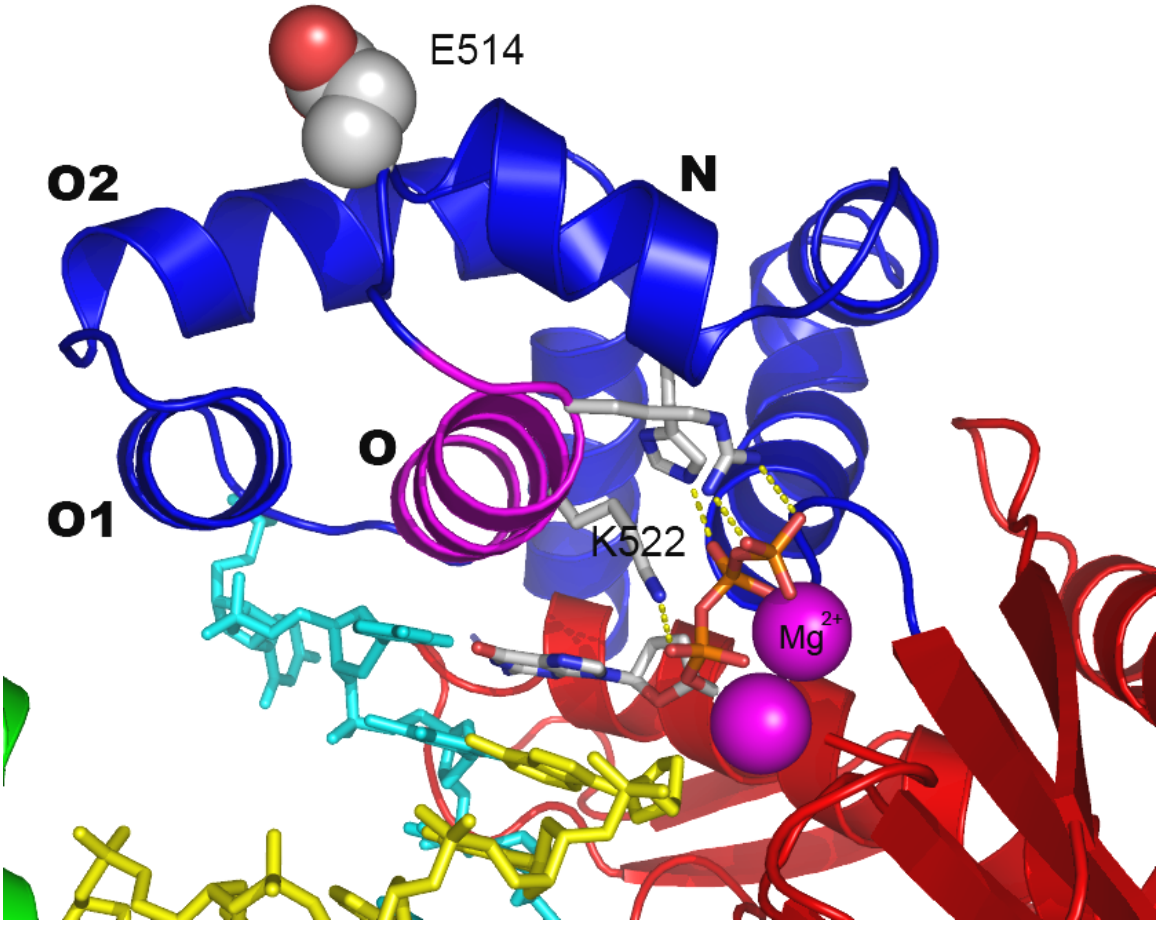


Figure 4.1: Catalytic residues on the fingers domain

Residues K522, H506, and R518 provide hydrogen bonds to the α-, β-, and γ-phosphates of the incoming dNTP (in CPK colors). Two metal ions are shown in magenta color. The O helix is also in magenta. The MDCC labeling site is shown with the E514 in stead (spheres in CPK colors). From structure file 1T7P.

catalytic residue reduces the net rate of the polymerization, we can determine if the effect is due to slowing the rate of the conformational change or the chemistry step. We even could estimate if the reduced rate of chemistry is due to the destabilization of the transition state or a misalignment of the recognition domain after the conformational change, based on the observation that a rise in fluorescence after nucleotide binding is correlated with misalignment of the recognition domain.

As shown in the closed form of the crystal structure of the T7 DNA polymerase ternary complex, three catalytic amino acids, K522, H506, and R518 on the fingers domain, form hydrogen bonds with the non-bridging oxygen atoms of the α-, β-, γ-phosphates of the incoming nucleotide, respectively (14). These amino acids may orient the α-phosphate for the phosphoryl transfer reaction, and help to neutralize the developing negative charge on the oxygen during the transition state. Mutations of the corresponding residues in the Klenow fragment of DNA Pol I to alanine caused a significant decrease in the rate of catalysis (77), but little was known as to whether the decrease in the rate of catalysis was due to the changes in conformational change, the chemistry step or both. With the newly developed methods, we can measure the rate of the conformational change directly by monitoring the fluorescence change. It is important to note that the MDCC label is linked to the R518 only by a three amino acid loop, and therefore, changes in fluorescence may reflect the positioning of this active site residue, as well as others on or near the O-helix. In this study, we analyzed the roles of these residues by site-directed mutagenesis and transient kinetic methods.

A set of mutations with varying severity to catalysis was made to more accurately assess the roles of these residues in the conformational change and chemistry. K522 and H506 are hydrogen bonded to α- and β- phosphates respectively, and therefore K522H and H506K were made to probe the importance of charge and precise geometry in catalysis. K522A and H506A were made to test the effect of entire elimination of the positive charge. Since R518 makes two hydrogen bonds with the γ-phosphate of the incoming nucleotide, R518K was made to test the

effect of a weaker hydrogen bond and R518A was made to test the effect of removing the positive charge.

The goal of the kinetic analyses on these mutants was to build the pathway of nucleotide incorporation for each mutant, from which we could establish the effect of a specific mutation. In this study, the nucleotide binding and incorporation pathway of each mutant was examined by chemical quench and stopped-flow assays, and the data were analyzed globally to obtain the pathway for each mutant.

Mutagenesis studies on these residues may help to better assess the thio-elemental effect of nucleotide incorporation. Thus the binding and incorporation of the thio substituted nucleotide analogs were also studied.

4.2 MATERIALS AND METHODS

Construction, expression, purification, and fluorescently labeling of the fingers domain mutants

The recognition domain mutants K522A, K522H, H506A, H506K, R518A and R518K were constructed from a plasmid encoding T7 DNA polymerase cys-light mutant (C20S-C88A-C275A-C313A-C451S-C660A-C688A-C703A-E514C) by using site-directed mutagenesis method. The mutant proteins were expressed, purified, and fluorescently labeled as described in Chapter Five. Compare to the procedure used in Chapter Three, the major change in this procedure is the addition of a mono-S column, which further cleans the final protein samples.

DNA substrates for kinetic studies

DNA oligomers were synthesized by Integrated DNA Technologies (IDT, Inc.) and purified by using 15% polyacrylamide/7 M urea denaturing gel electrophoresis in our lab. A 27-mer primer (5'GCC TCG CAG CCG TCC AAC CAA CTC AAC3') and a 45-mer template (5'GGA CGG CAT TGG ATC GAG GTT GAG TTG GTT GGA CGG CTG CGA GGC3') were adopted from a previous study to allow comparison between all our studies (69).

There were two versions of the 27-mer primer. For nucleotide binding studies only, a ddCMP-terminated 27-mer primer was used to prevent incorporation of the next incoming nucleotide, dCTP. The dCMP terminated primer was used for single nucleotide incorporation experiments and it was 5'-^{32}P-labeled using T4 polynucleotide kinase according to the manufacturer's instructions (Invitrogen). The labeling reaction was terminated by heating at 95 ºC for 5 min and excess ^{32}P-labeled nucleotide was removed by using a biospin 6 column (Bio-Rad).

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

Chemical quench experiments

Single nucleotide incorporation assays were performed at 20 ºC in a *reaction buffer* containing 40 mM Tris (pH 7.5), 1mM EDTA, 50 mM NaCl, 1 mM DTT, 12.5 mM $MgCl_2$. For dCTP and dCTPαS incorporation assays, a KinTek RQF-3 Rapid Quench Flow apparatus was used. One sample loop was loaded with preformed enzyme-DNA complex (600 nM enzyme, 12 μM thioredoxin and 200 nM 5'-^{32}P-labeled 27-mer/45-mer DNA duplex) in the *reaction buffer* without $MgCl_2$. Another loop was loaded with nucleotide in the *reaction buffer* with 25 mM $MgCl_2$. The reactions were started by rapidly mixing the two reactants and then were quenched by mixing with 500 mM EDTA after various time intervals ranging from several milliseconds to several seconds.

The products from the assays described above were resolved on a 15 % denaturing polyacrylamide sequencing gel and then the dried gel was exposed to a phosphor screen. The screen was scanned using a Storm 860 scanner (GE Healthcare) and the amount of product formation at each time point was analyzed by the ImageQuant software (GE Healthcare). The concentration of the product was calculated from the concentration of the starting isotope labeled DNA primer times the fraction of material in the product band.

Stopped-flow experiments

All experiments were performed at 20 ºC in *reaction buffer* using a SF2004 stopped-flow apparatus (KinTek Corp.). The excitation wavelength was set at 425 nm and a photomultiplier with a 450 nm cutoff high pass filter was used to observe fluorescence. For nucleotide incorporation stopped-flow assays, one syringe was loaded with an enzyme-DNA complex (200 nM enzyme, 8 μM thioredoxin, and 600 nM DNA with dCMP terminated 27-mer primer) in the *reaction buffer* and another syringe was loaded with the next incoming nucleotide at various concentrations in the *reaction buffer*. Aliquots of these two syringes were rapidly mixed and the fluorescence emission was recorded. For nucleotide binding assays, 600 nM DNA with ddCMP terminated 27-mer primer was used instead of normal primers.

Traditional data fitting and global data fitting

The data were analyzed by nonlinear regression using GraFit 5 software (Erithacus Software Limited). The time course of single nucleotide incorporation at each nucleotide concentration was fit to a single exponential equation, $y = A \cdot \exp(-k_{obs}\, t) + C$, where A is the amplitude, k_{obs} is the observed rate, and C is the endpoint. The concentration dependence of the rate of incorporation was fit to a hyperbola, $k_{obs} = k_{pol} \cdot S_0/(K_d+S_0)$, where k_{obs} is the observed rate, k_{pol} is maximum incorporation rate, S_0 is variable starting nucleotide concentration, and K_d is apparent ground-state binding constant. The data from the stopped-flow experiments were best fit to a burst equation, $y = A \cdot \exp(-k_{bst}\, t) + k_{ss} t + C$, where k_{bst} is the burst rate and k_{ss} accounts for a slow linear increase in fluorescence. The concentration dependence of the rate of the fluorescence change was fit to a hyperbola, $k_{obs} = k_2 \cdot S_0/(K_d+S_0)+k_{-2}$, where k_{obs} is the observed rate, k_2 is the forward rate, k_{-2} is the reverse rate, S_0 is the variable starting nucleotide concentration, and K_d is apparent ground-state binding constant.

The global data fitting was performed by using the KinTek Explorer software (KinTek Corp.), in which both quench flow and stopped-flow data were fitted simultaneously. The rate

constants from the conventional data fitting described above were used as initial estimate. The process of global data fitting involved trial, refinement and a final convergence to the best fit as describe in the user manual.

4.3 RESULTS

The kinetics of single nucleotide incorporation measured by quench-flow experiments

Correct nucleotide incorporation was measured by rapid quench-flow experiments. Pre-formed enzyme-DNA complex was rapidly mixed with a solution containing correct incoming nucleotide (dCTP) at various concentrations. Figure 4.2A shows the time courses of product (28-mer) formation observed with the H506K mutant. The time courses of product formation were fit to a single exponential function to obtain the rates at various dCTP concentrations. The rates were then plotted as a function of dCTP concentration (Figure 4.2B) and fit to a hyperbola, yielding a maximum nucleotide incorporation rate, k_{pol} = 0.35 ± 0.02 s^{-1}, and an apparent nucleotide dissociation constant, K_d = 32 ± 7 μM. Results of correct nucleotide incorporation for all the mutants are summarized in Table 4.1.

Table 4.1: Results of dCTP incorporation from the quench-flow assays

Mutant	K_d (μM)	k_{pol} (s^{-1})
K522A	184±57	0.39±0.05
K522H	180±30	1.1±0.1
H506A	66±6	1.12±0.03
H506K	32±7	0.35±0.02
R518A	86±25	5.9± 0.8
R518K	75±22	43±5

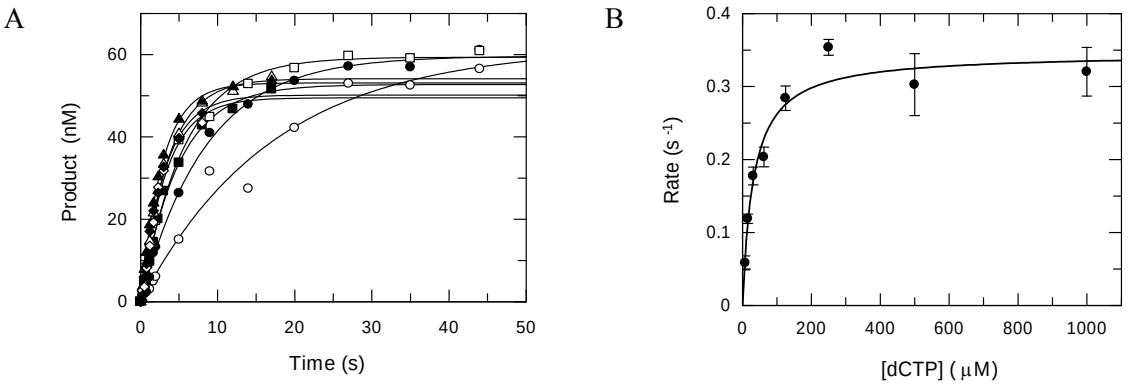


Figure 4.2: dCTP concentration dependence of the rate of singe nucleotide incorporation for H506K mutant

(A) A preformed enzyme-DNA complex (600 nM enzyme and 200 nM DNA) was rapidly mixed (1:1) with dCTP solution at various concentrations to start the experiments. The time course of product formation at each dCTP concentration (from bottom to top, the series of data points are for reactions with 7.8, 15.6, 31.2, 62.5, 125, 250, 500 and 1000 μM dCTP solutions, respectively) was fit to a single exponential to obtain the incorporation rate. The continuous lines are the fitted lines. (B) The dCTP concentration dependence of the incorporation rate was fit to a hyperbola, defining an apparent K_d of 32 ± 7 μM, and a maximal incorporation rate, k_{pol}, of 0.35 ± 0.02 s^{-1}.

Fluorescence spectra of the mutant enzymes with DNA_{dd} and nucleotide

To study the conformational change induced by nucleotide binding for the mutant enzymes, significant fluorescence changes of the MDCC labeled enzymes upon binding nucleotides were needed. This goal was already accomplished with the "wild type" cys-light enzyme (69). Here the fluorescence emission spectra of the MDCC labeled mutant enzymes upon binding its substrates were also characterized.

Figure 4.3 depicts the fluorescence spectra of the MDCC labeled H506K mutant obtained upon binding DNA, Mg^{2+}, and correct (dCTP) or incorrect (dGTP) nucleotides. Mutants such as H506A, K522A, and K522H all have very similar results, so the data were represented by the results of the H506K mutant in Figure 4.3. Similar to the "wild type" enzyme, binding of duplex DNA with a ddC-terminated primer (D_{dd}) and magnesium to the mutant enzyme (E) only caused slight changes in the fluorescence emission. Surprisingly, upon binding a correctly base paired nucleotide, dCTP, there was a significant fluorescence increase. This observation is reminiscent of the fluorescence increase upon binding a mismatched nucleotide to the "wild type" enzyme, suggesting a misaligned conformation could be formed after the binding of a correct nucleotide to this H506A mutant. If so, this would imply that H506 plays an important role in nucleotide alignment. When dCTP was replaced with mismatched nucleotide, dGTP, a smaller fluorescence increase was observed, in contrast to a large increase observed for the "wild type" enzyme. In summary, the results showed that the binding of dCTP to the enzyme-DNA_{dd} complex induced a large fluorescence increase which may correlate to the conformational change of the enzyme from an open state to a misaligned state preceding chemistry. This fluorescence signal was used for stopped-flow experiments to study the kinetics of the conformational change induced by nucleotide binding.

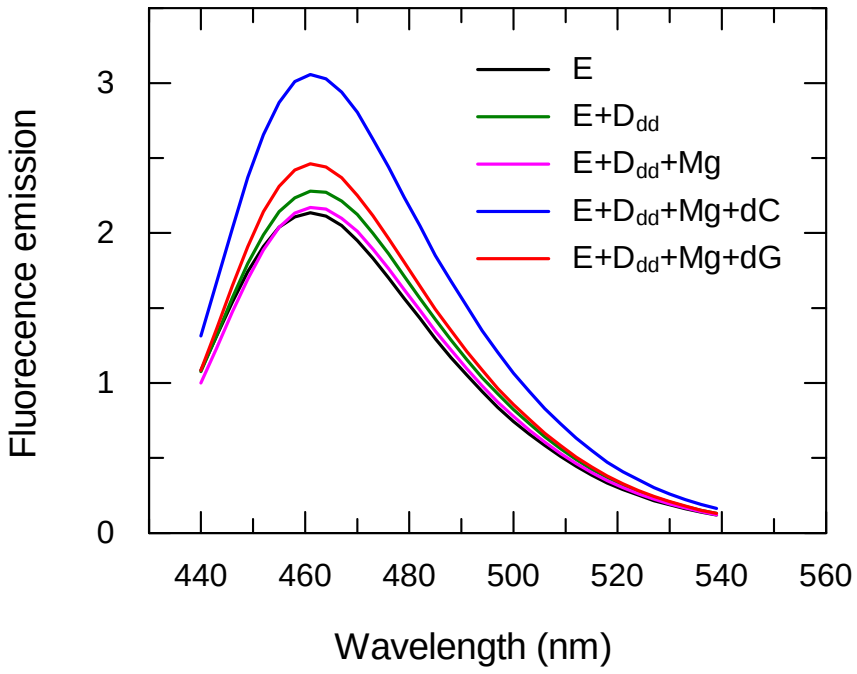


Figure 4.3: Fluorescence spectra of the H506K mutant with DNA_{dd} and nucleotide

Fluorescence emission scans of 0.2 μM enzyme were recorded after the addition of the substrates in the order of 0.3 μM DNA_{dd} (a DNA with dideoxy terminated primer), 12.5 mM Mg^{2+} and 0.5 mM dCTP or 1 mM dGTP.

Kinetics of nucleotide binding to E-DNA$_{dd}$ complex

The kinetics of the conformational change induced by nucleotide binding to the H506K mutant were characterized by a stopped-flow experiment in which DNA duplex with a ddCMP-terminated primer was used to block the chemistry step. As depicted in Figure 4.4A, for each reaction, an exponential increase of the fluorescence was observed after the enzyme-DNA$_{dd}$ complex was rapidly mixed with a solution containing the correct nucleotide, dCTP. The pattern of the fluorescence traces from the reactions at various dCTP concentrations implies a readily reversible conformational change. Each trace was fit to a burst equation to extract the rate and the amplitude of the rapid rise of the fluorescence. The dCTP concentration dependence of the rate was fit to a hyperbola, defining a K_d of 182 ± 49 μM, a k_2 of 9.7 ± 0.8 s^{-1}, and a k_{-2} of 1.4 ± 0.3 s^{-1} (Figure 4.4B). Since the rate of the fluorescence change approached a maximum at saturating dCTP concentration, the data indicated a two-step binding mechanism. During ground-state binding, dCTP rapidly reaches equilibrium with a K_d of 182 μM, then the enzyme undergoes a conformational change with a forward rate of 9.7 s^{-1} and a reverse rate of 1.4 s^{-1}. By calculation, the overall K_d for dCTP binding after these two steps is 182/(1+9.7/1.4) ≈ 23 μM. As an additional check, the dCTP concentration dependence of the amplitude was also fit to a hyperbola, yielding an apparent K_d of 13 ± 3 μM and a maximum amplitude of 0.287 ± 0.016 (Figure 4.4C). This apparent K_d is an estimate of the overall K_d of dCTP binding according to the two-step binding mechanism, and the value is close to the calculated overall K_d. Scheme 4.2 depicts the two-step nucleotide binding mechanism of the H506K mutant based on these data. In this Scheme, E stands for enzyme, D_{dd} is the DNA with a ddC-terminated primer, N is for nucleotide and F is the enzyme after the conformational change.

$$ED_{dd} + N \underset{}{\overset{182\,\mu M}{\rightleftharpoons}} ED_{dd}N \underset{1.4 s^{-1}}{\overset{9.7 s^{-1}}{\rightleftharpoons}} FD_{dd}N$$

Scheme 4.2: Two-step nucleotide binding mechanism for the H506K mutant

The observed rate of nucleotide incorporation for the H506 mutant is 0.35 s^{-1}, which is slower than the forward and reverse rates of the conformational change; therefore, we propose the observed slow rate of incorporation defines the rate of the chemistry step.

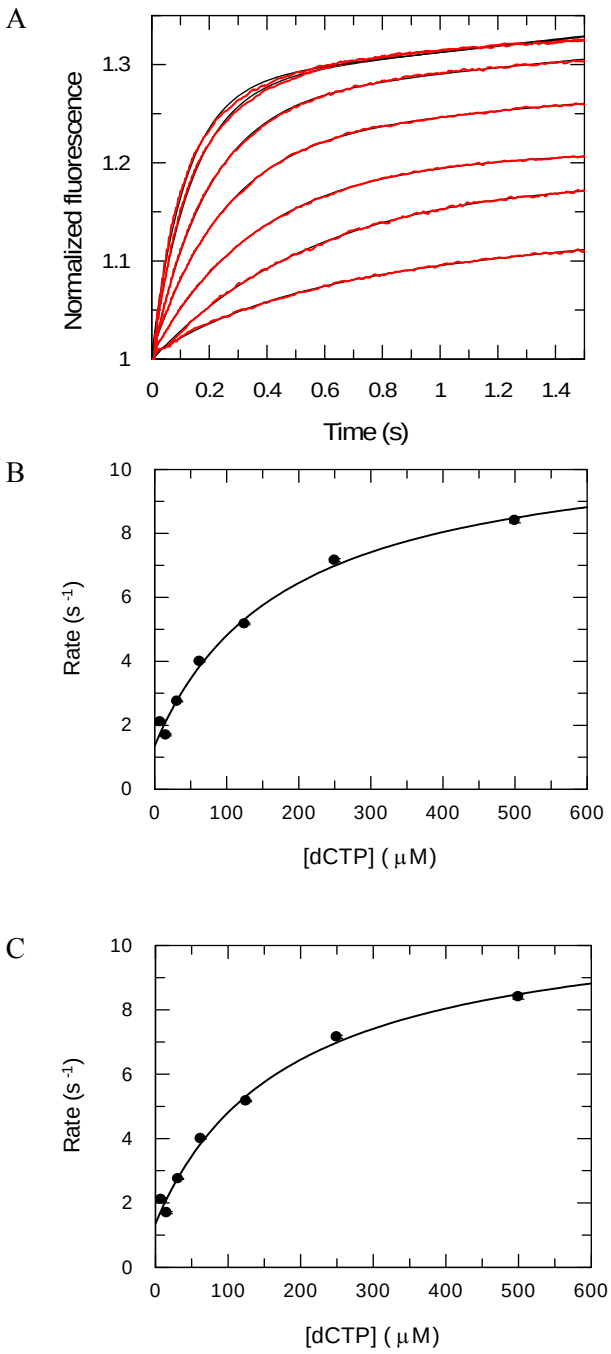


Figure 4.4: Conformational change induced by dCTP binding to the H506K mutant and DNA with dideoxy terminated primer

Figure 4.4: Conformational change induced by dCTP binding to the H506K mutant and DNA with dideoxy terminated primer.

(A) A preformed enzyme-DNA complex (200 nM enzyme and 600 nM DNA_{ddC}) was rapidly mixed (1:1) with dCTP at various concentrations in a stopped-flow instrument and the time courses of fluorescence changes were recorded (from bottom to top, the red traces are for reactions with 7.8, 15.6, 31.3, 62.5, 125, 250, and 500 μM dCTP solutions, respectively). Each trace was fit to a burst equation to extract the burst rate and the amplitude of the fluorescence change (the fitted lines are in black). (B) The fitted burst rates were plotted against dCTP concentrations and the data were fit to a hyperbola, yielding a K_d of 182 ± 49 μM, k_2 of 9.7 ± 0.8 s^{-1}, k_{-2} of 1.4 ± 0.3 s^{-1}. (C) The fitted amplitudes were plotted against dCTP concentrations and the data were fit to a hyperbola, yielding an apparent K_d of 13 ± 3 μM and a maximum amplitude of 0.287 ± 0.016.

The kinetics of single nucleotide incorporation measured by stopped-flow experiments

To obtain more information about the conformational change and the chemistry steps, we characterized the kinetics of single nucleotide incorporation by stopped-flow experiments. In these experiments, normal DNA was used so nucleotide could be incorporated while the fluorescence changes of the enzyme could also be monitored.

Figure 4.5 shows the results of dCTP incorporation for the H506K mutant. In each reaction, after the enzyme-DNA complex was rapidly mixed with a solution containing the correct nucleotide, dCTP, an increase in fluorescence was observed. As shown in Figure 4.5A, each fluorescence trace was fit to a burst equation to extract the rate and the amplitude of the fast rise of the fluorescence change. The dCTP concentration dependence of the rate was fit to a hyperbola, defining a K_d of 406 ± 93 μM, a k_2 of 10 ± 1 s^{-1}, and a k_{-2} of 1.2 ± 0.2 s^{-1} (Figure 4.5B). These values were very close to the kinetic constants observed from the study of dCTP binding to the E-D_{dd} complex, strongly suggesting that the fast fluorescence rise observed here is correlated with the conformational change preceding chemistry. The dCTP concentration dependence of the amplitude was also analyzed by fitting the data to a hyperbola, yielding an apparent K_d of 29 ± 3 μM and a maximum amplitude of 0.354± 0.008 (Figure 4.5C). This apparent K_d is close to the calculated K_d of 406/(1+10/1.2) ≈ 44 μM, suggesting the mechanism underlying the observed fluorescence change can be approximated by a two-step nucleotide binding mechanism because the step after the conformational change (chemistry) is slow enough to be negligible. The pathway of nucleotide incorporation for this mutant is a three-step mechanism as shown in Scheme 4.1 for the "wild type" enzyme. The observed nucleotide incorporation rate measured from the quench-flow experiment is 0.35 ± 0.02 s^{-1}, which is a function of the rates of the steps preceding and including the chemistry step. At saturating nucleotide concentration, this observed rate approximates to $k_2k_3/(k_2+k_{-2}+k_3)$, and when $(k_2+k_{-2}) \gg k_3$, it reduces to $k_2k_3/(k_2+k_{-2})$. More

intuitively, because of the reversible conformational change, about $k_2/(k_2+k_{-2})$ = 10/(10+1.2) *100% ≈ 90% of the enzyme is in the productive state to form product, and thus the rate of the chemistry step (k_3) is about 0.35/0.9≈ 0.39 ± 0.06 s^{-1}. Based on all these estimated kinetic parameters, a three-step reaction pathway of nucleotide incorporation for this H506K mutant is proposed as shown in Scheme 4.3. As a confirmation, the K_d measured from the quench-flow experiment is 32 ± 7 μM, which is close to the calculated K_d of $K_{d,1}(k_{-2}+k_3)/(k_2+k_{-2})$ ≈ 57 ± 14 μM.

$$ED_n + N \underset{}{\overset{406\,\mu M}{\rightleftarrows}} ED_nN \underset{1.2\,s^{-1}}{\overset{10\,s^{-1}}{\rightleftarrows}} FD_nN \overset{0.39\,s^{-1}}{\rightleftarrows} FD_{n+1}PP_i$$

Scheme 4.3: Pathway of dCTP incorporation for the H506K mutant from conventional data fitting

The kinetics of nucleotide incorporation measured by stopped-flow assays for the H506K mutant is representative of the kinetics of other tested mutants and thus the detailed results and data analysis of other mutants are not shown here. Table 4.2 summarizes the results. Combined with other data presented in previous sections, the three-step pathways of correct nucleotide incorporation for these mutants were also proposed and the kinetic parameters for each step were shown in Table 4.3.

Table 4.2: Data analysis results for correct nucleotide incorporation stopped-flow assays

Mutant	K_d (μM)	k_2 (s^{-1})	k_{-2} (s^{-1})	$K_{d,overall}$ (μM)	Amplitude
K522A	233±126	47±6	56±4	143±9	0.247±0.005
K522H	164±155	102±23	48±31	147±29	0.20±0.01
H506A	113±26	9.7±0.5	4.2±0.3	25.7±1.4	0.237±0.003
H506K	406±93	10.2±1.0	1.2±0.2	29±3	0.354±0.008

Table 4.3: Kinetic constants proposed for the three-step pathway of correct nucleotide incorporation of the mutants

Mutant	K_d (μM)	k_2 (s^{-1})	k_{-2} (s^{-1})	k_3 (s^{-1})
K522A	233±126	47±6	56±4	0.86±0.17
K522H	459±234	102±23	48±31	1.6±0.6
H506A	113±26	9.7±0.5	4.2±0.3	1.6±0.1
H506K	406±93	10.2±1.0	1.2±0.2	0.39±0.06

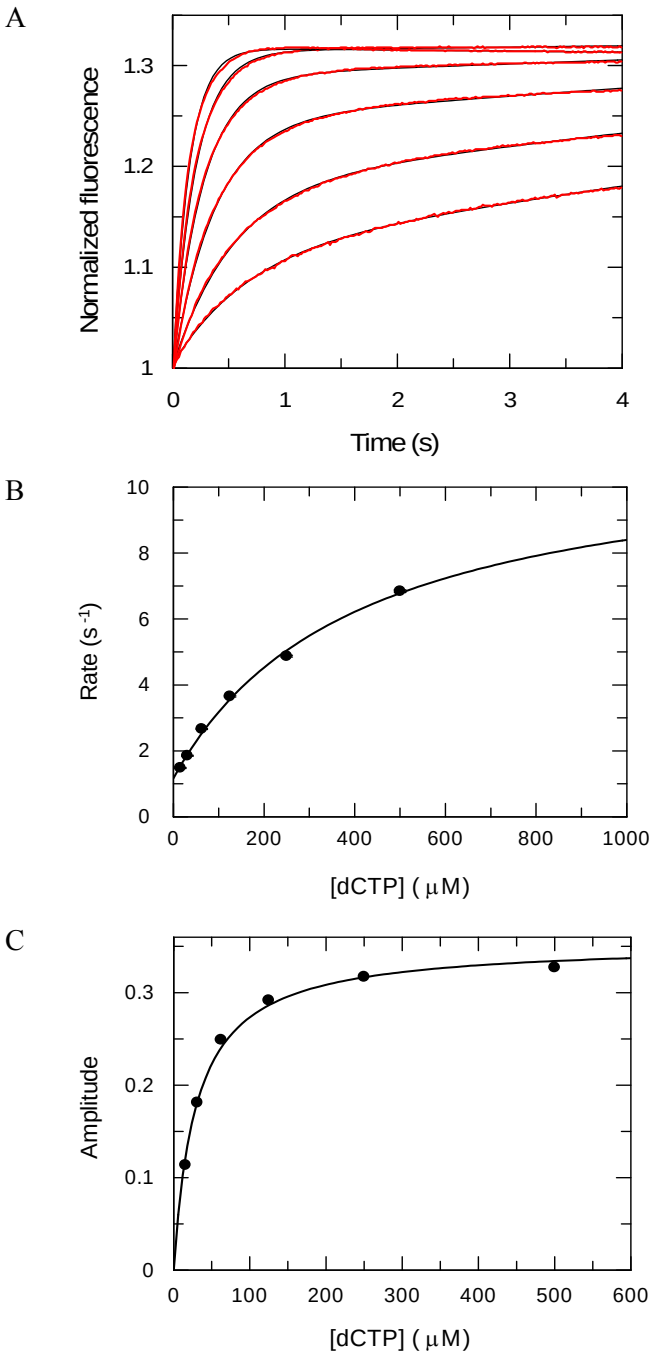


Figure 4.5: Conformational change induced by dCTP binding to the H506K mutant and normal DNA.

Figure 4.5: Conformational change induced by dCTP binding to the H506K mutant and normal DNA.

(A) A preformed enzyme-DNA complex (200 nM enzyme and 600 nM DNA_{dC}) was rapidly mixed (1:1) with dCTP solution at various concentrations in a stopped-flow instrument and the time courses of fluorescence changes were recorded (from bottom to top, the red traces are for reactions with 15.6, 31.2, 62.5, 125, 250, and 500 μM dCTP solutions, respectively). Each trace was fit to a burst equation to extract the burst rate and the amplitude of the fluorescence change (the fitted lines are in black). (B) The fitted burst rates were plotted against dCTP concentrations and the data were fit to a hyperbola, yielding a K_d of 406 ± 93 μM, k_2 of 10.2 ± 1.0 s^{-1}, k_{-2} of 1.2 ± 0.2 s^{-1}. (C) The fitted amplitudes were plotted against dCTP concentrations and the data were fit to a hyperbola, yielding an apparent K_d of 29 ± 3 μM and a maximum amplitude of 0.354 ± 0.008.

Fluorescence spectra of the mutant enzymes with normal DNA and nucleotide incorporation

One surprise from the stopped-flow assays of correct nucleotide incorporation was that only the fluorescence increase was observed while the expected fluorescence decrease after nucleotide incorporation was not seen. During processive nucleotide incorporation, DNA polymerases are thought to return to the open state after each cycle of polymerization, and thus a fluorescence decrease following the initial rise was expected. To understand why the fluorescence did not return to the lower level, we studied the fluorescence spectra of the mutant enzymes before and after several sequential correct nucleotide incorporation events in a fluorimeter. The spectra shown in Figure 4.6B demonstrates that after incorporation of one nucleotide, the fluorescence stayed at a higher level; however, after incorporation of three nucleotides, the fluorescence returned to the lower level. One interpretation could be that the MDCC fluorescent tag is sensitive enough to discriminate the subtle difference in the DNA sequence bound at the active site, so the fluorescent state of the labeled mutant enzymes are different depending on the sequence of the DNA. This hypothesis was supported by the results of another experiment designed to compare the fluorescence states of the mutant enzymes upon binding two different DNA duplexes. In this experiment, the bound DNA with a 28-mer primer was replaced by a DNA with a 27-mer primer added in large excess. As shown in Figure 4.6C, the fluorescence level of the enzyme decreased after addition of excess DNAs with 27-mer primer. More interestingly, after addition of 1mM dCTP to the reaction to incorporate one nucleotide, the fluorescence level returned back to a higher level. These results demonstrated that the fluorescence of the enzyme shows different levels when bound with different DNAs; in our assays, the fluorescence level of the enzyme is higher when it is bound with a DNA with a 28-mer primer. This provides a rationale for why the fluorescence of the enzyme did not return to the lower lever after incorporation of a dCTP onto the 27-mer primer in the stopped-flow assays. Moreover, this

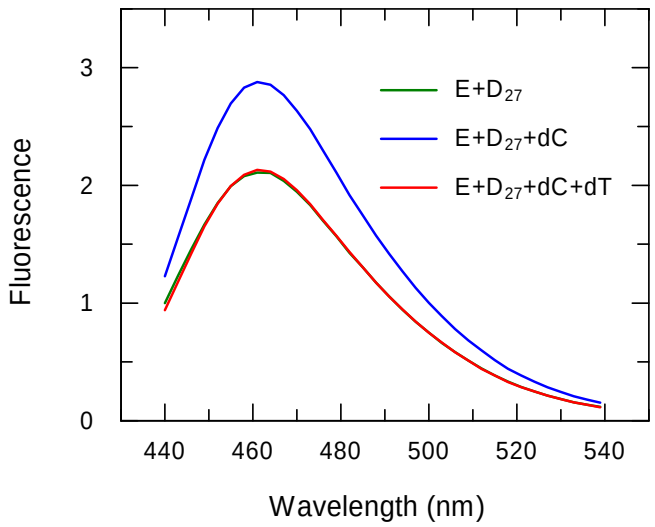
information was very useful for the global data fitting since the fluorescence level of the enzyme at different states should be defined during the fitting process.

A

D_{27}: 5’ GCC TCG CAG CCG TCC AAC CAA CTC AAC
3’ CGG AGC GTC GGC AGG TTG GTT GAG TTG GAG CGA GGT TAC GGC AGG 5’

D_{28}: 5’ GCC TCG CAG CCG TCC AAC CAA CTC AAC C
3’ CGG AGC GTC GGC AGG TTG GTT GAG TTG GAG CGA GGT TAC GGC AGG 5’

B



C

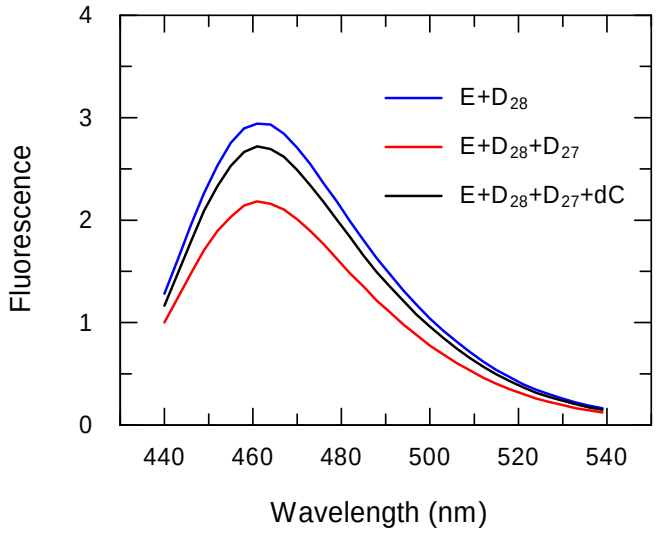


Figure 4.6: Fluorescence spectra of the H506K mutant with normal DNA and nucleotide

Figure 4.6: Fluorescence spectra of the H506K mutant with normal DNA and nucleotide.

(A) DNA substrates used in the assays. (B) Fluorescence emission scans of 0.2 μM enzyme were recorded after addition of the substrates in the order of 0.3 μM DNA_{27} (a DNA with a 27-mer primer), 0.5 mM dCTP and 0.5 mM dTTP. (C) Fluorescence emission scans of 0.2 μM enzyme were recorded after addition of the substrates in the order of 0.3 μM DNA_{28}, 2.3 μM DNA_{27} and 1 mM dCTP.

Global data fitting

Global data fitting has several advantages over the conventional data fitting. First, in global data fitting, all the data from different experiments are fit simultaneously to a unique mechanism; however, in conventional fitting, the data from different experiments are fit separately to equations derived from simplified mechanisms, and then the fitted results are further analyzed to derive the rate constants governing each step. In this way, error is propagated and a large number of extraneous parameters are included in the process of data fitting. Second, all aspects of the data including the reaction amplitudes are used to obtain intrinsic constants in global data fitting, so the fitted constants are more constrained by the data and thus are more accurate. Third, by using global data fitting, more complicated data can be analyzed; in contrast, it is hard for conventional data fitting to analyze a data set that are best fit to equations beyond single exponential. For example, when the data are best fit to a double exponential during conventional data fitting, the two phases sometimes are not well resolved.

However, conventional data fitting has its own merits, because patterns indicative of a particular mechanism can be revealed. Each kinetic experiment is designed to probe the rate constants governing a kinetic mechanism and can be fit conventionally by the equations derived from this mechanism. Thus conventional data fitting can generate an estimate of the rate constants for each step with the certainty that these rates are constrained by the data from a specific experiment. During global data fitting, the rate constants obtained from the conventional data fitting can be used as initial estimates to guide the fitting direction. After several refinements, a convergence to a final mechanism representing the best fit to all experimental data can be obtained.

Here, the global data fitting of the H506K mutant will be described in detail. Both the stopped-flow and the quench flow data of the H506K mutant were fit simultaneously. The rate

constants from the conventional data fitting shown in Scheme 4.3 were applied to this mechanism and for the extra steps such as DNA binding and dissociation the rate constants were from previously measured rates for T7 DNA polymerase (2). The best fit is shown in Figure 4.7 with the mechanism shown in Scheme 4.4. To fit the ground-state dissociation constant, we set the binding rate to 100 $\mu M^{-1}s^{-1}$ and allowed the dissociation rate to float during fitting to derive a value of 26000 s^{-1} to define the K_d, 260 μM, for the rapid equilibrium.

$$E + D_n \underset{0.2 s^{-1}}{\overset{10 \mu M^{-1} s^{-1}}{\rightleftharpoons}} ED_n$$

$$ED_n + N \overset{260 \mu M}{\rightleftharpoons} ED_n N \underset{0.7 s^{-1}}{\overset{10 s^{-1}}{\rightleftharpoons}} FD_n N \overset{0.36 s^{-1}}{\rightleftharpoons} ED_{n+1} + PP_i$$

$$ED_{n+1} \underset{10 \mu M^{-1} s^{-1}}{\overset{0.2 s^{-1}}{\rightleftharpoons}} E + D_{n+1}$$

Scheme 4.4: dCTP incorporation reaction pathway for H506K mutant from global data fitting

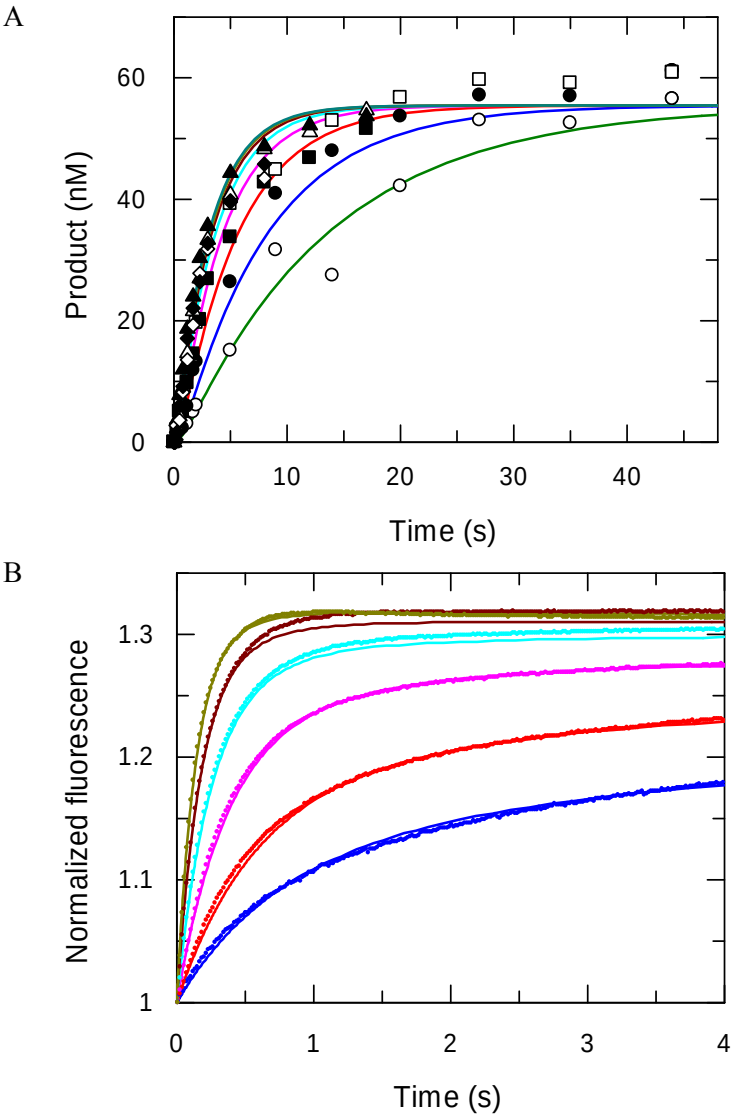


Figure 4.7: Global data fitting of dCTP incorporation for H506K mutant

(A) The series of data points were from quench flow experiments described in Figure 4.2A. The colored lines show results of fitting to the mechanism shown in Scheme 4.4. (B) The series of data points represented the fluorescence changes from the stopped-flow experiments described in Figure 4.5A. The solid lines show results of fitting to the mechanism shown in Scheme 4.4.

As shown in Scheme 4.4, the ground-state binding of the correct nucleotide with a K_d of about 260 μM is followed by one isomerization preceding chemistry. The forward rate of this isomerization is reduced about 660/10 = 66 fold from the rate of the wild type enzyme while the reverse rate is close to the rate observed for the wild type enzyme. The chemistry rate is about 0.36 s^{-1} which could be considered as a rate-limiting step during processive replication for this mutant, assuming the reopenning of the fingers domain and pyrophosphate release steps are fast.

Global data fitting were performed on mutants K522A, K522H, H506A using the same routine as well and the results of the fitting are shown in Appendix and the rate constants are summarized in Table 4.4.

Table 4.4: Kinetic constants for dCTP incorporation from global data fitting

Mutant	K_d, dNTP (μM)	k_2 (s^{-1})	k_{-2} (s^{-1})	k_3 (s^{-1})
K522A	270	35	31	0.7
K522H	623	90	47	2
H506A	70	15	3.5	0.9
H506K	260	10	0.7	0.36

The thio-elemental effects on nucleotide incorporation and enzyme conformational change

The thio-elemental effect was defined by the rate of incorporating a normal dNTP divided by the rate of incorporating a corresponding α-thio-substituted dNTP analog with a non-bridging oxygen of the α-phosphate replaced by a sulfur (S_p stereoisomer) ($k_{pol,dNTP}/k_{pol,dNTP\alpha S}$). Since the sulfur group decreases the chemistry rate of phosphoryl transfer due to its decreased electronegativity compared to oxygen, it has been suggested that a full thio-elemental effect would be observed if the chemistry step is rate-limiting (2,39,41). It would be interesting to study the thio-elemental effect of these fingers domain mutants since the data analysis shown in Table 4.4 strongly suggested that they have a rate-limiting chemistry step during nucleotide incorporation.

We studied the dCTPαS incorporation by using the quench-flow method first. Figure 4.8 depicts results for the H506K mutant. The dCTPαS concentration dependence of the incorporation rates were fitted to a hyperbola, defining an apparent K_d of 42 ± 6 μM, and a maximal incorporation rate k_{pol} of 0.24 ± 0.01 s^{-1}. The same assays were done using other mutants and the results are shown in Table 4.5.

Table 4.5: Results of dCTPαS incorporation from the quench-flow assays

Mutant	K_d (μM)	k_{pol} (s^{-1})
K522A	152±17	0.106±0.004
K522H	175±25	0.80±0.04
H506A	66±13	0.37±0.02
H506K	42±6	0.24±0.01
R518A	n/d	n/d
R518K	n/d	n/d

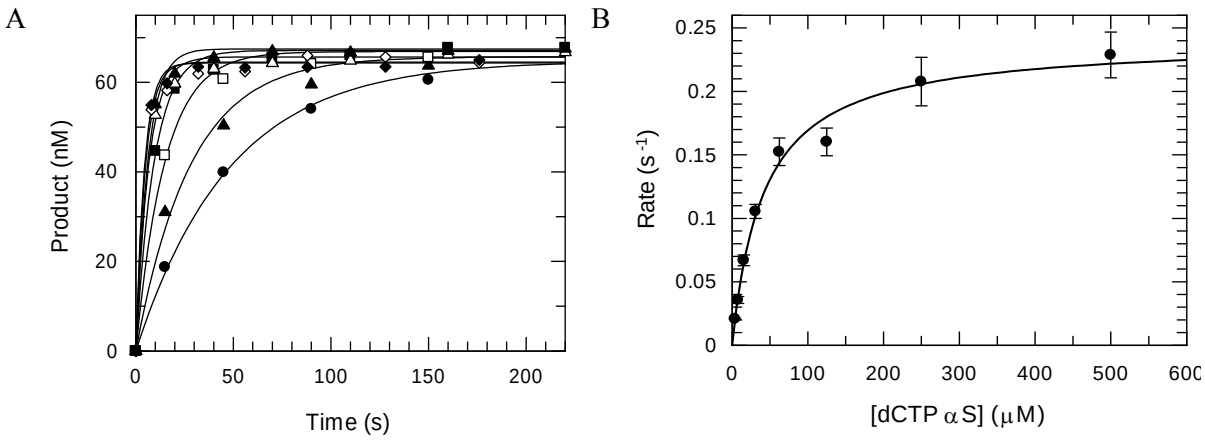


Figure 4.8: dCTPαS concentration dependence of the rate of singe nucleotide incorporation for the H506K mutant

(A) A preformed enzyme-DNA complex (600 nM enzyme and 200 nM DNA) was rapidly mixed (1:1) with a solution containing dCTPαS at various concentrations to start the experiments. The time course of product formation at each dCTPαS concentration (from bottom to top, the series of data points are for reactions with 3.9, 7.8, 15.6, 31.2, 62.5, 125, 250, and 500 μM dCTPαS concentrations, respectively) was fit to a single exponential to obtain the incorporation rate. The continuous lines are the fitted lines. (B) The dCTPαS concentration dependence of the incorporation rates was fit to a hyperbola, defining an apparent K_d of 42 ± 6 μM, and a maximal incorporation rate kpol of 0.24 ± 0.01 s^{-1}.

We also studied whether dCTPαS (S_p) affects the kinetics of the conformational change of this H506K mutant. Figure 4.9 shows the results of dCTPαS incorporation using the stopped-flow method. Each fluorescence trace in Figure 4.9A was fit to a burst equation to extract the amplitude and the rate of the fast increase of fluorescence. The dCTPαS concentration dependence of the burst rates was fit to equation, $k_{obs} = k_2[S]/([S] + K_d) + k_{-2}$, yielding an apparent K_d of 164 ± 14 μM, k_2 of 6.9 ± 0.1 s^{-1}, and k_{-2} of 2.3 ± 0.1 s^{-1} (Figure 4.9B). The dCTPαS concentration dependence of the amplitude was also fit to a hyperbola, yielding an apparent K_d of 32 ± 4 μM and a maximum amplitude of 0.36 ± 0.01 (Figure 4.9C). These data are comparable to the data from dCTP incorporation (Table 4.4.), suggesting the conformational change is not affected by the sulfur substitution. Since the conformational change step is readily reversible, only about 6.9/(6.9+2.3) *100% ≈ 75% of the enzyme is bound in the state ready for chemistry. By calculation, the rate of the chemistry step is about 0.24/0.75 ≈ 0.32 s^{-1}. Based on all these estimated kinetic parameters, a three-step reaction pathway of dCTPαS incorporation for this H506K mutant was proposed and shown in Scheme 4.5.

$$ED_n + N \xrightleftharpoons{164\mu M} ED_nN \underset{2.3s^{-1}}{\overset{6.9s^{-1}}{\rightleftharpoons}} FD_nN \xrightleftharpoons{0.32s^{-1}} FD_{n+1}PP_i$$

Scheme 4.5: Pathway of dCTPαS incorporation for H506K mutant from conventional data fitting

The kinetics of dCTPαS incorporation measured by stopped-flow assays for the other fingers domain mutants were also studied and the results are summarized in Table 4.6. Combined with quench-flow data, the three-step pathways of dCTPαS incorporation for these mutants were also proposed and the kinetic parameters for each step were shown in Table 4.7.

Table 4.6: Data analysis results for dCTPαS incorporation stopped-flow assays

Mutant	K_d (μM)	k_2 (s^{-1})	k_{-2} (s^{-1})	$K_{d,overall}$ (μM)	Amplitude
K522A	102±104	88±16	51±23	46±12	0.104±0.007
K522H	243±96	81±3.9	114±16	66±8	0.182±0.007
H506A	n.d.	14-28	14-28	13±6	0.135±0.012
H506K	164±14	6.9±0.1	2.3±0.1	32±4	0.36±0.01

Table 4.7: Kinetic constants proposed for the three-step pathway of dCTPαS incorporation of the mutants

Mutant	$K_{d, dNTP}$ (μM)	k_2 (s^{-1})	k_{-2} (s^{-1})	k_3 (s^{-1})
K522A	102±104	88±16	51±23	0.17±0.05
K522H	243±96	81±3.9	114±16	1.9±0.2
H506A	n.d.	n.d.	n.d.	n.d.
H506K	164±14	6.9±0.1	2.3±0.1	0.32±0.01

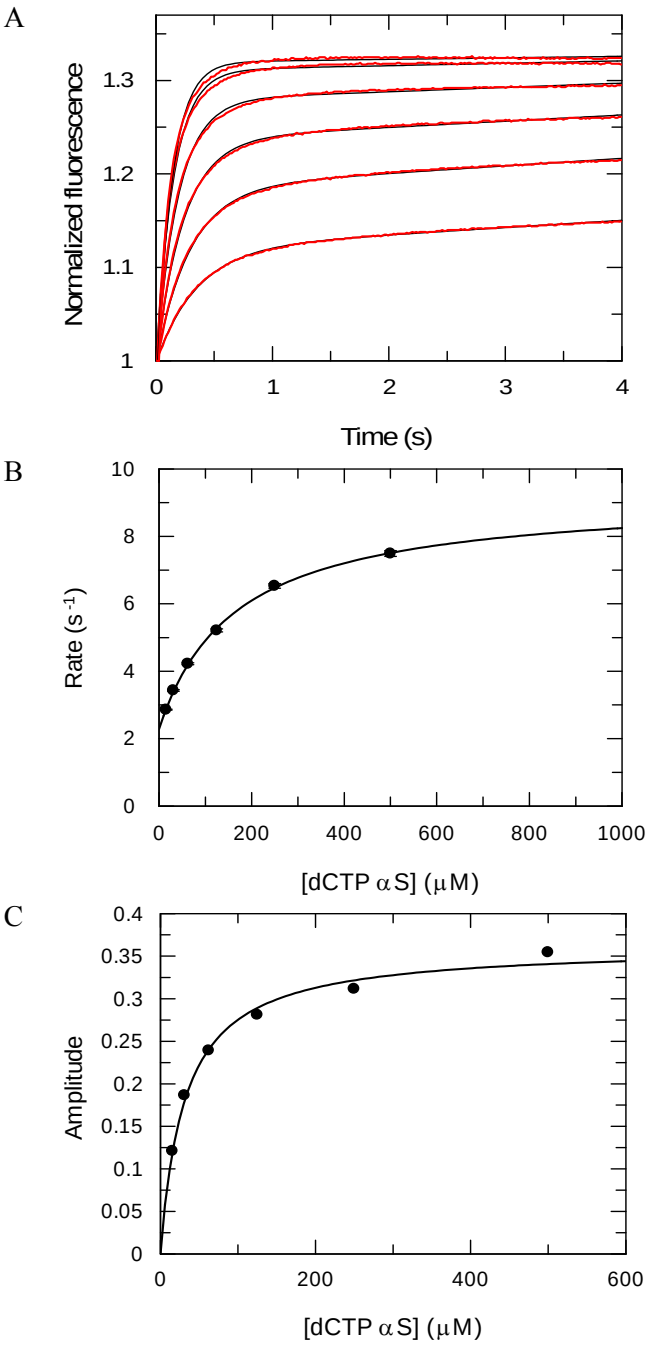


Figure 4.9: The rate of the conformational change induced by dCTPαS binding measured by stopped-flow experiments for H506K mutant

Figure 4.9: The rate of the conformational change induced by dCTPαS binding measured by stopped-flow experiments for H506K mutant.

(A) A preformed enzyme-DNA complex (200 nM enzyme and 600 nM DNA_{dC}) was rapidly mixed (1:1) with dCTPαS at various concentrations in a stopped-flow instrument and the time course of fluorescence change was recorded (from bottom to top, the red traces are for reactions with 15.6, 31.2, 62.5, 125, 250, and 500 μM dCTPαS concentrations, respectively). Each trace was fit to a burst equation to extract the rate and the amplitude of the fluorescence change (the fitted lines are in black). (B) The fitted burst rates were plotted against dCTPαS concentrations and the data were fit to a hyperbola, yielding a K_d of 164 ± 14 μM, k_2 of 6.9 ± 0.1 s^{-1}, k_{-2} of 2.3 ± 0.1 s^{-1}. (C) The fitted amplitudes were plotted against dCTPαS concentrations and the data were fit to a hyperbola, yielding an apparent K_d of 32 ± 4 μM and a maximum amplitude of 0.36 ± 0.01.

For global data fitting of the H506K mutant, both the stopped-flow data and the quench flow data were fit simultaneously and the best fit is shown in Figure 4.10 with the mechanism shown in Scheme 4.6.

$$E + D_n \underset{0.42 s^{-1}}{\overset{10 \mu M^{-1} s^{-1}}{\rightleftharpoons}} ED_n$$

$$ED_n + N \overset{125 \mu M}{\rightleftharpoons} ED_n N \underset{1.9 s^{-1}}{\overset{8.3 s^{-1}}{\rightleftharpoons}} FD_n N \overset{0.26 s^{-1}}{\rightleftharpoons} ED_{n+1} + PP_i$$

$$ED_{n+1} \underset{10 \mu M^{-1} s^{-1}}{\overset{0.42 s^{-1}}{\rightleftharpoons}} E + D_{n+1}$$

Scheme 4.6: dCTPαS incorporation reaction pathway for H506K mutant from global data fitting

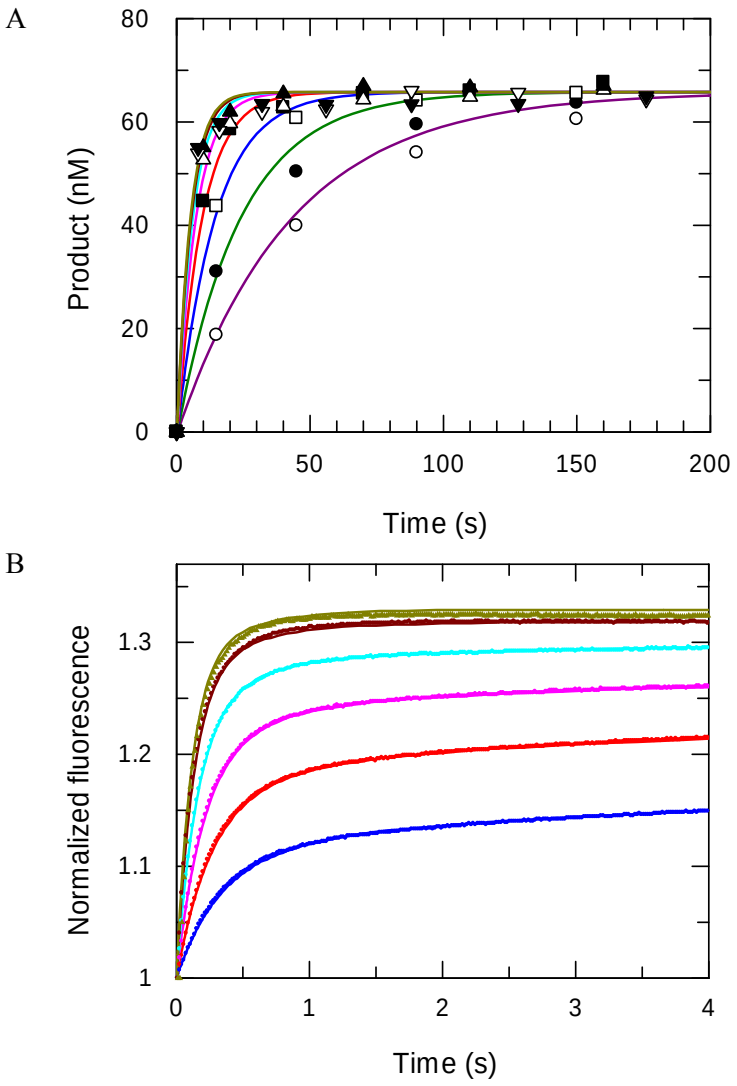


Figure 4.10: Global data fitting of dCTPαS incorporation for H506K mutant

(A) The series of data points were from quench flow experiments which were described in Figure 4.8A. The colored lines show the results of fitting to the mechanism shown in Scheme 4.6. (B) The series of data points represented the fluorescence changes from the stopped-flow experiments described in Figure 4.9A. The solid lines show the results of fitting to the mechanism shown in Scheme 4.6.

Global data fitting were also performed on mutants K522A, K522H, H506A using the same routine as well and the results of the fitting are shown in Appendix and the rate constants are summarized in Table 4.8, .

Table 4.8: Kinetic constants for dCTPαS incorporation from global data fitting

Mutant	$K_{d,\ dNTP}$ (μM)	k_2 (s^{-1})	k_{-2} (s^{-1})	k_3 (s^{-1})
K522A	146	30	53	0.24
K522H	132	39	46	1.35
H506A	100	11	7	0.5
H506K	125	8.3	1.9	0.26

Compared with the mechanism of dCTP incorporation shown in Scheme 4.4, the mechanism of dCTPαS suggests that the sulfur substitution does not affect the binding of nucleotide and the chemistry step is slower than the conformational change. By calculation, the thio-elemental effect is $k_{pol,dCTP} / k_{pol,dCTP\alpha S} = 0.36/0.26 \approx 1.4$. If the chemistry step is rate-limiting, then the small thio-elemental effect may suggest that the transition state of the phosphoryl transfer reaction for the H506K mutant is different from the wild type enzyme. Alternatively, if the full thio-elemental effect of phosphoryl transfer reaction is not large at all, then the larger thio-elemental effect seen from the wild type enzyme might be from a steric effect of the sulfur substitution. The thio-elemental effects for the fingers domain mutants are summarized in Table 4.9.

Table 4.9: Summary of kinetic constants for dCTP and dCTPαS incorporation and the thio elemental effect

Mutant	Nucleotide incorporated	$K_{d, dNTP}$ (μM)	k_2 (s-1)	k_{-2} (s^{-1})	k_3 (s^{-1})	$k_{3,dCTP}/k_{3,dCTP\alpha S}$
K522A	dCTP	270	35	31	0.7	2.9
	dCTPαS	146	30	53	0.24	
K522H	dCTP	623	90	47	2	1.5
	dCTPαS	132	39	46	1.35	
H506A	dCTP	70	15	3.5	0.9	1.8
	dCTPαS	100	11	7	0.5	
H506K	dCTP	260	10	0.7	0.36	1.4
	dCTPαS	125	8.3	1.9	0.26	
Wild type	dCTP	28	660	1.6	360	16.7*
	dCTPαS	1.9		14*		

The role of R518 during nucleotide binding and incorporation

R518 forms two hydrogen bonds with the non-bridge oxygen of the γ-phosphate of the incoming nucleotide in the closed form of T7 DNA polymerase tertiary complex. To probe the function of this residue during nucleotide incorporation, we made an R518A mutant with all hydrogen bonds removed and an R518K mutant where only one possible hydrogen bond is remained.

The kinetics of single nucleotide incorporation for the two mutants were studied by quench-flow method and the results are summarized in Table 4.1. The R518K mutant restored part of the function at this position as the rate of the nucleotide incorporation increased from 5.9 s^{-1} to 43 s^{-1}.

To test how these mutants affect the conformational change, nucleotide binding kinetics was characterized by stopped-flow assays and the results are shown in Figure 4.11 and Figure 4.12. In these experiments, the fluorescence changes corresponding to the conformational

changes induced by correct nucleotide binding to the DNA with dideoxy terminated primer were recorded. Different from the K522 and H506 mutants, the R518 mutants had smaller fluorescence change amplitude and the signal to noise ratio was lower which deteriorates the quality of the data. The smaller amplitude may imply the misalignment of the active site after binding of the nucleotide is less severe in the R518 mutants. The pattern of the fluorescence traces for the R518A implies a readily reversible conformational change with comparable forward and reverse rate. However, for the R518K mutant, the pattern implies the forward rate is much faster than the reverse rate and thus all the fluorescence traces reach same end point.

The concentration dependence of the rates of the fluorescence change was plotted in Figure 4.11B and Figure 4.12B for R518A and R518K respectively. The rates for R518A ranged from 25 s^{-1} to 14 s^{-1}, and the data could not be fitted to hyperbola equation. The rates for R518K increased with increasing dCTP concentration and the data were fit to a hyperbola, giving an apparent K_d of 11 ± 29 μM, k_2 of 104 ± 60 s^{-1}, and k_{-2} of 12 ± 62 s^{-1}.

Based on the available data, the tentative models for dCTP incorporation of the R518A and R518K mutants were proposed. For R518A, we proposed that the rate of the conformational change was close to the chemistry rate, and the conformational change was readily reversible with equal rates. For R518K, we thought the rates of conformational change and the chemistry are comparable and the reverse rate of the conformational change is same as the rate observed for the wild type enzyme. Scheme 4.7 shows the pathways and the microscopic rate constants were estimated from the quench-flow and stopped-flow results.

R518A $$ED_n + N \underset{}{\overset{200\mu M}{\rightleftharpoons}} ED_nN \underset{10s^{-1}}{\overset{10s^{-1}}{\rightleftharpoons}} FD_nN \overset{10s^{-1}}{\rightleftharpoons} FD_{n+1}PP_i$$

R518K $$ED_n + N \underset{}{\overset{75\mu M}{\rightleftharpoons}} ED_nN \underset{1.6s^{-1}}{\overset{100s^{-1}}{\rightleftharpoons}} FD_nN \overset{100s^{-1}}{\rightleftharpoons} FD_{n+1}PP_i$$

Scheme 4.7: Proposed pathway of dCTP incorporation for R518 mutants

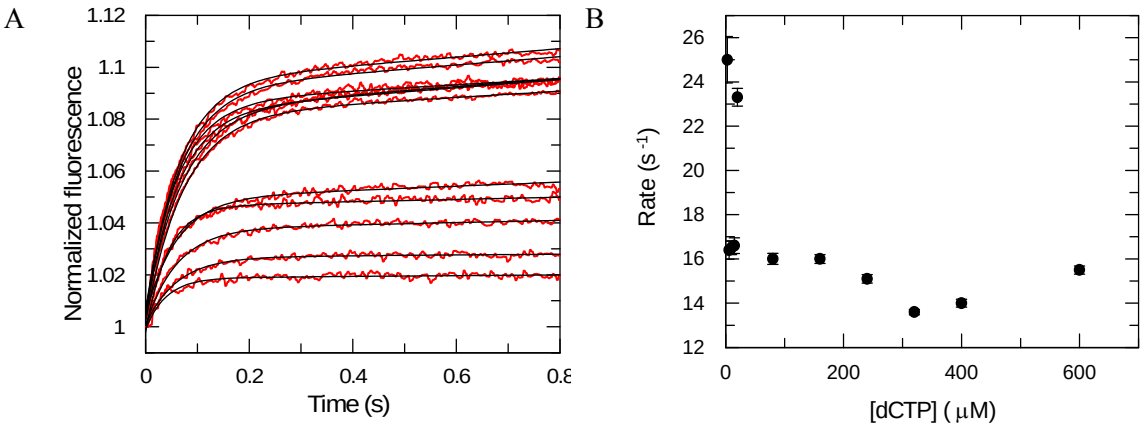


Figure 4.11: Conformational change induced by dCTP binding to the R518A mutant and DNA with dideoxy terminated primer

(A) Stopped-flow assays for the R518A mutant. A preformed enzyme-DNA complex (200 nM enzyme and 600 nM DNA_{ddC}) was rapidly mixed (1:1) with dCTP solution at various concentrations in a stopped-flow instrument and the time courses of fluorescence changes were recorded (from bottom to top, the red traces are for reactions with 3, 6, 10, 15, 20, 80, 160, 240, 320, 400, and 600 μM dCTP solutions, respectively). Each trace was fit to a burst equation to extract the burst rate and amplitude of the fluorescence change (the fitted lines are in black). (B) The fitted burst rates were plotted against dCTP concentrations and the rates range from 25 to 14 s^{-1} within the tested dCTP concentrations.

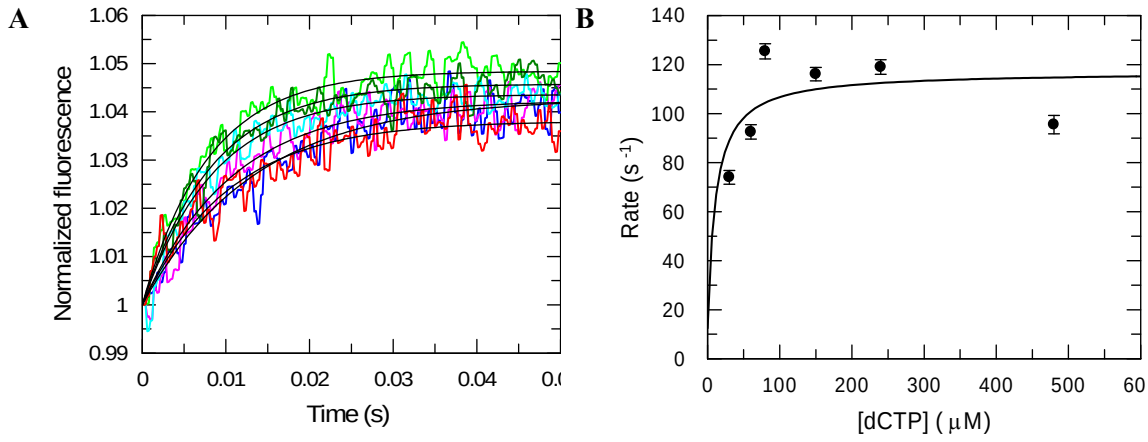


Figure 4.12: Conformational change induced by dCTP binding to R518K mutant and DNA with dideoxy terminated primer

(A) The stopped-flow assays for the R518K mutant. The reaction setup was same as the conditions for R518A as described in Fingure 4.9 (A) and the fluorescence traces are for reactions at 30 (blue trace), 60 (magenta), 80 (green), 150 (cyan), 240 (dark green), 480 (red) μM dCTP concentration. (B) The fitted rates were plotted with dCTP concentration and the data were fit to a hyperbola, giving an apparent K_d of 11 ± 29 μM, k_2 of 104 ± 60 s^{-1}, and k_{-2} of 12 ± 62 s^{-1}.

4.4 DISCUSSION

Previously, the pathway of nucleotide incorporation with ground-state binding of a nucleotide followed by an isomerization of the enzyme preceding the phosphoryl transfer reaction was established (Scheme 4.1). Extensive mutagenesis and kinetic studies on the DNA polymerase I have identified several active site residues which affect the nucleotide incorporation (78,79), but how these residues contribute to the kinetic constants along the pathway including the isomerization step was not clear. The goal of this study was to use site-directed mutagenesis and transient kinetics to characterize how these active site residues affect the reaction pathway, and thus define their mechanistic roles during nucleotide incorporation. To fulfill one part of this goal, I chose three fingers domain residues, K522, H506, and R518, which form ionic interactions with the α-, β-, γ-phosphates of the incoming nucleotide as shown in a high resolution crystal structure of the enzyme in its closed form. This crystal snap shot implies that these residues must play important roles during catalysis, but it does not give any dynamic information. Here, by using quench-flow and stopped-flow methods and global data fitting, the pathways of correct nucleotide incorporation of the mutants of these residues were solved and summarized in Table 4.9. The free energy differences for each step along the pathway between the mutant and wild type enzyme were calculated and presented in Figure 4.13.

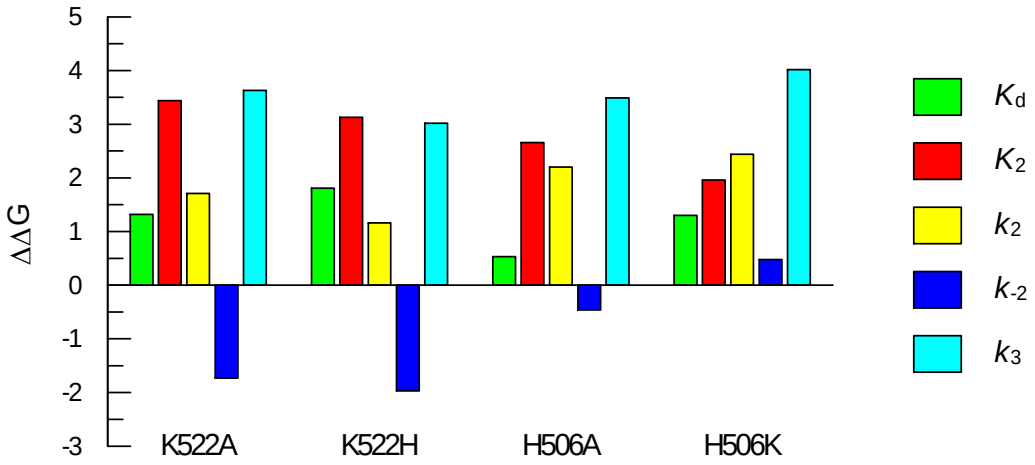


Figure 4.13: ΔΔG profiles of the fingers domain mutants

The role of K522 during conformational change and chemistry

The effect of the mutations of this lysine on the ground-state nucleotide binding is about 1.3 and 1.8 kcal/mol for K522A and K522H, respectively. These energy changes are close to the loss of one hydrogen bond, implying the K522 contributes a hydrogen bond during ground-state binding of a nucleotide. Since the histidine should be partially charged under the reaction condition (pH7.5) and it is polar, we might expect to see partial restoration of the function. Interestingly, the histidine substitution did not restore any affinity and even decreased a bit, implying the altered side chain may affect the geometry of the binding surface for nucleotide.

Both the forward and reverse rates of the conformational change were changed about 20-fold with the mutants. The forward rate is slower but the reverse rate becomes faster. As a result, the net change in the equilibrium constant of the conformational change step is significant. Loss of the lysine made isomerization following the binding of a correct nucleotide unfavorable, suggesting that this lysine may be part of the fast nucleotide recognition element and it may stabilize the binding of a correct nucleotide.

It is not surprising that the loss of lysine affected the chemistry step significantly. The electrostatic interaction between the lysine and the non-bridge oxygen of the α-phosphate can

stabilize the developing negative charge on the oxygen during the transition state. Thus the data suggest that the lysine stabilizes the transition state of the reaction and the state captured in the crystal structure snap shot may accurately represent the reactive complex.

The role of H506 during nucleotide binding and incorporation

The energy difference between the wild type and the H506A and H506K mutants for the ground-state nucleotide binding are 0.5 and 1.3 kcal/mol, respectively. These relatively small energy differences imply the polar interaction between this H506 and the nucleotide during ground-state binding is weaker than the interaction between K522 and nucleotide. In a crystal structure of T7 DNA polymerase in its open state (PDB entry 1SL0), the incoming nucleotide was bound on the fingers domain and formed polar interactions with the H506 and the R518 residues (59). The conservative histidine at position 734 of the *E. coli* DNA polymerase I was also close to the phosphate groups of a bound nucleoside triphosphate in the open state of the enzyme (PDB entry 1KFD) (19). Thus the H506 may be involved in nucleotide ground-state binding but the interaction is not as strong as the one formed between K522 and the nucleotide.

Interestingly, only the forward rate of the conformational change was significantly decreased by the H506 mutations, while no change on the reverse rate. This result implies the histidine is involved in the fast delivery of the nucleotide to the closed state of the enzyme and it does not affect the reverse rate. Since the rate decreased more in the H506 mutants than in the K522 mutants, H506 may play a more important role for the fast delivery. The absence of an affect on the reverse rate could be explained in two ways. First, the possible lock function of the K522 may keep the reverse rate slow. Second, the nearby Y526 also forms a hydrogen bond with the non-bridge oxygen of the β-phosphate of the nucleotide, so it can stabilize the bound nucleotide allowing it go forward to the chemistry.

The H506 mutations caused a large reduction in the rate of the chemistry step implying that H506 helps to stabilize the transition state of the reaction. This is not surprising too, since

H506 forms hydrogen bond with the non-bridge oxygen on the β-phosphate of the incoming nucleotide sitting in the enzyme active site in the closed state.

The role of R518 during nucleotide binding and incorporation

Since the signal-to-noise ratio of the stopped-flow data was relatively low, the data analysis and the modeling of the reaction mechanism for the two R518 mutants were done separately from other mutants. However, we still derived valuable data to evaluate the role of R518.

From the quench-flow data, there was modest effect on the ground-state binding, suggesting R518 may be involved in this step. As discussed above, in the open state of the enzyme, there could be one hydrogen bond formed between this R518 and the incoming nucleotide. Estimated from the stopped-flow data, the R518A mutant had a much slower forward rate of the conformational change than the wild type enzyme, suggesting this arginine can speed up the forward conformational change. As supportive evidence, the lysine substitution partially restored the rate of this forward conformational change. The R518A mutant caused the reverse rate of the conformational change to be faster than for the wild type enzyme, judged from the pattern and the rates of the fluorescence traces. The R518K mutant restored this rate to the level of the wild type enzyme. These results suggest this R518 and the K522 may work together to stabilize nucleotide binding. The R518 mutations greatly reduced the rate of chemistry step, suggesting the R518 also plays a role in stabilizing the transition state. However, the R518 mutants had less severe effect on the chemistry step due to the fact that the polar interaction formed between the R518 and the γ-phosphate is the furthest from the reaction center, the α-phosphate.

Understanding the thio-elemental effect on nucleotide incorporation

Thio-elemental effect in nucleotide incorporation refers to the effect on the phosphoryl transfer reaction caused by a thio substitution on non-bridge oxygens of the phosphate group of

the nucleotide. Sulfur has weaker electron withdrawing ability relative to oxygen, so the transition state of the phosphoryl transfer reaction is less stable in the case of a sulfur substitution leading to a slower chemistry rate. The magnitude of the thio-elemental effect for nucleotide incorporation was quantified by measuring the ratio of the rate for normal nucleotide to the rate of the reaction for the sulfur substituted nucleotide ($k_{dNTP}/k_{dNTP\alpha S}$).

Our data showed that the thio-elemental effects for all the mutants were from 1.4 to 2.9, which are relatively small compared to the value of 16.7 observed for the wild type enzyme. This could be explained by at least three possibilities. First, there is a rate-limiting conformational change preceding the chemistry. However, our stopped-flow data showed that the sulfur substitution did not affect the ground-state nucleotide binding and the conformational change step significantly. It is possible that there is an extra conformational change step preceding chemistry but it is silent in our data.

Second, the transition state of the mutants might be different from the wild type enzyme. For a reaction in free solution, the magnitude of the thio-elemental effect is related to the type of the transition state of the chemistry step. Usually a more associative transition state has the larger thio-elemental effect while a dissociative state has smaller thio-elemental effect. Furthermore, there is a trend in which the monoester has the smallest thio-elemental effect, the diester is intermediate and the triester has the largest effect, suggesting the triester's reaction transition state is more associative (48,80,81). However, in an enzyme catalyzed reaction, the situation becomes complicated. Usually the type of the transition state for phosphoryl transfer reaction in an enzyme catalyzed reaction is unknown due to the fact that the contact between the enzyme and the phosphate may also change the thio-elemental effect. The large thio-elemental effect, about 60, seen from misincorporation of T7 DNA polymerase, implied that the transition state is more associative in the enzyme catalyzed reaction than in solution(39). Since much smaller thio-elemental effects were seen for these mutants, the transition state in these mutants might be less

associative than that in the wild-type enzyme. However, the studies from alkaline phosphatase catalyzed phosphoryl transfer reaction suggested that the transition state in enzyme catalyzed reaction is same as that in solution (82,83).

Finally, it is possible that the actual thio-elemental effect caused by the electronegativity of the sulfur is not large at all for DNA polymerase catalyzed phosphoryl transfer reaction. Rather, the larger thio-elemental effect observed with the wild type enzyme during correct nucleotide incorporation might be caused by steric effects. The closed T7 DNA polymerase active site is tight, so replacement of the smaller oxygen (1.5 Å Van der Waals radius) with a larger sulfur (1.8 Å) may cause steric effects on the side chains of the enzyme, resulting in slight misalignment of the active site. Thus the reduced rate of the reaction with a thio-substituted nucleotide analogue might be from this misalignment not from the weaker electronegativity of the sulfur atom. Then the even larger thio-elemental effect of about 130, seen from the misincorporation reaction of MDCC labeled T7 DNA polymerase (unpublished data from Yu-Chih Tsai), can be explained by a more severe distortion of the active site when sulfur was used. For the fingers domain mutants we tested, K522A, K522H and H506A, all have a smaller side chain than the wild type enzyme. These smaller side chains may relax the active site leading to smaller or no steric effect. Therefore the observed thio-elemental effects could reflect the intrinsic thio-elemental effect for the phosphoryl transfer reaction catalyzed by T7 DNA polymerase.

The induced-fit mechanism of nucleotide binding and incorporation

Based on the information from the crystal structures of T7 DNA polymerase and other DNA polymerases and our kinetic analysis on these fingers domain mutants, a mechanism of nucleotide binding and incorporation for T7 DNA polymerase was proposed.

The next incoming nucleotide bound with one magnesium ion enters the active site and makes electrostatic interactions with catalytic residues D475 and D654 that are negatively charged. At the meantime, the base of the correct nucleotide might be able to form more polar

interactions with the templating base than an incorrect nucleotide, which can explain why a mismatch has a weaker ground-state binding affinity (39). The positively charged K522 and R518 with the partially charged H506 form a dynamic binding surface for the negatively charged phosphate group of the incoming nucleotide through long-range electrostatic interactions. These long-range interactions are stronger when a correct nucleotide enters the active site than when an incorrect nucleotide enters, since the initial alignment of the active site is better for binding a correct nucleotide than binding an incorrect nucleotide. The correct nucleotide induces a faster closure of the fingers domain through these interactions. The interactions between K522 and the incoming nucleotide might vary depending on whether a correct or incorrect nucleotide is bound, and upon recognition of a correct nucleotide, K522 can lock a correct nucleotide in the active site while release an incorrect nucleotide. Without this lock, as shown for the K522A mutant, the fingers domain can move back rapidly. After a perfectly aligned active site is formed, all three fingers domain residues contribute polar interactions to stabilize the transition state of the phosphoryl transfer reaction for fast catalysis. Since the K522 and H506 are closer to the α-phosphate, their contributions to the rate of chemistry are more than the R518 residue.

Chapter 5: Crystallization of CSF Labeled T7 DNA Polymerase.

5.1 INTRODUCTION

The fluorescence intensity of MDCC labeled T7 DNA polymerase decreases upon binding a correct nucleotide but increases upon binding an incorrect nucleotide (69). The decrease in fluorescence suggests quenching of fluorescence by solvent and the increase suggests the fluorophore shifts to a more hydrophobic environment. Crystal structures of the ternary complex of MDCC labeled T7 DNA polymerase with DNA and correct or incorrect nucleotide will be helpful to explain the origin of the fluorescence signal. Moreover, the crystal structures will show the difference between an aligned and a misaligned active site. Tsai and Johnson have proposed that upon binding an incorrect nucleotide, the binding energy was utilized to misalign the recognition domain, leading to slow catalysis. Therefore, the enzyme could exist in three possible states: open, closed and misaligned. Studies of these three complexes will help to reveal the structural basis for nucleotide selectivity.

A crystal structure of the T7 DNA polymerase ternary complex with a correct nucleotide was solved to 2.2 Å resolution in Ellenberger's lab (14), but attempts to obtain a structure with an incorrect nucleotide at the insertion position have not succeeded. The low affinity ($K_d \approx$ 8mM) of binding a mismatched nucleotide and the high salt crystallization conditions may account for the inability to stabilize a mismatch in the binding site. However, the MDCC labeled T7 DNA polymerase showed a higher affinity for the mismatch ($K_d \approx$ 130 μM measured from a titration experiment), making crystallization of a mismatch at the insertion position more feasible (69). Although such structures might be considered artifacts, we should note that the catalytic rate of incorporation of an incorrect nucleotide by an MDCC labeled enzyme was not changed compared to the unlabeled enzyme, therefore the MDCC tag might not alter the structure of the active site

but increase the affinity to an incorrect nucleotide to the active site by an unknown mechanism. A structure of such complex may still reflect the properties of an active site with a mismatch.

Crystals of the ternary complex of the MDCC labeled enzyme and double stranded DNA, with either a correct or an incorrect nucleotide, were obtained by following the procedures established in the Ellenberger lab (14,56,59). The X-ray diffraction data were collected, and structures were determined by molecular replacement. The atomic structures are still under refinement.

5.2 MATERIAL AND METHODS

Expression, purification and MDCC labeling of the cys-light T7 DNA polymerase

The procedures were modified from the previous protocol for the cys-light T7 DNA polymerase described in Chapter 2. Frozen cells with the plasmid containing the cys-light T7 DNA polymerase was inoculated into 60 ml of NZY medium (10 g of NZ amine, 5 g of NaCl, 5 g of bacto-yeast extract, and 2 g of $MgSO_4$-$7H_2O$ dissolved in 1 L of ddH_2O) containing 50 μg/ml of ampicillin and 12.5 μg/ml of chloramphenicol. The culture was incubated at 37 ºC overnight on a shaker rotating at 250 rpm. An 8-ml overnight culture then was transferred to 1.5 L fresh NZY medium containing 50 μg/ml of ampicillin. This culture was incubated for about 3 hours at 37 ºC with shaking at 175 rmp till OD_{600} reached 0.3~0.4. The culture then was slowly cooled to 14 ºC, and then IPTG and chloramphenicol were added at final concentration of 0.4 mM and 12.5 μg/ml, respectively. After about 20 hours' induction, the cells were harvested by centrifugation at 5000 x g for 15 min at 4 ºC.

The cells were frozen at –80 ºC for more than half hour and then thawed in lysis buffer (50 mM Tris-HCl, pH 8, 2.5 mM of EDTA, 150 mM NaCl, 1 mM β-mercaptoethanol, 0.1 mM DTT, and10 mM of phenylmethylsulfonyl fluoride). A Kontes douncer was used to fully suspend cells in the lysis buffer. After addition of 0.3 mg/ml lysozyme, the cell suspension was stirred at room temperature for 15 min, followed by sonication on ice (50% duty cycle, 5 output control)

till the suspension could drip as discontinuous drops from a pipette tip. For the remaining protein purification steps, the solution was either kept on ice or at 4 ºC. With continuous stirring of the solution, sodium deoxycholate and sodium chloride were added to a final concentration of 0.1% (w/v) and 0.5 M, respectively. PEI was then slowly added to a final concentration of 0.3 % (w/v) to precipitate DNA. The solution was stirred for another 30 min and was cleared by centrifugation at 16,000 x g. The supernatant was fractionated by 35% and 70% ammonium sulfate precipitation (discarded 35% pellet while saved 70% pellet). The pellet was resuspended in a dilution buffer (50 mM Tris-HCl, pH 7.5, 0.1 mM EDTA, 10% glycerol, and 5 mM DTT) and dialyzed against this buffer for about 2 hours until its conductivity was equal to that of the low salt column running buffer (50 mM Tris-HCl, pH 7.5, 0.1 mM EDTA, 100 mM NaCl, 10% glycerol, and 5 mM DTT). The protein solution then was loaded onto a 200 ml DEAE-sepharose column at 2 ml/min flow rate using a Pharmacia FPLC. The column was washed with low salt running buffer and bound proteins were eluted by a 1.6 L 100-400 mM NaCl gradient in column running buffer. The fractions containing T7 DNA polymerase (eluted between 200-250 mM NaCl gradient) were identified by 8 % SDS-PAGE gel and pooled. The pooled fractions were diluted with the dilution buffer to a conductivity equal to that of the low salt running buffer. The diluted protein solution was loaded onto a 10 ml single-stranded DNA cellulose (ssDNA-cellulose) column. The column was washed with a low salt running buffer containing 1 mM Tris(2-carboxyethyl)phosphine hydrochloride (TCEP) without DTT. The bound T7 DNA polymerase was eluted with a high salt running buffer (50 mM Tris-HCl, pH 7.5, 0.1 mM EDTA, 400 mM NaCl, 10% glycerol, and 1 mM TCEP).

The fractions containing T7 DNA polymerase were pooled and dialyzed against 1 L of MDCC labeling buffer (40 mM HEPES, pH 8, 0.1 mM EDTA, 10% glycerol, 50 mM NaCl, and 1 mM TCEP) twice for total 4 hours. MDCC (10 mM stock solution in DMSO) was added to and immediately mixed with the protein solution to a final concentration of 0.3 mM. The labeling

reaction was incubated with constant slow mixing for about 16 hours at 4 ºC. The labeled protein solution was loaded onto a 10 ml ssDNA-cellulose column to remove free MDCC with the low salt running buffer. The bound MDCC labeled protein was eluted by the high salt running buffer. The peak fractions were pooled and diluted with dilution buffer to an estimated salt concentration close to the low salt running buffer. Excess thioredoxin (about 2 times molar ratio relative to polymerase catalytic subunit) was added to the diluted protein, and the holoenzyme was loaded onto a 1 ml Mono-S column. The column was washed with the low salt running buffer and the bound protein was eluted by 16 ml 100-400 mM NaCl gradient. Two peaks were seen from the UV absorbance and only the first peak contained T7 DNA polymerase bound with thioredoxin as identified by 8% SDS-PAGE. The eluted MDCC labeled protein was concentrated by using a Vivaspin 4 concentrator (5 KDa, Vivascience) to 400 μM.

Protein concentration measurement

The DU 800 spectrophotometer (Beckman Coulter) was set to a fixed wavelength 280 nm. The machine then was blanked with 91 μl of 6.6 M guanidium chloride in a cuvette with 1 cm length. 9 μl of sample was added into the cuvette and thoroughly mixed with a pipette. For calculation of the protein concentration, an extinction coefficient of 0.1344 cm^{-1} μM^{-1} was used for T7 DNA polymerase only and 0.148 cm^{-1} μM^{-1} for T7 DNA polymerase bound with thioredoxin.

Crystallization

The procedure of the crystallization experiments was modified from the published procedures (14,56,59). The complex of MDCC labeled T7 DNA polymerase and thioredoxin (100 μM) was mixed with equimolar amounts of DNA duplex (21-mer primer: 5' CGAAAACGACGGCCAGTGCCA3'; 26-mer template: 5' ATGGATGGCA CTG GCCGTCGTTTTCG3'), and 0.5 mM ddTTP (a chain terminator). The reaction mixture was incubated on ice for 30 minutes to allow the incorporation of ddTTP onto the primer and then 10

mM ddCTP was added as a next incoming correct nucleotide. In the case of crystallization of T7 DNA polymerase with a mismatched nucleotide, 10 mM ddGTP was added into the solution. The open complex was made with T7 DNA polymerase and DNA only. An example of a crystallization reaction setup was shown in Table 5.1.

Table 5.1: Preparation of 10 µl crystallization reaction mixture

Ingredient	Correct complex	Mismatch complex	Open complex
10x reaction buffer1	1	1	1
10 mM DTT	2	2	2
water	1.8	1.8	2.8
MDCC labeled enzyme (0.4 mM)	2.5	2.5	2.5
DNA (21/26-mer, 0.85 mM)	1.2	1.2	1.2
ddTTP (10 mM)	0.5	0.5	0.5
ddCTP (100 mM)	1	0	0
ddGTP (100 mM)	0	1	0

Note:

1. 10x reaction buffer: 500 mM HEPES, pH 7.5, 100 mM $MgCl_2$.

The hanging drop method was used to set up crystallization experiments. The reservoir contained 0.5 ml reservoir buffer, and 1 µl of reaction mixture was mixed with 1 µl reservoir buffer on a coverslip. Then the reservoir well was sealed by the coverslip with vacuum grease. The trays were incubated in the cold room (4 ºC). The reservoir buffers with increasing concentrations of polyethylene glycol (PEG) 8000 (12-22%) were tested (Table 5.2).

Table 5.2: Preparation of 10 ml reservoir buffer with increasing percentage of PEG8000

Stock solution	12%	13%	14%	15%	16%	17%	18%	19%	20%	21%	22%
500mM ACES, pH7.5 (5X)	2	2	2	2	2	2	2	2	2	2	2
600mM Ammo. Sulf (5X)	2	2	2	2	2	2	2	2	2	2	2
250mM $MgCl_2$ (25/3X)	1.2	1.2	1.2	1.2	1.2	1.2	1.2	1.2	1.2	1.2	1.2
100mM DTT (20X)	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
50% PEG8000	2.4	2.6	2.8	3	3.2	3.4	3.6	3.8	4	4.2	4.4
water	1.9	1.7	1.5	1.3	1.1	0.9	0.7	0.5	0.3	0.1	0

Crystal freezing conditions

Before X-ray data collection, the crystals were stepwisely transferred to a solution containing all the reservoir components plus 15% glycerol to prevent the damages to the crystals and ice formation. In the first step, the crystals were soaked in reservoir buffer with 5% glycerol for 5 minutes, then the crystals were transferred to the buffer with 10% glycerol and soaked for another 5 minutes. Finally the crystals were soaked in the buffer with 15% glycerol for 5 minutes and then mounted on nylon loops and flash-frozen in liquid nitrogen.

X-ray data collection and structure determination

X-ray data were collected with the Rigaku H3R rotating anode generator and a Raxis4++ detector. The X-ray data were indexed, integrated and scaled with the program HKL. The structures were determined by molecular replacement with the Molrep program in the CCP4 program suite 6.0 (84). The models then were refined with the Refmac5 program in the same program suite and were built with the program O.

5.3 RESULTS AND DISCUSSION

The purification procedure is very important for crystallization of the MDCC labeled enzyme. Initially the crystallization of the complexes was attempted without the mono S column

purification, but the crystallization failed. These crystallization reactions might be contaminated by single stranded DNAs from the ssDNA column. After the enzyme-thioredoxin complex was further purified over a Mono S column, a strong cation exchanger, to remove DNA, the enzyme complexes was finally crystallized.

Thin sheet-like crystals of the correct complex appeared after 6 days. These crystals were crushed into small seeds that were used to microseed the crystallization reactions for the correct, open, and mismatch complexes. After microseeding, crystals grew overnight. For the open and mismatch complexes, the initial seeds from the correct nucleotide mixture introduced certain quantity of the correct nucleotide, so that the crystals grown using these seeds may contain complexes with correct nucleotide. To dilute the contaminated correct nucleotide, another two rounds of seeding were performed. Many crystals grew overnight for the reactions of the open and mismatch complexes after every microseeding, and we think these crystals were less likely to be contaminated by the correct nucleotide after at least three times of microseeding. Crystals grown in lower concentration of PEG8000 (8-11%) reservoir solutions with microseeding were bigger and in better quality.

Figure 5.1 shows the crystals of each complex grown after microseedings. MDCC is a yellow-colored compound under ambient light, and all the crystals observed were yellow under the stereo microscope, strongly suggesting that these crystals contain MDCC labeled T7 DNA polymerase. Majority of the crystals were a rhombic shape with the longest dimension of about 0.2-0.3 mm.

Before x-ray data collection, the crystals were frozen in liquid nitrogen on a nylon loop. Cryoprotectants such as PEG400, ethylene glycol, and glycerol were tested. Stepwise transfering of the crystals into 15% glycerol with all other reservoir buffer components worked best as no damage in the crystals was observed during transfer and no ice formation in the crystals judged from the x-ray data.

The crystals diffracted to 2.8 Å, 3.2 Å, 3.0 Å for the correct complex (Figure 5.1D), the open complex and the mismatch complex, respectively, and the data collection and processing statistics are summarized in Table 5.3. The crystals of the three complexes belong to space group $P2_12_12$, with unit cell dimensions similar to those reported from the Ellenberger lab (14,56). For the correct complex, a structure of T7 DNA polymerase and DNA with a correct nucleotide, dCTP, at insertion site (PDB entry 1T8E) (56) was used in molecular replacement. For the mismatch complex and the open complex, an open structure (PDB entry 1TK5) (56) was used in molecular replacement. From the electron density maps of the correct complex, extra densities were seen around the exonuclease active site, as a result of the difference between the amino acid compositions of our protein and the ones published by Ellenberger; six residues near the exonuclease active site had been removed to deactivate the exonuclease in the 1T8E and 1TK5 models. There was not much extra electron density around Cys 514, where the MDCC label was located, suggesting that this area of the molecule is dynamic. The models are still under refinement; and with the 3.0 Å data, we may finally get some informative models to explain the fluorescence state of the enzyme when bound with different substrates.

Table 5.3: Data collection and processing statistics

	Correct complex	Open complex	Mismatch complex
area detector	RAXIS4++	RAXIS4++	RAXIS4++
wavelength (Å)	1.5418	1.5418	1.5418
temperature (K)	100	100	100
no. of mole. In asu	1	1	1
space group	$P2_12_12$	$P2_12_12$	$P2_12_12$
cell dimenstions			
a (Å)	214.84	214.6	216.06
b (Å)	63.37	63.03	63.28
c (Å)	105.16	103.13	105.27
resolution (Å)	25.0 ~ 2.8	27.8 ~ 3.2	30.0 ~ 3.0
no. of processed reflections	32,508	21,834	26,891
model used in Molecular Replacement	1T8E	1TK5	1TK5
R_{work}	0.274	0.245	0.225
R_{free}	0.381	0.374	0.349

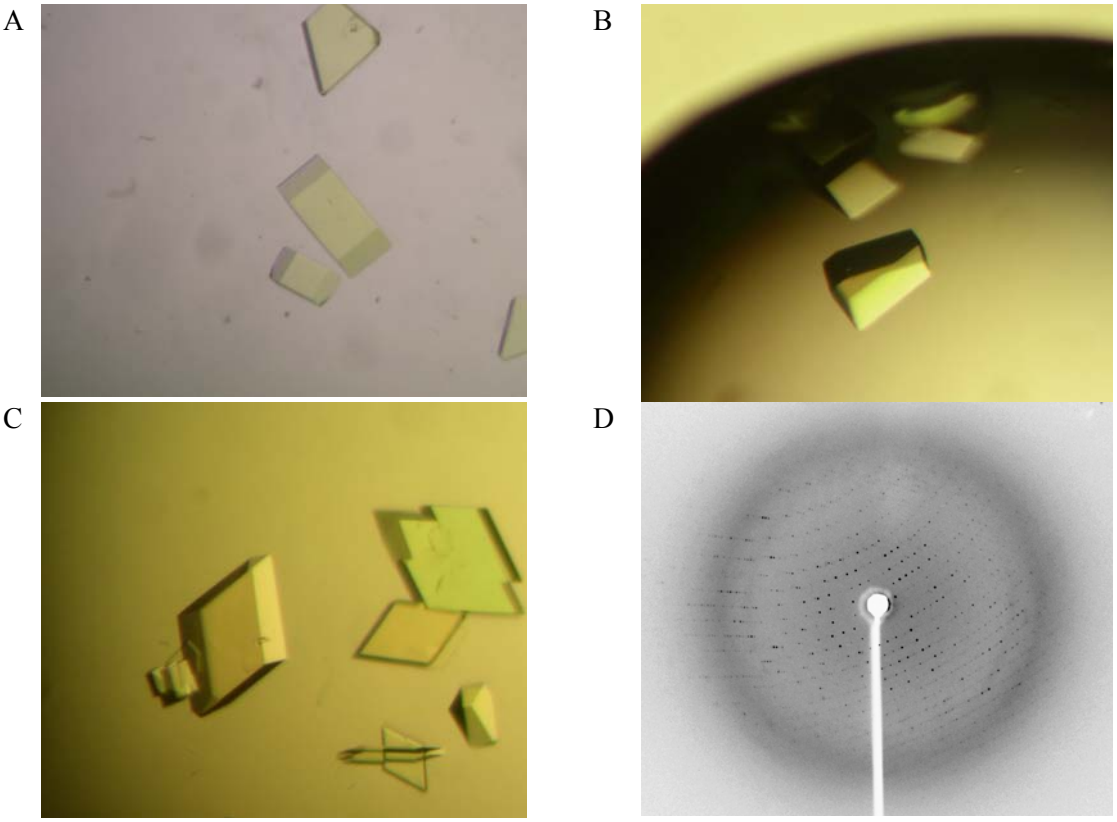


Figure 5.1: Crystallographic study of MDCC labeled T7 DNA polymerase

(A) Open complex crystals were grown in a 9% PEG8000 solution with 100 μM enzyme/thioredoxin, 100 μM DNA, and 0.5 mM ddTTP. Note the crystals are yellow due to the yellow color from the MDCC label. (B) Correct complex crystals were grown in 9% PEG8000 solution with 100 μM enzyme/thioredoxin, 100 μM DNA, 0.5 mM ddTTP, and 10 mM ddCTP. (C) Mismatch complex crystals were grown in 9% PEG8000 solution with 100 μM enzyme/thioredoxin, 100 μM DNA, 0.5 mM ddTTP, and 10 mM ddGTP. (D) A diffraction pattern from a correct complex crystal. The crystal diffracted to 2.8 Å.

One might argue that the crystals of the complex containing a mismatch might be similar to other open complexes observed before, thus not giving any valuable information for misincorporation. So far, no one has achieved capture of a mismatch at the catalytic site for T7 DNA polymerase, presumably due to the weak binding affinity of a mismatch and high salt concentration in the crystallization reactions. However, for this MDCC labeled enzyme, we observed higher binding affinity for the mismatch nucleotides than the non-labeled enzyme while the rates of misincorporation were close for labeled and unlabeled enzyme. Figure 5.2 and 5.3 depict the results of single nucleotide incorporation assays for misincorporation of dGTP with unlabeled enzyme and MDCC-labeled enzyme. The MDCC-labeled enzyme binds the mismatched dGTP 17.5-fold more tightly than the unlabeled enzyme. The results for all the misincorporation onto the templating dGMP are summarized in Table 5.4. The labeled enzyme showed tighter binding for all the mismatched nucleotide opposite templating dGMP than the unlabeled enzyme. Therefore it is more likely that we may get a mismatched nucleotide bound at the active site; however, only solving the mismatch complex structure will tell the truth.

Table 5.4: Effect of the MDCC tag on nucleotide misincorporation

	unlabeled		MDCC-labeled		Affinity Change
	K_d (mM)	k_{pol} (s^{-1})	K_d (mM)	k_{pol} (s^{-1})	
dATP	2.7±0.8	0.58±0.09	0.22±0.05	0.71±0.04	12
dGTP	2.1±0.8	0.13±0.02	0.12±0.03	0.13±0.01	17.5
dTTP	3.4±0.5	0.65±0.05	0.54±0.06	0.76±0.02	6.8

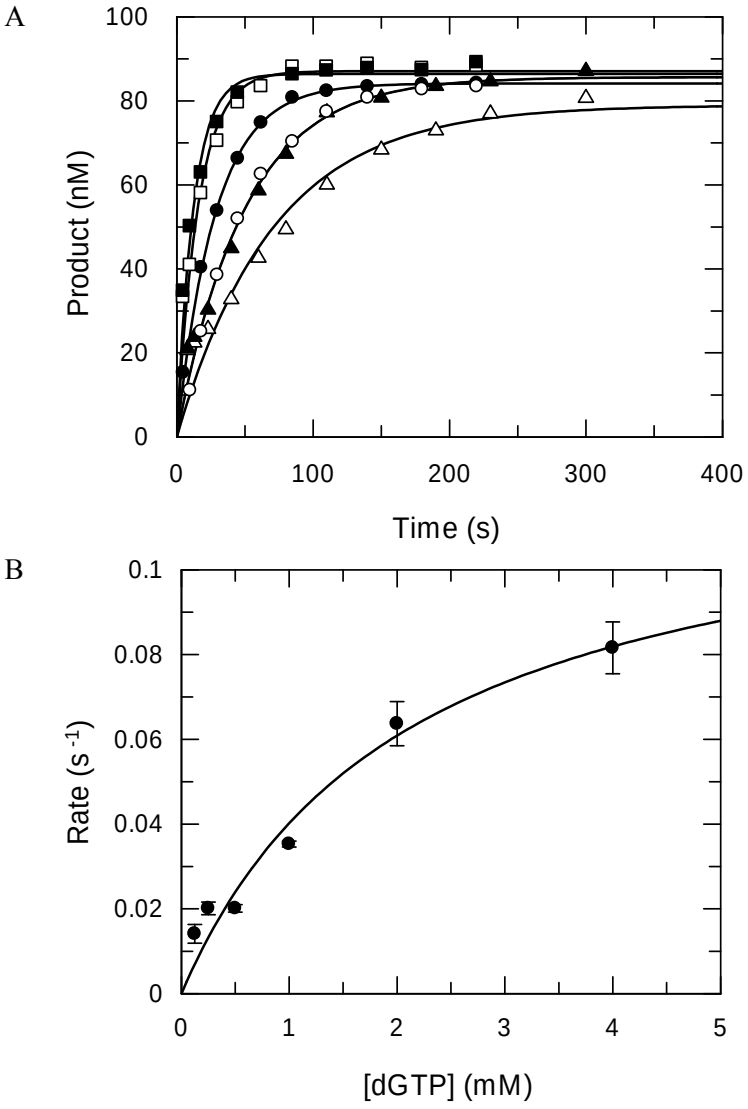


Figure 5.2: Misincorporation of dGTP onto templating dGMP with unlabeled enzyme

(A) The kinetics of misincorporation of dGTP was assayed under single nucleotide incorporation condition over a range of 0.125 (Δ) ~ 4 mM (■) dGTP. The time courses of product formation were fit to single exponentials to extract the observed rates of misincorporation. (B) The observed rates were plotted against dGTP concentration and fit to a hyperbola, yielding a K_d of 2.1 ± 0.8 mM and a k_{pol} of 0.13 ± 0.02 s^{-1}.

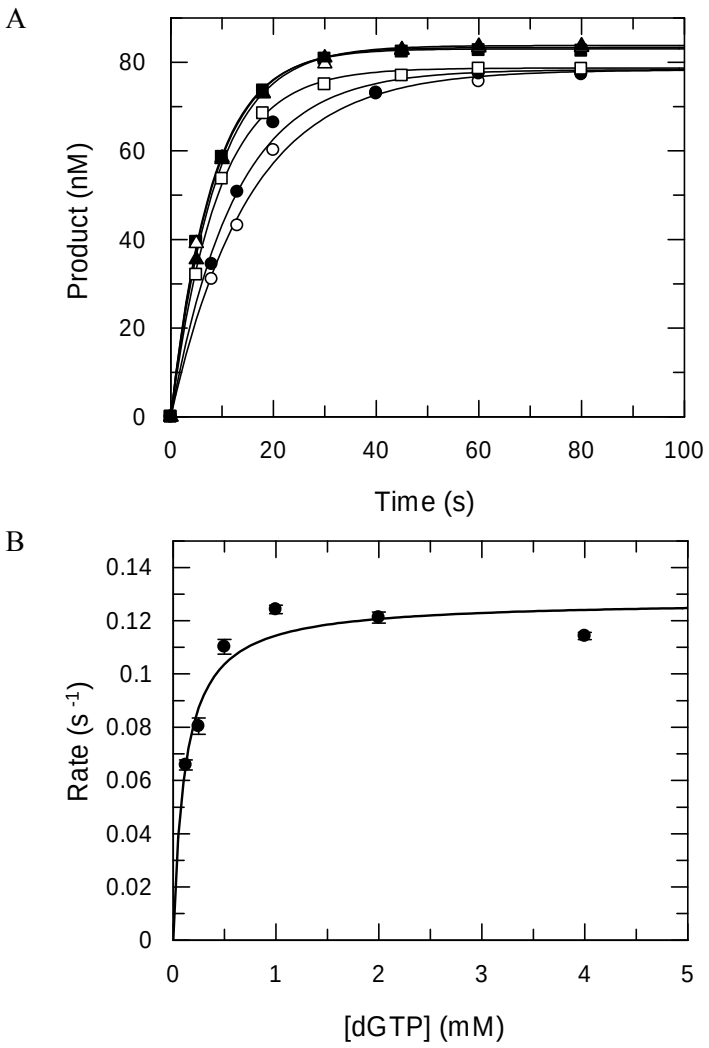


Figure 5.3: Misincorporation of dGTP onto templating dGMP with MDCC-labeled enzyme

(A) The kinetics of misincorporation of dGTP was assayed under single nucleotide incorporation condition over a range of 0.125 (○) ~ 4 mM (▲) dGTP. The time courses of product formation were fit to single exponentials to extract the observed rates of misincorporation. (B) The observed rates were plotted against dGTP concentration and fit to a hyperbola, yielding a K_d of 0.12 ± 0.03 mM and a k_{pol} of 0.13 ± 0.01 s^{-1}.

Appendix

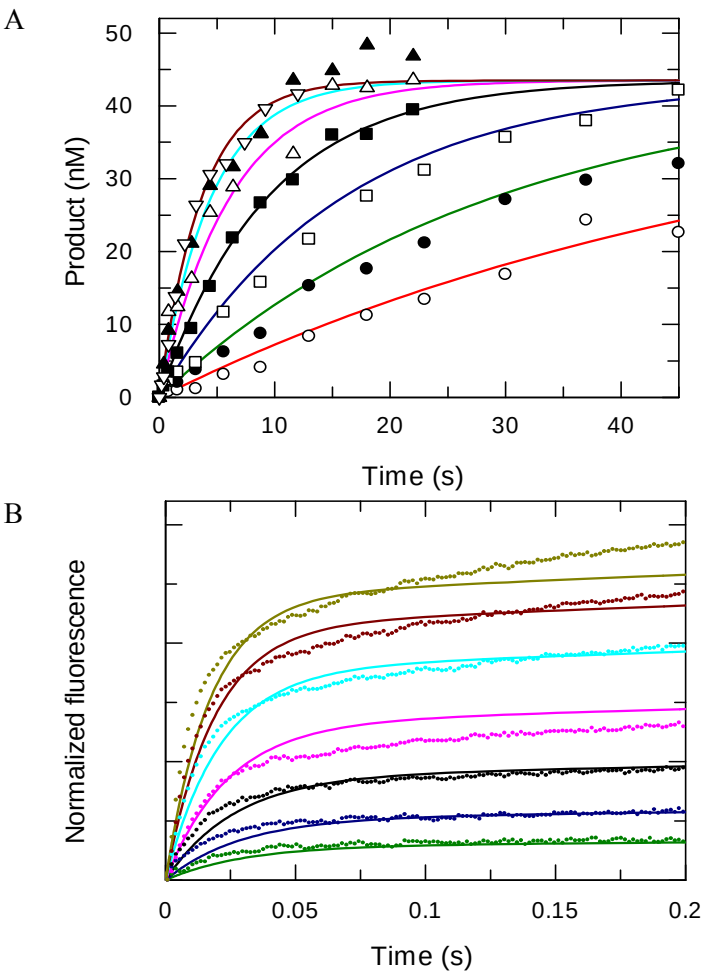


Figure 6.1: Global data fitting of dCTP incorporation for K522A mutant

(A) The series of data points were from quench flow experiments. The colored lines show results of fitting to the mechanism shown in Scheme 6.1. (B) The series of data points represented the fluorescence changes from the stopped-flow experiments. The solid lines show results of fitting to the mechanism shown in Scheme 6.1.

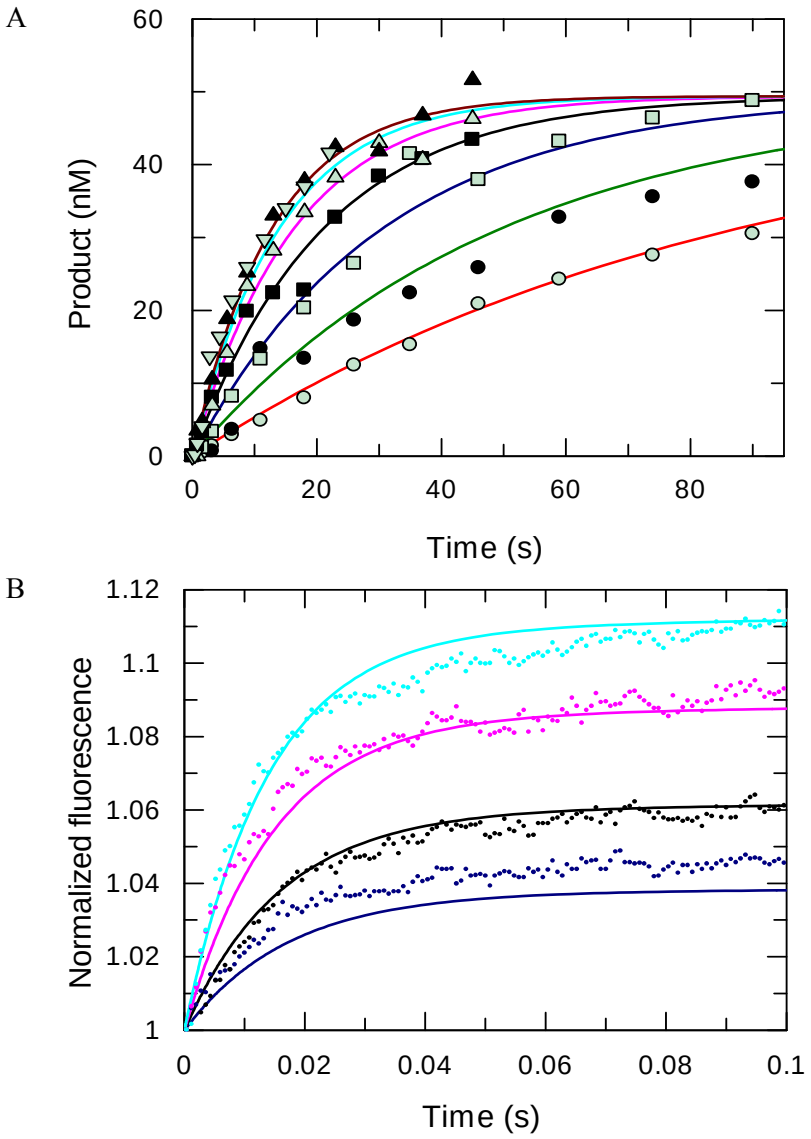


Figure 6.2: Global data fitting of dCTPαS incorporation for K522A mutant

(A) The series of data points were from quench flow experiments. The colored lines show results of fitting to the mechanism shown in Scheme 6.2. (B) The series of data points represented the fluorescence changes from the stopped-flow experiments. The solid lines show results of fitting to the mechanism shown in Scheme 6.2.

$$E + D_n \underset{0.2 s^{-1}}{\overset{10 \mu M^{-1} s^{-1}}{\rightleftharpoons}} ED_n$$

$$ED_n + N \overset{270 \mu M}{\rightleftharpoons} ED_n N \underset{31 s^{-1}}{\overset{35 s^{-1}}{\rightleftharpoons}} FD_n N \overset{0.7 s^{-1}}{\rightleftharpoons} ED_{n+1} + PP_i$$

$$ED_{n+1} \underset{10 \mu M^{-1} s^{-1}}{\overset{0.2 s^{-1}}{\rightleftharpoons}} E + D_{n+1}$$

Scheme 6.1: dCTP incorporation reaction pathway for K522A mutant

$$E + D_n \underset{0.2 s^{-1}}{\overset{10 \mu M^{-1} s^{-1}}{\rightleftharpoons}} ED_n$$

$$ED_n + N \overset{146 \mu M}{\rightleftharpoons} ED_n N \underset{53 s^{-1}}{\overset{30 s^{-1}}{\rightleftharpoons}} FD_n N \overset{0.24 s^{-1}}{\rightleftharpoons} ED_{n+1} + PP_i$$

$$ED_{n+1} \underset{10 \mu M^{-1} s^{-1}}{\overset{0.2 s^{-1}}{\rightleftharpoons}} E + D_{n+1}$$

Scheme 6.2: dCTPαS incorporation reaction pathway for K522A mutant

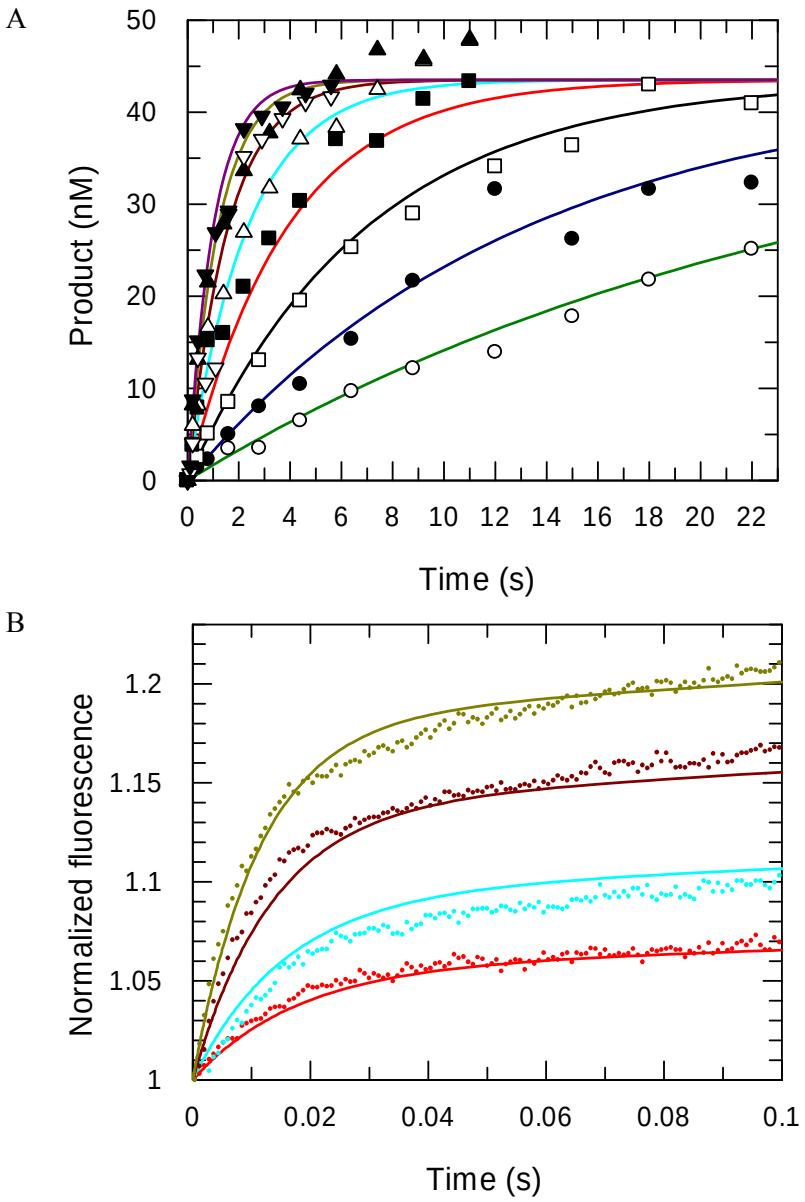


Figure 6.3: Global data fitting of dCTP incorporation for K522H mutant

(A) The series of data points were from quench flow experiments. The colored lines show results of fitting to the mechanism shown in Scheme 6.3. (B) The series of data points represented the fluorescence changes from the stopped-flow experiments. The solid lines show results of fitting to the mechanism shown in Scheme 6.3.

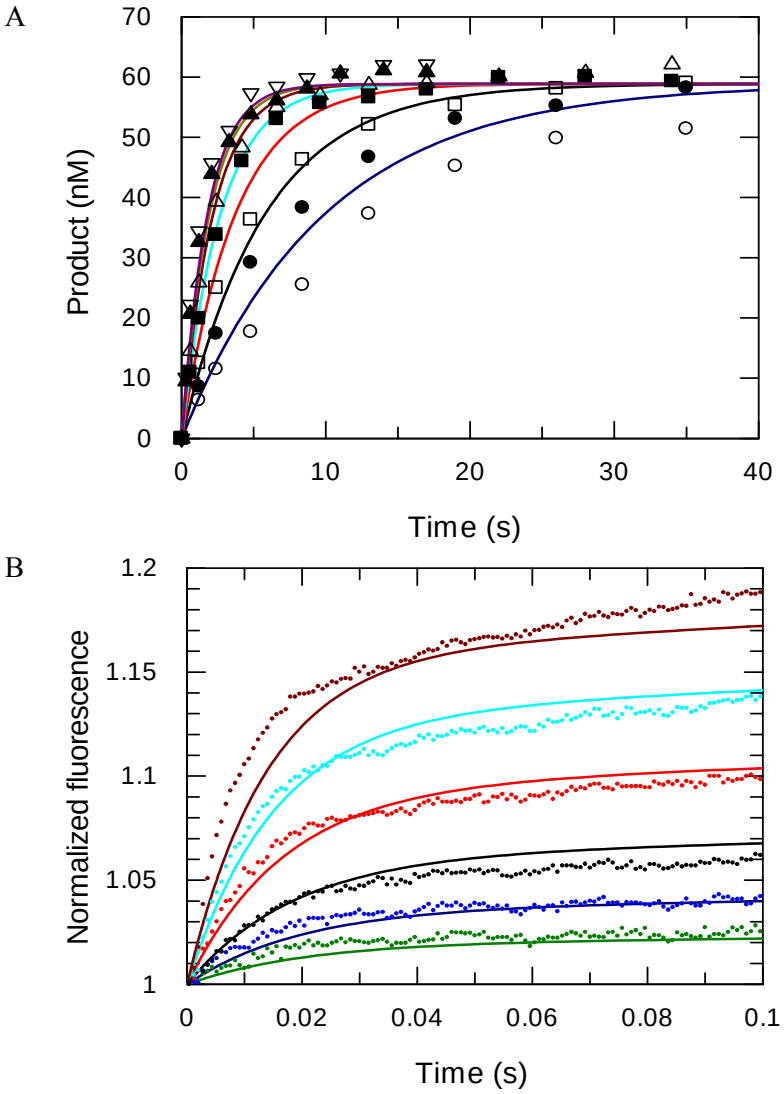


Figure 6.4: Global data fitting of dCTPαS incorporation for K522H mutant

(A) The series of data points were from quench flow experiments. The colored lines show results of fitting to the mechanism shown in Scheme 6.4. (B) The series of data points represented the fluorescence changes from the stopped-flow experiments. The solid lines show results of fitting to the mechanism shown in Scheme 6.4.

$$E + D_n \underset{0.2 s^{-1}}{\overset{10 \mu M^{-1} s^{-1}}{\rightleftarrows}} ED_n$$

$$ED_n + N \overset{584 \mu M}{\rightleftarrows} ED_n N \underset{44 s^{-1}}{\overset{82 s^{-1}}{\rightleftarrows}} FD_n N \overset{1.8 s^{-1}}{\rightleftarrows} ED_{n+1} + PP_i$$

$$ED_{n+1} \underset{10 \mu M^{-1} s^{-1}}{\overset{0.2 s^{-1}}{\rightleftarrows}} E + D_{n+1}$$

Scheme 6.3: dCTP incorporation reaction pathway for K522H mutant

$$E + D_n \underset{0.2 s^{-1}}{\overset{10 \mu M^{-1} s^{-1}}{\rightleftarrows}} ED_n$$

$$ED_n + N \overset{132 \mu M}{\rightleftarrows} ED_n N \underset{46 s^{-1}}{\overset{39 s^{-1}}{\rightleftarrows}} FD_n N \overset{1.35 s^{-1}}{\rightleftarrows} ED_{n+1} + PP_i$$

$$ED_{n+1} \underset{10 \mu M^{-1} s^{-1}}{\overset{0.2 s^{-1}}{\rightleftarrows}} E + D_{n+1}$$

Scheme 6.4: dCTPαS incorporation reaction pathway for K522H mutant

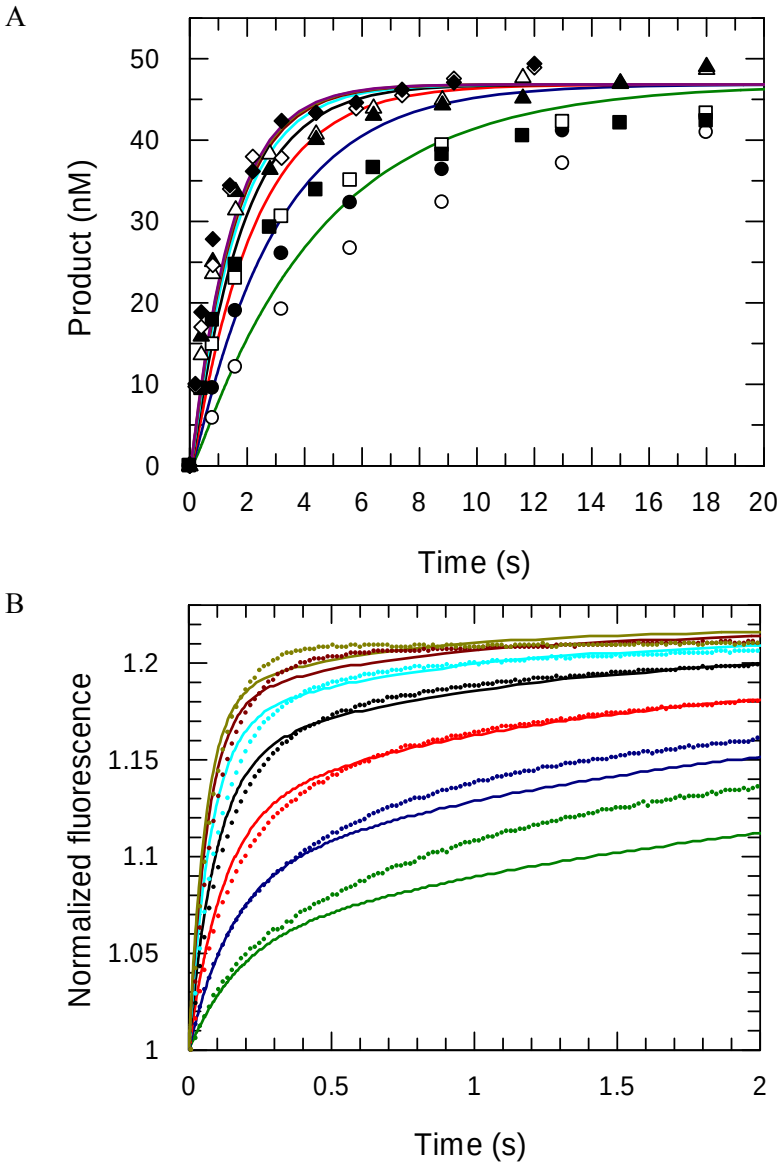


Figure 6.5: Global data fitting of dCTP incorporation for H506A mutant

(A) The series of data points were from quench flow experiments. The colored lines show results of fitting to the mechanism shown in Scheme 6.5. (B) The series of data points represented the fluorescence changes from the stopped-flow experiments. The solid lines show results of fitting to the mechanism shown in Scheme 6.5.

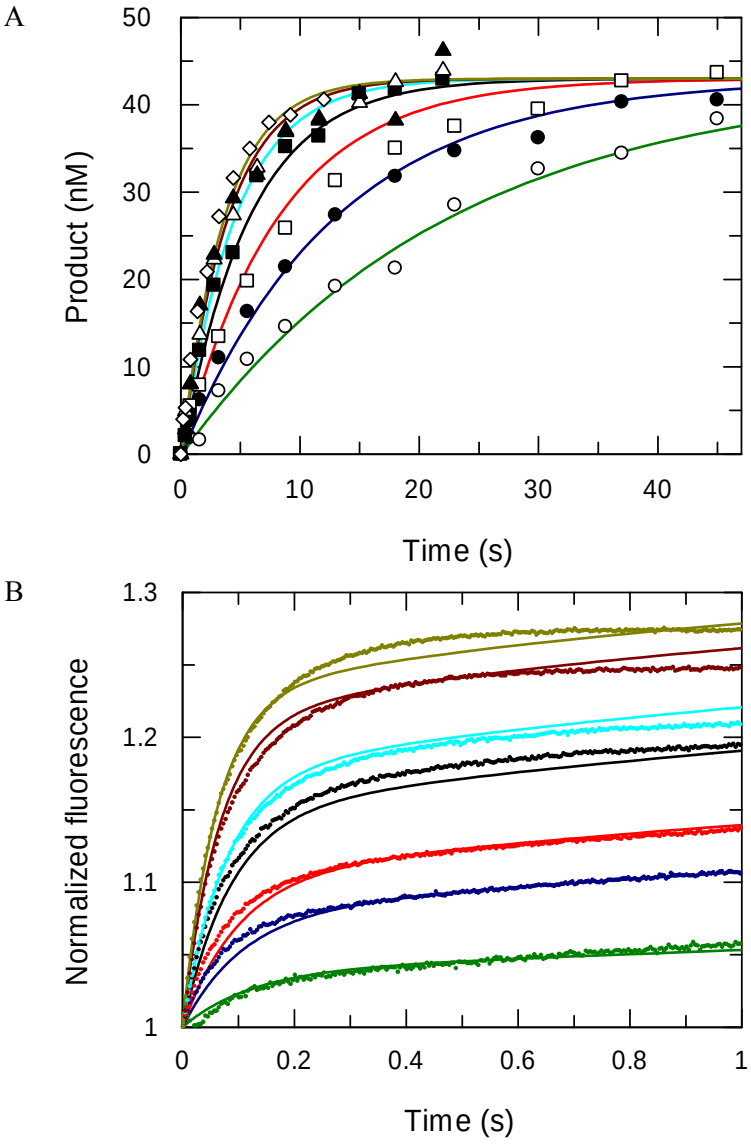


Figure 6.6: Global data fitting of dCTPαS incorporation for H506A mutant

(A) The series of data points were from quench flow experiments. The colored lines show results of fitting to the mechanism shown in Scheme 6.6. (B) The series of data points represented the fluorescence changes from the stopped-flow experiments. The solid lines show results of fitting to the mechanism shown in Scheme 6.6.

$$E + D_n \underset{0.2 s^{-1}}{\overset{10 \mu M^{-1} s^{-1}}{\rightleftarrows}} ED_n$$

$$ED_n + N \overset{70 \mu M}{\rightleftarrows} ED_n N \underset{3.5 s^{-1}}{\overset{15 s^{-1}}{\rightleftarrows}} FD_n N \overset{0.9 s^{-1}}{\rightleftarrows} ED_{n+1} + PP_i$$

$$ED_{n+1} \underset{10 \mu M^{-1} s^{-1}}{\overset{0.2 s^{-1}}{\rightleftarrows}} E + D_{n+1}$$

Scheme 6.5: dCTP incorporation reaction pathway for H506A mutant

$$E + D_n \underset{0.2 s^{-1}}{\overset{10 \mu M^{-1} s^{-1}}{\rightleftarrows}} ED_n$$

$$ED_n + N \overset{100 \mu M}{\rightleftarrows} ED_n N \underset{7 s^{-1}}{\overset{11 s^{-1}}{\rightleftarrows}} FD_n N \overset{0.5 s^{-1}}{\rightleftarrows} ED_{n+1} + PP_i$$

$$ED_{n+1} \underset{10 \mu M^{-1} s^{-1}}{\overset{0.2 s^{-1}}{\rightleftarrows}} E + D_{n+1}$$

Scheme 6.6: dCTPαS incorporation reaction pathway for H506A mutant

Reference

1. **Fersht, A.** 1998. Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding. W. H. Freeman and Company, New York.

2. **Patel, S. S., I. Wong, and K. A. Johnson**. 1991. Pre-steady-state kinetic analysis of processive DNA replication including complete characterization of an exonuclease-deficient mutant. Biochemistry **30**:511-525.

3. **Kunkel, T. A.** 2003. Considering the cancer consequences of altered DNA polymerase function. Cancer Cell **3**:105-110.

4. **Trifunovic, A., A. Wredenberg, M. Falkenberg, J. N. Spelbrink, A. T. Rovio, C. E. Bruder, Y. Bohlooly, S. Gidlof, A. Oldfors, R. Wibom, J. Tornell, H. T. Jacobs, and N. G. Larsson**. 2004. Premature ageing in mice expressing defective mitochondrial DNA polymerase. Nature **429**:417-423.

5. **Lee, H., J. Hanes, and K. A. Johnson**. 2003. Toxicity of nucleoside analogues used to treat AIDS and the selectivity of the mitochondrial DNA polymerase. Biochemistry **42**:14711-14719.

6. **Ito, J. and D. K. Braithwaite**. 1991. Compilation and alignment of DNA polymerase sequences. Nucl.Acids Res. **19**:4045-4057.

7. **Filee, J., P. Forterre, T. Sen-Lin, and J. Laurent**. 2002. Evolution of DNA polymerase families: Evidences for multiple gene exchange between cellular and viral proteins. Journal of Molecular Evolution **54**:763-773.

8. **McCulloch, S. D. and T. A. Kunkel**. 2008. The fidelity of DNA synthesis by eukaryotic replicative and translesion synthesis polymerases. Cell Res **18**:148-161.

9. **Ollis, D. L., P. Brick, R. Hamlin, N. G. Xuong, and T. A. Steitz**. 1985. Structure of large fragment of Escherichia coli DNA polymerase I complexed with dTMP. Nature **313**:762-766.

10. **Kohlstaedt, L. A., J. Wang, J. M. Friedman, P. A. Rice, and T. A. Steitz**. 1992. Crystal structure at 3.5 A resolution of HIV-1 reverse transcriptase complexed with an inhibitor. Science **256**:1783-1790.

11. **Kim, Y., S. H. Eom, J. Wang, D. S. Lee, S. W. Suh, and T. A. Steitz**. 1995. Crystal structure of Thermus aquaticus DNA polymerase. Nature **376**:612-616.

12. **Wang, J., A. K. M. A. Sattar, C. C. Wang, J. D. Karam, W. H. Konigsberg, and T. A. Steitz**. 1997. Crystal structure of a pol alpha family replication DNA polymerase from bacteriophage RB69. Cell **89**:1087-1099.

13. **Kiefer, J. R., C. Mao, C. J. Hansen, S. L. Basehore, H. H. Hogrefe, J. C. Braman, and L. S. Beese**. 1997. Crystal structure of a thermostable Bacillus DNA polymerase I large fragment at 2.1 angstrom resolution. Structure **5**:95-108.

14. **Doublie, S., S. Tabor, A. M. Long, C. C. Richardson, and T. Ellenberger**. 1998. Crystal structure of a bacteriophage T7 DNA replication complex at 2.2 angstrom resolution. Nature **391**:251-258.

15. **Silvian, L. F., E. A. Toth, P. Pham, M. F. Goodman, and T. Ellenberger**. 2001. Crystal structure of a DinB family error-prone DNA polymerase from Sulfolobus solfataricus. Nature Structural Biology **8**:984-989.

16. **Ling, H., F. Boudsocq, R. Woodgate, and W. Yang**. 2001. Crystal structure of a Y-family DNA polymerase in action: A mechanism for error-prone and lesion-bypass replication. Cell **107**:91-102.

17. **Uijon, S. N., R. E. Johnson, T. A. Edwards, S. Prakash, L. Prakash, and A. K. Aggarwal**. 2004. Crystal structure of the catalytic core of human DNA polymerase kappa. Structure **12**:1395-1404.

18. **Lamers, M. H., R. E. Georgescu, S. G. Lee, M. O'Donnell, and J. Kuriyan**. 2006. Crystal structure of the catalytic alpha subunit of E. coli replicative DNA polymerase III. Cell **126**:881-892.

19. **Beese, L. S., J. M. Friedman, and T. A. Steitz**. 1993. Crystal structures of the Klenow fragment of DNA polymerase I complexed with deoxynucleoside triphosphate and pyrophosphate. Biochemistry **32**:14095-14101.

20. **Pelletier, H., M. R. Sawaya, W. Wolfle, S. H. Wilson, and J. Kraut**. 1996. Crystal structures of human DNA polymerase beta complexed with DNA: Implications for catalytic mechanism, processivity, and fidelity. Biochemistry **35**:12742-12761.

21. **Sawaya, M. R., R. Prasad, S. H. Wilson, J. Kraut, and H. Pelletier**. 1997. Crystal Structures of Human DNA Polymerase Beta Complexed with Gapped and Nicked DNA: Evidence for an Induced Fit Mechanism. Biochemistry **36**:11205-11215.

22. **Li, Y., S. Korolev, and G. Waksman**. 1998. Crystal structures of open and closed forms of binary and ternary complexes of the large fragment of Thermus aquaticus DNA polymerase I: structural basis for nucleotide incorporation. Embo Journal **17**:7514-7525.

23. **Sawaya, M. R., H. Pelletier, A. Kumar, S. H. Wilson, and J. Kraut**. 1994. Crystal-Structure of Rat Dna-Polymerase-Beta - Evidence for A Common Polymerase Mechanism. Science **264**:1930-1935.

24. **Korolev, S., M. Nayal, W. M. Barnes, E. Dicera, and G. Waksman**. 1995. Crystal-Structure of the Large Fragment of Thermus-Aquaticus Dna-Polymerase-I at 2.5-Angstrom Resolution - Structural Basis for Thermostability. Proceedings of the National Academy of Sciences of the United States of America **92**:9264-9268.

25. **Beese, L. S., V. Derbyshire, and T. A. Steitz**. 1993. Structure of DNA polymerase I Klenow fragment bound to duplex DNA. Science **260**:352-355.

26. **Eom, S. H., J. Wang, and T. A. Steitz**. 1996. Structure of Taq polymerase with DNA at the polymerase active site. Nature **382**:278-281.

27. **Garcia-Diaz, M., K. Bebenek, G. Gao, L. C. Pedersen, R. E. London, and T. A. Kunkel**. 2005. Structure-function studies of DNA polymerase lambda. DNA Repair **4**:1358-1367.

28. **Johnson, S. J. and L. S. Beese**. 2004. Structures of Mismatch Replication Errors Observed in a DNA Polymerase. Cell **116**:803-816.

29. **Pelletier, H., M. R. Sawaya, A. Kumar, S. H. Wilson, and J. Kraut**. 1994. Structures of Ternary Complexes of Rat DNA-Polymerase-Beta, A DNA Template-Primer, and ddCTP. Science **264**:1891-1903.

30. **Yang, W., J. Y. Lee, and M. Nowotny**. 2006. Making and breaking nucleic acids: Two-Mg^{2+}-ion catalysis and substrate specificity. Molecular Cell **22**:5-13.

31. **Steitz, T. A. and J. A. Steitz**. 1993. A General 2-Metal-Ion Mechanism for Catalytic RNA. Proceedings of the National Academy of Sciences of the United States of America **90**:6498-6502.

32. **Joyce, C. M. and T. A. Steitz**. 1994. Function and Structure Relationships in DNA Polymerases. Annual Review of Biochemistry **63**:777-822.

33. **Steitz, T. A.** 1999. DNA polymerases: structural diversity and common mechanisms. J Biol Chem **274**:17395-17398.

34. **Steitz, T. A.** 1993. Dna-Dependent and Rna-Dependent Dna-Polymerases. Current Opinion in Structural Biology **3**:31-38.

35. **Doublie, S., M. R. Sawaya, and T. Ellenberger**. 1999. An open and closed case for all polymerases. Structure with Folding & Design **7**:R31-R35.

36. **Johnson, K. A.** 1992. Transient-state kinetic analysis of enzyme reaction pathways. The Enzyme **XX**:1-61.

37. **Johnson, K. A.** 1993. Conformational coupling in DNA polymerase fidelity. Annu.Rev.Biochem. **62**:685-713.

38. **Johnson, K. A.** 1995. Rapid quench kinetic analysis of polymerases, adenosinetriphosphatases, and enzyme intermediates. Methods Enzymol. **249**:38-61.

39. **Wong, I., S. S. Patel, and K. A. Johnson**. 1991. An induced-fit kinetic mechanism for DNA replication fidelity: direct measurement by single-turnover kinetics. Biochemistry **30**:526-537.

40. **Kuchta, R. D., V. Mizrahi, P. A. Benkovic, K. A. Johnson, and S. J. Benkovic**. 1987. Kinetic mechanism of DNA polymerase I (Klenow). Biochemistry **26**:8410-8417.

41. **Mizrahi, V., R. N. Henrie, J. F. Marlier, K. A. Johnson, and S. J. Benkovic**. 1985. Rate-limiting steps in the DNA polymerase I reaction pathway. Biochemistry **24**:4010-4018.

42. **Kati, W. M., K. A. Johnson, L. F. Jerva, and K. S. Anderson**. 1992. Mechanism and fidelity of HIV reverse transcriptase. J Biol Chem **267**:25988-25997.

43. **Johnson, A. A. and K. A. Johnson**. 2001. Fidelity of nucleotide incorporation by human mitochondrial DNA polymerase. J Biol Chem **276**:38090-38096.

44. **Ahn, J., B. G. Werneburg, and M. D. Tsai**. 1997. DNA Polymerase beta: Structure-Fidelity Relationship from Pre-Steady-State Kinetic Analyses of All Possible Correct and Incorrect Base Pairs for Wild Type and R283A Mutant. Biochemistry **36**:1100-1107.

45. **Carroll, S. S., M. Cowart, and S. J. Benkovic**. 1991. A mutant of DNA polymerase I (Klenow fragment) with reduced fidelity. Biochemistry **30**:804-813.

46. **Fiala, K. A. and Z. Suo**. 2004. Mechanism of DNA polymerization catalyzed by Sulfolobus solfataricus P2 DNA polymerase IV. Biochemistry **43**:2116-2125.

47. **Showalter, A. K. and M. D. Tsai**. 2002. A reexamination of the nucleotide incorporation fidelity of DNA polymerases. Biochemistry **41**:10571-10576.

48. **Herschlag, D., J. A. Piccirilli, and T. R. Cech**. 1991. Ribozyme-catalyzed and nonenzymatic reactions of phosphate diesters: rate effects upon substitution of sulfur for a nonbridging phosphoryl oxygen atom. Biochemistry **30**:4844-4854.

49. **Bakhtina, M., S. Lee, Y. Wang, C. Dunlap, B. Lamarche, and M. D. Tsai**. 2005. Use of viscogens, dNTP alpha S, and rhodium(III) as probes in stopped-flow, experiments to obtain new evidence for the mechanism of catalysis by DNA polymerase beta. Biochemistry **44**:5177-5187.

50. **Bakhtina, M., M. P. Roettger, S. Kumar, and M. D. Tsai**. 2007. A unified kinetic mechanism applicable to multiple DNA polymerases. Biochemistry **46**:5463-5472.

51. **Dunlap, C. A. and M. D. Tsai**. 2002. Use of 2-Aminopurine and Tryptophan Fluorescence as Probes in Kinetic Analyses of DNA Polymerase beta. Biochemistry **41**:11226-11235.

52. **Purohit, V., N. D. F. Grindley, and C. M. Joyce**. 2003. Use of 2-aminopurine fluorescence to examine conformational changes during nucleotide incorporation by DNA polymerase I (Klenow fragment). Biochemistry **42**:10200-10211.

53. **Rothwell, P. J., V. Mitaksov, and G. Waksman**. 2005. Motions of the Fingers Subdomain of Klentaq1 Are Fast and Not Rate Limiting: Implications for the Molecular Basis of Fidelity in DNA Polymerases. Molecular Cell **19**:345-355.

54. **Zhong, X. J., S. S. Patel, B. G. Werneburg, and M. D. Tsai**. 1997. DNA polymerase beta: Multiple conformational changes in the mechanism of catalysis. Biochemistry **36**:11891-11900.

55. **Yang, L. J., W. A. Beard, S. H. Wilson, S. Broyde, and T. Schlick**. 2002. Polymerase beta simulations suggest that Arg258 rotation is a slow step rather than large subdomain motions per se. Journal of Molecular Biology **317**:651-671.

56. **Brieba, L. G., B. F. Eichman, R. J. Kokoska, S. Doublie, T. A. Kunkel, and T. Ellenberger**. 2004. Structural basis for the dual coding potential of 8-oxoguanosine by a high-fidelity DNA polymerase. EMBO J **23**:3452-3461.

57. **Brieba, L. G., R. J. Kokoska, K. Bebenek, T. A. Kunkel, and T. Ellenberger**. 2005. A Lysine Residue in the Fingers Subdomain of T7 DNA Polymerase Modulates the Miscoding Potential of 8-Oxo-7,8-Dihydroguanosine. Structure **13**:1653-1659.

58. **Dutta, S., Y. Li, D. Johnson, L. Dzantiev, C. C. Richardson, L. J. Romano, and T. Ellenberger**. 2004. Crystal structures of 2-acetylaminofluorene and 2-aminofluorene in complex with T7 DNA polymerase reveal mechanisms of mutagenesis. Proceedings of the National Academy of Sciences of the United States of America **101**:16186-16191.

59. **Li, Y., S. Dutta, S. Doublie, H. M. Bdour, J. S. Taylor, and T. Ellenberger**. 2004. Nucleotide insertion opposite a cis-syn thymine dimer by a replicative DNA polymerase from bacteriophage T7. Nat Struct Mol Biol **11**:784-790.

60. **Dunn, J. J., F. W. Studier, and M. Gottesman**. 1983. Complete nucleotide sequence of bacteriophage T7 DNA and the locations of T7 genetic elements. Journal of Molecular Biology **166**:477-535.

61. **Mark, D. F. and C. C. Richardson**. 1976. Escherichia coli thioredoxin: a subunit of bacteriophage T7 DNA polymerase. Proc.Natl.Acad.Sci U.S.A **73**:780-784.

62. **Modrich, P. and C. C. Richardson**. 1975. Bacteriophage T7 deoxyribonucleic acid replication invitro. Bacteriophage T7 DNA polymerase: an an emzyme composed of phage- and host-specific subunits. J.Biol.Chem. **250**:5515-5522.

63. **Modrich, P. and C. C. Richardson**. 1975. Bacteriophage T7 Deoxyribonucleic acid replication in vitro. A protein of Escherichia coli required for bacteriophage T7 DNA polymerase activity. J.Biol.Chem. **250**:5508-5514.

64. **Dyson, H. J., M. F. Jeng, L. L. Tennant, I. Slaby, M. Lindell, D. S. Cui, S. Kuprin, and A. Holmgren**. 1997. Effects of Buried Charged Groups on Cysteine Thiol Ionization and Reactivity in Escherichia coli Thioredoxin: Structural and Functional Characterization of Mutants of Asp 26 and Lys 57. Biochemistry **36**:2622-2636.

65. **Singha, N. C., A. Vlamis-Gardikas, and A. Holmgren**. 2003. Real-time Kinetics of the Interaction between the Two Subunits, Escherichia coli Thioredoxin and Gene 5 Protein of Phage T7 DNA Polymerase. J.Biol.Chem. **278**:21421-21428.

66. **Huber, H. E., S. Tabor, and C. C. Richardson**. 1987. Escherichia coli thioredoxin stabilizes complexes of bacteriophage T7 DNA polymerase and primed templates. J.Biol.Chem. **262**:16224-16232.

67. **Tabor, S., H. E. Huber, and C. C. Richardson**. 1987. Escherichia coli thioredoxin confers processivity on the DNA polymerase activity of the gene 5 protein of bacteriophage T7. J Biol Chem **262**:16212-16223.

68. **Donlin, M. J., S. S. Patel, and K. A. Johnson**. 1991. Kinetic partitioning between the exonuclease and polymerase sites in DNA error correction. Biochemistry **30**:538-546.

69. **Tsai, Y. C. and K. A. Johnson**. 2006. A new paradigm for DNA polymerase specificity. Biochemistry **45**:9675-9687.

70. **Johnson, K. A.** 2008. Role of induced-fit in enzyme specificity: a molecular forward/reverse switch. J.Biol.Chem.in press.

71. **Johnson, K. A.** 1992. Conformational Coupling in Dna-Polymerase Information-Transfer. Philosophical Transactions of the Royal Society of London Series B-Biological Sciences **336**:107-112.

72. **Brune, M., J. L. Hunter, J. E. Corrie, and M. R. Webb**. 1994. Direct, real-time measurement of rapid inorganic phosphate release using a novel fluorescent probe and its application to actomyosin subfragment 1 ATPase. Biochemistry **33**:8262-8271.

73. **Hirshberg, M., K. Henrick, L. L. Haire, N. Vasisht, M. Brune, J. E. T. Corrie, and M. R. Webb**. 1998. Crystal structure of phosphate binding protein labeled with a coumarin fluorophore, a probe for inorganic phosphate. Biochemistry **37**:10381-10385.

74. **Koshland, D. E.** 1958. Application of a Theory of Enzyme Specificity to Protein Synthesis. Proc.Natl.Acad.Sci U.S.A **44**:98-104.

75. **Herschlag, D.** 1988. The role of induced fit and conformational changes of enzymes in specificity and catalysis. Bioorganic Chemistry **16**:62-96.

76. **Johnson, K. A.** 1998. Advances in transient-state kinetics. Curr.Opin.Biotechnol. **9**:87-89.

77. **Astatke, M., N. D. F. Grindley, and C. M. Joyce**. 1995. Deoxynucleoside Triphosphate and Pyrophosphate Binding Sites in the Catalytically Competent Ternary Complex for the Polymerase Reaction Catalyzed by DNA Polymerase I (Klenow Fragment). J.Biol.Chem. **270**:1945-1954.

78. **Polesky, A. H., M. E. Dahlberg, S. J. Benkovic, N. D. Grindley, and C. M. Joyce**. 1992. Side chains involved in catalysis of the polymerase reaction of DNA polymerase I from Escherichia coli. J.Biol.Chem. **267**:8417-8428.

79. **Minnick, D. T., K. Bebenek, W. P. Osheroff, R. M. Turner, Jr., M. Astatke, L. Liu, T. A. Kunkel, and C. M. Joyce**. 1999. Side Chains That Influence Fidelity at the Polymerase Active Site of Escherichia coli DNA Polymerase I (Klenow Fragment). J.Biol.Chem. **274**:3067-3075.

80. **Herschlag, D. and W. P. Jencks**. 1989. Phosphoryl Transfer to Anionic Oxygen Nucleophiles - Nature of the Transition-State and Electrostatic Repulsion. Journal of the American Chemical Society **111**:7587-7596.

81. **Hengge, A. C. and W. W. Cleland**. 1990. Direct Measurement of Transition-State Bond-Cleavage in Hydrolysis of Phosphate-Esters of P-Nitrophenol. Journal of the American Chemical Society **112**:7421-7422.

82. **Zalatan, J. G., I. Catrina, R. Mitchell, P. K. Grzyska, P. J. O'Brien, D. Herschlag, and A. C. Hengge**. 2007. Kinetic isotope effects for alkaline phosphatase reactions: Implications for the role of active-site metal ions in catalysis. Journal of the American Chemical Society **129**:9789-9798.

83. **Zalatan, J. G. and D. Herschlag**. 2006. Alkaline phosphatase mono- and diesterase reactions: Comparative transition state analysis. Journal of the American Chemical Society **128**:1293-1303.

84. Collaborative Computational Project, No. 4. The CCP4 Suite: Programs for Protein Crystallography. 1994. Acta Cryst. D50, 760-763.

Vita

Zhinan Jin was born in 1972 in Shanghai, People's Republic of China. After completing his study at Minhang High School in Shanghai in July 1991, he attended Fudan University in Shanghai and received the degree of Bachelor of Science in Genetics in July 1996. During the following years, He worked as a research assistant in Shanghai Institute of Pharmaceutical Industry. In August 2001, he entered the department of Biochemistry in University of Houston for his graduate studies. In June 2002, he joined the graduate school at the University of Texas at Austin and started his graduate studies in Dr. Kenneth Johnson's lab.

Permanent address: 100 Haifang Road, Bldg 17, Rm 203, Shanghai 200041, P. R. China

This dissertation was typed by the author.