The Dissertation Committee for Yu-Chih Tsai Certifies that this is the approved version of the following dissertation:

Kinetics of DNA Polymerase Conformational Changes during Nucleotide Binding and Incorporation

Committee:

Kenneth A. Johnson, Supervisor

Andrew Ellington

David W. Hoffman

Henry R. Bose

Robert M. Krug

Kinetics of DNA Polymerase Conformational Changes during Nucleotide Binding and Incorporation

by

Yu-Chih Tsai, 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
May, 2005

Dedication

To my parents: Wen-Hsiung Tsai and Li-Hua Chen

Acknowledgments

I would like to thank my supervisor, Dr. Kenneth Johnson, who patiently taught and assisted me during my graduate work toward finishing this dissertation. I also want to thank Dr. JoAnn Johnson for providing me enormous support while I was mourning the loss of my dear mother.

I am also grateful to my committee members for their precious time and invaluable suggestions: Drs. Andrew Ellington, David Hoffman, Henry Bose, and Robert Krug.

I also want to thank the past and present Johnson lab members for discussions and help with lab work: Harold Lee, Jeremiah Hanes, Vi Dougherty, Zhinan Jin, Jamie Caras, Jarle Lillemoen, Jeff Bartron, and Louise Wang.

Finally, I want to express my greatest gratitude to my parents. Without the constant support from both of you, I wouldn't be able to achieve what I can today.

Kinetics of DNA Polymerase Conformational Changes during Nucleotide Binding and Incorporation

Yu-Chih Tsai, PhD

The University of Texas at Austin, 2005

Supervisor: Kenneth A. Johnson

DNA polymerases are enzymes responsible for replicating DNA molecules in cells. They have evolved to carry out replication efficiently and highly accurately. For example, the T7 DNA polymerase catalyzes nucleotide incorporation at a rate of 300 s^{-1} with an error frequency of only one per 10^5-10^6 nucleotides incorporated. Previous studies examining the underlying mechanisms for the exceptional nucleotide selectivity have suggested an induced-fit model based on the structural and kinetic evidence. The key feature of the model is that a rate-limiting conformational change step preceding the phosphoryl transfer step controls the incorporation reaction. If the incoming nucleotide is not correctly base-paired with the template base, the conformational change proceeds at a greatly reduced rate to prevent misincorporation. However, the lack of direct measurement of the kinetics of this rate-limiting conformational change step leads to questions whether the proposed model could account for the nucleotide selectivity.

I created a mutant T7 DNA polymerase labeled with an environmentally sensitive fluorophore at an amino acid residue in the recognition domain. The signal from this

enzyme is sensitive to the conformational state of the polymerase and was used for studying the kinetics of the conformational changes. Results from my kinetic studies suggested the conformational change step preceding the chemistry step is at least partially rate-limiting during correct nucleotide incorporation. When a mismatched nucleotide binds to the enzyme active site, the conformational change steps not only become slow and unfavorable but also lead to misalignment of active site residues to prevent incorporation of incorrect nucleotide. The new model for nucleotide selectivity is compatible with the previously established induced-fit mechanism, but it also provides new insights on how a small free energy difference between correct and incorrect nucleotide binding leads to the remarkable selectivity the polymerase can achieve. Furthermore, I demonstrated that the fluorescence signal from this novel T7 DNA polymerase construct can be utilized to detect point mutations in DNA sequences. Other than the T7 polymerase project, my attempts to obtain an active HCV RNA polymerase are also presented in this dissertation.

Table of Contents

x

List of Abbreviations

dCTPαS: 2'-deoxycytidine-5'-O-(1-thiotriphosphate)

ddNMP: Dideoxynucleoside monophosphate

ddNTP: Dideoxynucleoside triphosphate

DNA_{dd}: DNA duplex with a dideoxynucleotide terminated primer

dNTP: Deoxynucleoside triphosphate

DTT: 1,4-Dithio-DL-threitol

EDTA: Ethylenediaminetetraacetic acid

HCV: Hepatitis C virus

HEPES: 4-(2-Hydroxyethyl)piperazine-1-ethanesulfonic acid

IPTG: Isopropyl β-D-1-thiogalactopyranoside

IRES: Internal ribosome entry site

KF: Klenow fragment

MDCC: 7-diethylamino-3-((((2-maleimidyl)ethyl) amino)carbonyl)coumarin

NTP: Nucleoside triphosphate

NTPase: Nucleoside triphosphatase

PBP: Phosphate binding protein

PMSF: Phenylmethylsulfonyl fluoride

PNPase: Polynucleotide polymerase

RdRp: RNA dependent RNA polymerase

RT: Reverse transcriptase

scNS3-NS4A: Single-chain NS3-NS4A fusion protein

SNP: Single nucleotide polymorphism

ssDNA: Single-strand DNA

TLCK: N_α-Tosyl-L-lysine chloromethyl ketone hydrochloride

TNTase: Terminal transferase

TPCK: N-p-Tosyl-L-phenylalanine chloromethyl ketone

UTR: Untranslated region

List of Tables

List of Figures

xv

List of Illustrations

Chapter 1: Overview of Polymerase Mechanisms

1.1 CLASSIFICATION OF DNA POLYMERASES

One of the biggest question biologists have been trying to answer since the discovery of deoxyribonucleic acid (DNA) is how genetic information stored in this molecule can be faithfully replicated generation after generation. DNA polymerases are enzymes responsible for this important biological process and can be classified into at least five different families based on their amino acid sequence homologies (1) and structural similarities (2). Many well studied DNA polymerases belong to the DNA polymerase I (pol I) or A family DNA polymerase, which includes Klenow fragment of *E. coli* Pol I (KF fragment), *T. aquaticus* (Taq) DNA polymerase, and T7 DNA polymerase (3). These enzymes are single subunit proteins and represent the simplest molecular machinery to carry out DNA replication. The DNA polymerase α (pol α) or B family DNA polymerase is consisting of all eukaryotic replicating DNA polymerases and the polymerases from bacteriophage T4 and RB69. A crystal structure of RB69 polymerase shows some similarities to the pol I family enzymes (4). The third family contains reverse transcriptase (RT), RNA-dependent RNA polymerases (RdRp), and telomerase. These RNA-dependent polymerases also show structural similarities to the pol I enzymes. However, the structure of DNA polymerase β, which is a member of another family of the DNA-dependent DNA polymerases, does not show structural resemblance to any of these previous families (5). The last family consists of bacterial DNA polymerase III enzymes based only on the amino acid sequence homologies. Despite the differences in sequence homologies and three-dimensional domain arrangements, the overall shape of DNA polymerases with known structures show significant similarities. Their overall shape can be described as a right hand with fingers,

palm, and thumb domains. Although the fingers and thumb domains are different between DNA polymerase families, the palm domain appears to be more highly conserved (6). In addition, all crystal structures of DNA polymerases suggest a common two-metal ion mechanism for the phosphoryl transfer reaction (7;8). Because the catalytic mechanism is highly preserved in different families of DNA polymerases, conclusions drawn form biochemical studies using relatively simple pol I enzymes have served as a valid model for describing DNA polymerase mechanisms in general.

1.2   THE FIDELITY OF DNA REPLICATION

The replication of DNA is achieved with a semiconservative mechanism in which two original strands of DNA are separated, and each acts as a template for DNA synthesis. DNA polymerases catalyze the incorporation of nucleotides in a template-directed manner, and their specificity changes for every incoming nucleotide according to the Watson-Crick base-pairing rules. In order to perform this task, DNA polymerases have evolved to efficiently select for correct nucleotides in each incorporation cycle. For example, the T7 DNA polymerase catalyzes nucleotide incorporation at a rate of 300 s^{-1} and makes one error out of every 10^5 to 10^6 nucleotides incorporated (9). When the incorporation of mismatched nucleotide occurs, the rate of reaction is at least 10^4-fold slower than correct incorporation and the polymerase stalls and utilizes its 3'-5' exonuclease activity to remove this misincorporated nucleotide. This exonuclease proof reading function provides an additional 10^3 to 10^4 fold accuracy for DNA replication (10). As a result, the overall error rate for T7 DNA polymerase is 10^{-8} to 10^{-9}. To achieve such a high fidelity during DNA replication, DNA polymerases can not just simply zip together deoxynucleoside triphosphate molecules based on the correctness of base-pairing. The differences in binding free energy for correct versus incorrect base pairs measured in solution is in the range of 1 to 3 kcal/mole (11-13) and could only afford an

error rate of one out of 5 to 150 bases by calculation. Therefore, the DNA polymerases must have a mechanism to differentiate incorrect and correct base pairs before the phosphoryl transfer reaction takes place. Theoretical studies suggest that multiple conformational changes contribute to the fidelity of DNA polymerization and theories have been postulated to explain the underlying mechanisms for the high-fidelity DNA replication (14).

1.3   THE INDUCED-FIT MODEL FOR NUCLEOTIDE SELECTIVITY

An important clue for understanding the underlying mechanisms by which small free energy differences are translated into nucleotide selectivity may come from structural studies of polymerases. Crystal structures of several DNA polymerases show a significant rearrangement of fingers domain when the E-DNA-dNTP ternary complexes are formed (15-18). This conformational change step is essential for bringing active site residues into optimal positions to catalyze the phosphoryl transfer reaction. Furthermore, ternary complexes showing a closed conformation are only observed in the presence of correctly base-paired nucleotides and suggest that the polymerase could somehow sense the differences between the correct and incorrect base pairs before or during the conformational change step to prevent misincorporation.

In order to delineate such a mechanism, the kinetics of each elementary step during DNA polymerization has to be defined. Kinetic studies using transient kinetic methods have established a complete reaction pathway of DNA polymerization for the T7 DNA polymerase (9) (Figure 1.1). The kinetic analysis revealed that a rate-limiting step occurs after the ground-state nucleotide binding and before the fast chemistry step in the reaction pathway of correct incorporation. Results form both structural and biochemical studies suggest that the rate-limiting step corresponds to the conformational change of the enzyme. However, the kinetics of this rate-limiting conformational change could not be

3

measured directly. The evidence that supports the conclusion came from the experiment using the α phosphothioate analogues (dNTPαS) of deoxynucleoside triphosphates (dNTP). In theory, the analogue would slow down the chemical reaction step (19-21). Therefore observing the full elemental effect in misincorporation but not during normal incorporation suggests that the chemical reaction is rate-limiting during misincorporation, but a conformational change preceding chemistry limits the rate of correct incorporation (22).
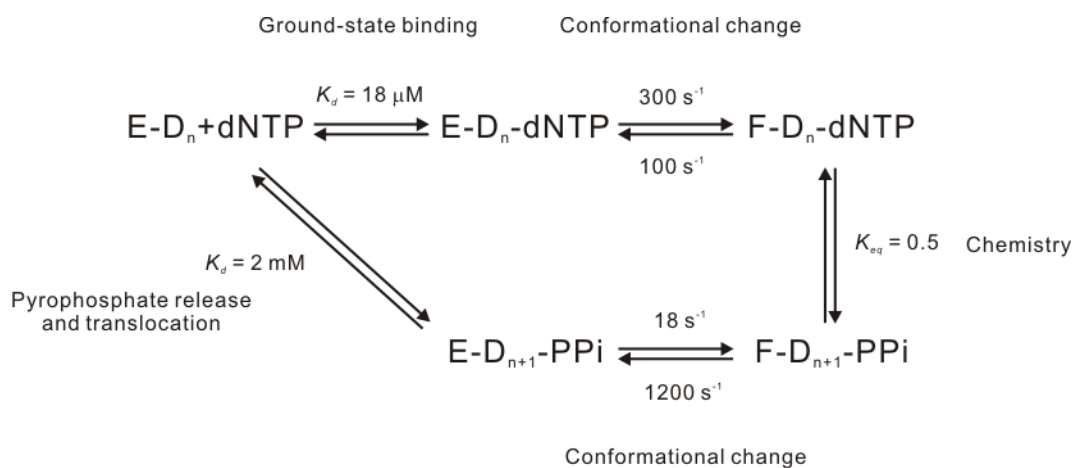
Figure 1.1:   Kinetic mechanism of DNA polymerization

The pathway of the DNA polymerization was determined using T7 DNA polymerase. The conformational change following ground-state binding is the rate-limiting step for the reaction pathway. Kinetic parameters are adopted from (9).

Kinetic data obtained from the T7 DNA polymerase (22) and the HIV-RT (23) suggest that a two-step binding mechanism is required for the observed nucleotide selectivity. The initial binding step is the ground-state binding, in which the nucleotide binds to the enzyme-DNA complex according to base pairing rules. At this stage, the enzyme is in its open conformation and the selectivity of nucleotides comes from the free energy differences in base pair formation. The initial binding of a correct nucleotide may induce a conformational change of the polymerase from the open to the closed state and result in tighter substrate binding. Also, the rearrangement of the active site residues provides optimal binding to the substrates and stabilizes the transition state for the subsequent phosphoryl transfer reaction to occur (24). It is suggested that in the closed conformation the enzyme may make contacts with both the incoming nucleotide and the template base from the minor groove side, allowing the polymerase to test for proper alignment of the base pair (25;26). In addition, the orientation of the template base changes during the conformational change of the polymerase and could possibly contribute to nucleotide selectivity (27). If the conformational change can not proceed normally due to a mismatch, the phosphoryl transfer reaction also slows down significantly. In this induced-fit model the rate-limiting conformational change provides the largest discrimination for incoming nucleotides.

However, it was argued that such a mechanism cannot enhance enzyme substrate selectivity any more than a simple one-step binding mechanism if the chemistry step is rate-limiting (28;29). A single-turnover stopped-flow fluorescence study with 2-aminopurine in the template DNA of correct nucleotide incorporation on pol β showed two detectable fluorescent transitions, one preceding and one following the chemistry step of nucleotide incorporation (30;31). The result suggests that there is no detectable rate-limiting conformational change step at least in the case of pol β although crystal

structures show a similar conformational change upon nucleotide binding (5;17). However, similar studies carried out with the Klenow fragment and HIV-RT lead to conclusions that conformational change is indeed rate-limiting (23;32). The validity of these conclusions is questionable since the signal arises due to positioning of the DNA. In order to further test the validity of the induced-fit model for nucleotide selectivity, the relationship of protein conformational change and its catalytic activity has to be established. I developed a method to directly measure the kinetics of conformational change step as described in chapter 3.

1.4   THE HEPATITIS C VIRUS RNA DEPENDENT RNA POLYMERASE

Hepatitis C virus (HCV) infection is the major cause of a chronic liver disease originally named as non-A, non-B hepatitis (33) and is becoming a significant medical problem (34). About 80% of infected patients fail to clear the virus after the acute infection stage and the virus infection persists for decades. The chronic hepatitis may progress to liver cirrhosis and patients have a high risk of developing hepatocellular carcinoma (35). In addition, the failure of humoral immunity to prevent re-infection and the diversity of the virus genome hamper the development of an effective HCV vaccine (36;37). The most effective antiviral therapies for HCV infection, involving combine treatment with interferon-α and ribavirin (Figure 1.2), have limited effectiveness. Only 40% of treated patients have sustained virological response (38;39).
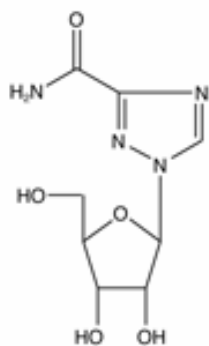
Figure 1.2: Ribavirin

HCV has similar genome organization and protein sequences to flaviviruses and pestiviruses. It is classified as a third genus of the *Flaviviridae* family (40) and consists of a single-stranded RNA genome of approximately 9.5 kb in length (41-43). The genome contains a 5' untranslated region (UTR) of 341 nucleotides (44;45), a large open reading frame encoding a single polyprotein (41-43), and a 3'UTR of variable length (46-48). The polyprotein is processed into structural (core, E1, and E2) and nonstructural (P7, NS2, NS3, NS4A, NS4B, NS5A, and NS5B) viral proteins. Among these viral proteins, NS5B is identified as an RNA dependent RNA polymerase (RdRp) and has been investigated extensively as one of the primary targets for antiviral drugs (49-52). The NS5B RdRp was shown to use either a self-primed (copy-back) or a *de novo* initiation mechanism to replicate RNA molecules which were revealed in *in vitro* studies (53-56). It can initiate RNA polymerization *de novo* from the HCV genome and heterologous viral RNA templates (53). Homopolymeric templates, poly(C) and poly(U), were also demonstrated to be replicated by NS5B in a primer-independent manner (54).

The expression and purification of soluble NS5B RdRp enzyme is problematic and requires the addition of salts, glycerol, and detergents (49). When 21 C-terminal hydrophobic amino acids of the protein were removed, the truncated NS5Bs was reported to show improved solubility and higher RdRp activities (50;51). Functional NS5B RdRp requires Mg^{2+} ion as a cofactor (52). Low concentration of Mn^{2+} ion was also found to support NS5B RdRp function and achieve even higher enzyme activities in many studies (50;54;57). Other than the inherent RdRp activity, a terminal transferase (TNTase) activity was detected with purified NS5B proteins. However, it may be a contaminating activity resulting from other proteins being co-purified with the NS5B (49). A high resolution crystal structure of the HCV NS5B RdRp shows that the enzyme has a typical polymerase structure resembling a right hand shape. However, extra loops and $\alpha-$helices

8

extended from the fingers domain connect it to the thumb domain to encircle the active site (58). This feature suggests that the NS5B RdRp structure is more rigid than other RNA-dependent polymerases, e.g. HIV-RT (18), and that the conformation of the enzyme does not change significantly during nucleotide incorporation cycles. The crystal structure of NS5B also reveals that a C-terminal loop extends into the RNA binding cleft and could potentially interfere with RNA binding (Figure 1.3A) (59).

The HCV replication is postulated to be carried out in two steps. The genome is first copied into (-) strand RNA. The (-) strand RNA then serves as a template for the synthesis of (+) strand RNA genome to complete the replication cycle. The 5'UTR has an internal ribosomal entry site (IRES) for initiating the translation of viral polyprotein (44;45) and it forms a secondary structure which is highly conserved between many pestiviruses (44). The HCV 3'UTR is approximately 200-300 bases long and consists of three elements: a short variable region, a poly(U)/(C) tract, and a highly conserved X-region at its 3' end. The X-region was predicted to form a conserved three stem-loop secondary structure (60). It has been speculated that the 3'UTR constitutes a *cis*-acting element that is necessary for directing *de novo* initiation of (-) strand RNA synthesis (46-48). Although the recognition domain for NS5B on the viral genome is not clearly defined, it is hypothesized that each UTRs provides a *cis*-acting element necessary for specific *de novo* initiation of RNA replication of each strand.

The composition of the HCV replication complex is not known. Previous studies showed that NS5B forms a complex with NS3 and NS4A (61) and the NS5B RdRp activity is positively modulated by the NS3 helicase (62). The NS3 is a multifunctional protein with a protease domain and a helicase domain (63). Its serine protease activity is important for viral polyprotein processing and requires NS4A as a cofactor (64;65). The RNA helicase activity and the RNA-stimulated nucleoside triphosphatase (NTPase)

activity are associated with the C-terminal helicase domain (63;66). To overcome the problem of expressing NS3/NS4A complex in eukaryotic systems, a fusion protein has been created (67;68). This single-chain NS3-NS4A (scNS3-NS4A) protein possesses physiological properties equivalent to those of the NS3/NS4A complex (67). Based on the structural study of the scNS3-NS4A, a model of polyprotein processing and formation of the replication complex was proposed (69). Chapter 2 describes my efforts to achieve active forms of the HCV RdRp enzyme.

1.5   THE T7 DNA POLYMERASE

The T7 DNA polymerase is a heterodimer consisting of the phage T7 gene 5 protein (80 kDa) and an *E. coli* host cofactor, thioredoxin (12 kDa) (70;71). The gene 5 protein has two enzyme activities, a 5'-3' DNA polymerization activity and a 3'-5' exonuclease activity. The exonuclease active site resides near the amino terminus, and the polymerase active site is located between the fingers and thumb domains toward the carboxyl terminus. *E. coli* thioredoxin has no catalytic activity on DNA substrates. The two component proteins form a 1:1 complex with an equilibrium dissociation constant of less than 5 nM (72). The increase in the DNA binding affinity of the gene 5 protein in the presence of thioredoxin transforms it from a distributive DNA polymerization enzyme to a highly processive enzyme (73;74), which can synthesize more than a thousand bases before dissociating from the DNA duplex in a single binding event. The T7 DNA polymerase is also a fast enzyme catalyzing DNA polymerization at a rate of 300 s^{-1} with an error frequency of 10^{-5}~10^{-6} (22). Including the extra fidelity provided by the exonuclease proof reading activity, the overall error frequency is only 10^{-8}~10^{-9} (10). Because T7 DNA polymerase is such a fast and accurate DNA polymerase responsible for genome replication, it serves as an important model for understanding high-fidelity DNA replication. The structure of T7 DNA polymerase has a typical right hand shape

with domains similar to other pol I family enzymes (Figure 1.3B). A 71-residue insertion in the tip of the thumb forms an extended loop that binds to thioredoxin. The 3'-5' exonuclease domain is at the N-terminus of the protein. Figure 1.2B shows the T7 DNA polymerase at its closed state with DNA substrate and ddGTP bound to its active site. Two Mg^{2+} ions in the active site of the polymerase suggest that the enzyme utilizes a two-metal mechanism to carry out the phosphoryl transfer reaction (25). Another Mg^{2+} ion is bound to the exonuclease domain in the crystal structure, although the exonuclease activity is thought to also require two-metal ions.
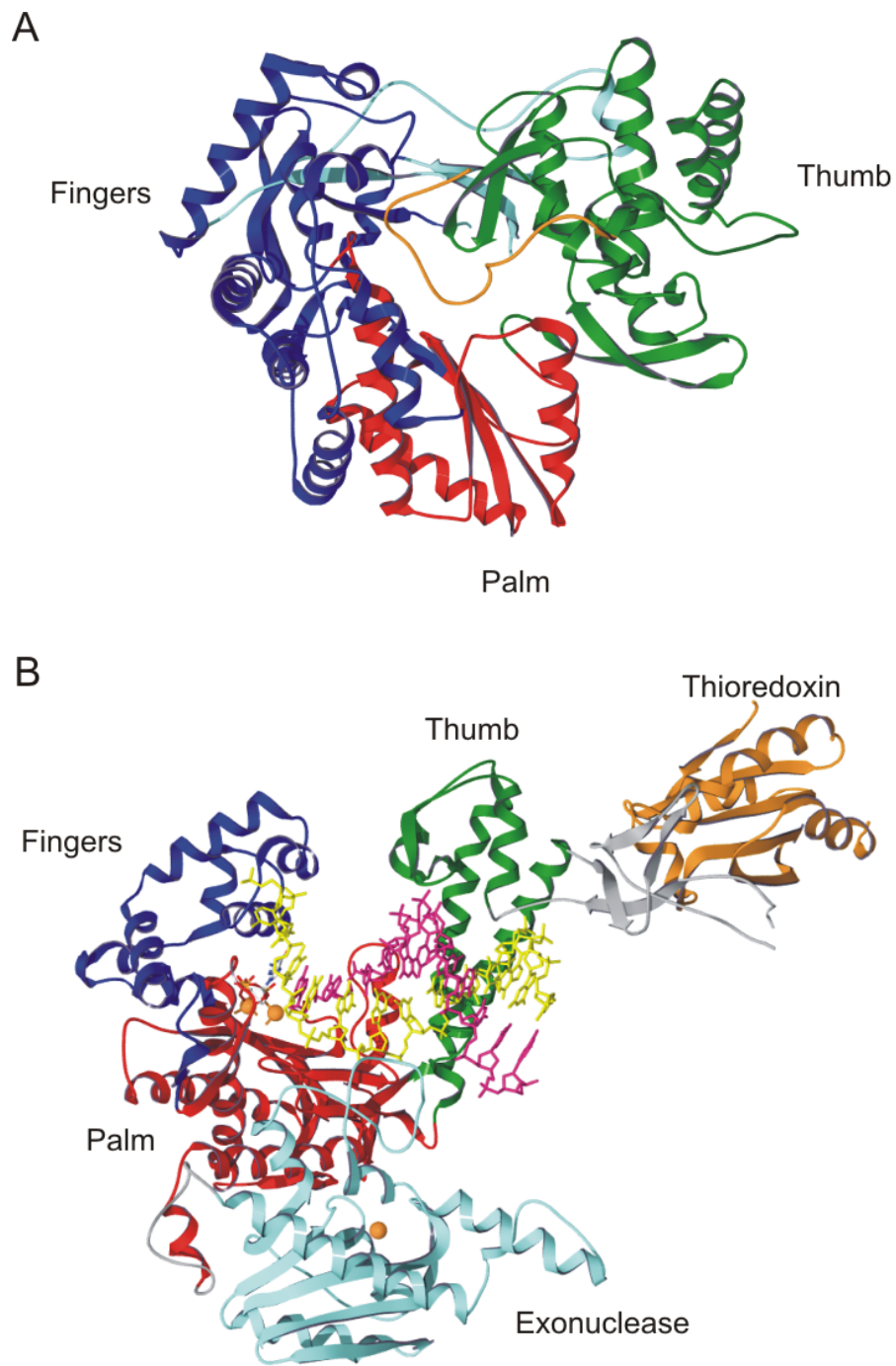
Figure 1.3: The structure of HCV RNA dependent RNA polymerase and T7 DNA polymerase

Figure 1.3:   The structure of HCV RNA dependent RNA polymerase and T7 DNA
              polymerase

(A) The structure of HCV NS5B RdRp shows extra amino acid loops interconnecting
fingers and thumb domains (Cyan). The hydrophobic tail consisting of 21 C-terminal
amino acids extends into the RNA binding cleft (Orange). (B) The T7 DNA polymerase
contains polymerase and 3'-5' exonuclease domains. A loop extending from tip of the
thumb domain binds to thioredoxin (Gray). DNA template and primer are shown in
yellow and magenta colors. A ddGTP substrate complex with two Mg^{2+} ions is bound to
the enzyme active site. The third Mg^{2+} ion binds to the exonuclease domain. Figures are
reproduced from 1C2P.pdb and 1T7P.pdb.

1.6   PROJECT SUMMARY

The research projects report in this dissertation can be divided into four sections. Chapter two presents our attempts to reconstitute an active HCV RdRp enzyme in order to develop *in vitro* assay systems for kinetic studies. Although an optimal reaction condition was established for *in vitro* assays, the recombinant enzymes including novel NS5B-NS3 fusion constructs showed low RdRp activities. The studies of T7 DNA polymerase mechanisms are presented in the rest of the dissertation. Chapter three reports kinetic and equilibrium studies of the polymerase conformational change detected using a fluorescently labeled enzyme. Our data suggested a novel mechanism of nucleotide selectivity involving the nucleotide-induced conformational change step. The kinetics of the polymerase conformational changes during continuous DNA polymerization is discussed in chapter four. The last chapter reports our results from experiments utilizing the properties of fluorescently labeled T7 DNA polymerase to detect single nucleotide mutations in DNA sequences.

Chapter 2: A Novel HCV RNA Polymerase-helicase Fusion Protein and Polynucleotide Phosphorylase Activity from HCV RNA Dependent RNA Polymerse

2.1   INTRODUCTION

In order to characterize the kinetic properties of NS5B, an *in vitro* assay system with purified, active enzyme is required. However, several studies showed that the rate of NS5B catalyzed RNA polymerization *in vitro* is in the order of 10^{-3}-10^{-4} s^{-1} (51;75) which is 10^5 fold slower than other well-characterized RNA and DNA polymerases. In this study, different approaches were taken to look for keys to improve the RdRp activity of recombinant NS5B. The wild-type and C-terminal truncated mutants of NS5B were expressed in *E. coli* and purified for biochemical studies. Activities of these recombinant NS5Bs were assayed using several different RNA templates derived from the UTRs of HCV genome under optimal Mg^{2+} concentration determined in titration experiments. To investigate the effect of NS3 on the NS5B RdRp activity, novel NS5B-NS3 fusion enzymes were created by fusing the scNS3-NS4A in-frame to the C-terminus of the 21 amino acid truncated form of NS5B, NS5BΔ21. Experimental data suggested that the novel fusion proteins have retained both the enzyme activities from NS3 and NS5B. Furthermore, a template independent, but MgCl$_2$ concentration dependent RNA polymerization activity was observed in the purified NS5B.

2.2   MATERIALS AND METHODS

2.2.1   Construction of NS5B-NS3 fusion enzymes

A plasmid, pCV-J4L6S (76), containing the genomic sequence of an infectious HCV-1b strain was used as a template for constructing NS5B-NS3 fusion proteins. Two different in-frame fusions were created by sequentially inserting three DNA fragments

15

into the *E. coli* expression vector, pET21a (Novagen). The first fragment of the fusion protein is a C-terminal 21 amino acid truncated NS5B amplified by PCR with primers, NS5B5'*Eco*RI (5'- ATA TGA ATT CAT GTC CTA TAC GTG GAC AGG C -3') and NS5B3'Δ21 (5'- TAT ATC TAG AAC CTC CGC GGC CGC GGG GTC GGG CAC GAG ACA G -3'). The second fragment is a linker between NS5B and scNS3-NS4A, which encodes repeating –Gly-Gly-Gly-Ser- sequences. Linkers with two different lengths were created by annealing complementary DNA oligomers to form double stranded DNAs. The 60Å linker was created by annealing DNA primer L60F (5'- GGC CGC GGA GGT GGA TCT GGA GGT GGG TCT GGC GGA TCT GGT T -3') to L60R (5'- CTA GAA CCA GAT CCG CCA GAC CCA CCT CCA GAT CCA CCT GGC C -3') while the 90Å linker was created with primers L90F (5'- GGC CGC GGA GGT GGA TCT GGT GGC GGA TCT GGT GGT GGC TCT GGA GGT GGG TCT GGC GGA TCT GGT T -3') and L90R (5'- CTA GAA CCA GAT CCG CCA GAC CCA CCT CCA GAG CCA CCA CCA GAT CCG CCA CCA GAT CCA CCT CCG C -3'). The complementary DNA pairs were mixed at a 1:1 molar ratio, heated to 95°C for 3 min, and cooled down slowly to room temperature to allow annealing of primer and template DNA. The helicase portion of the fusion protein is a single chain NS3-NS4A fusion protein based on a published design (67;68). DNA primers NS4A-3 (5'- ATA TTC TAG AGG AGG CAG TGT GGT CAT TGT GGG CAG GAT CAT CTT GTC CGG GTC GGG GTC GAT CAC GGC CTA CTC CCA ACA A -3') and NS3-3' (5'- ATA TCT CGA GAG TGA CGA CCT CCA GGT CAG C -3') were used in PCR to synthesize the encoding DNA fragments. Expression vectors for NS5B-NS3 fusion proteins were constructed using restriction enzyme sites, *Eco*RI, *Not*I, *Xba*I, and *Xho*I to subsequently clone the three DNA fragments into pET21a (Figure 2.1). Resulting expression plasmids were verified by sequencing analysis and the final fusion constructs were expressed as C-

terminal 6-histidine tagged proteins for affinity purification. Final fusion proteins were designated as NS5B60NS3-His and NS5B90NS3-His according to the length of the inserted linker.
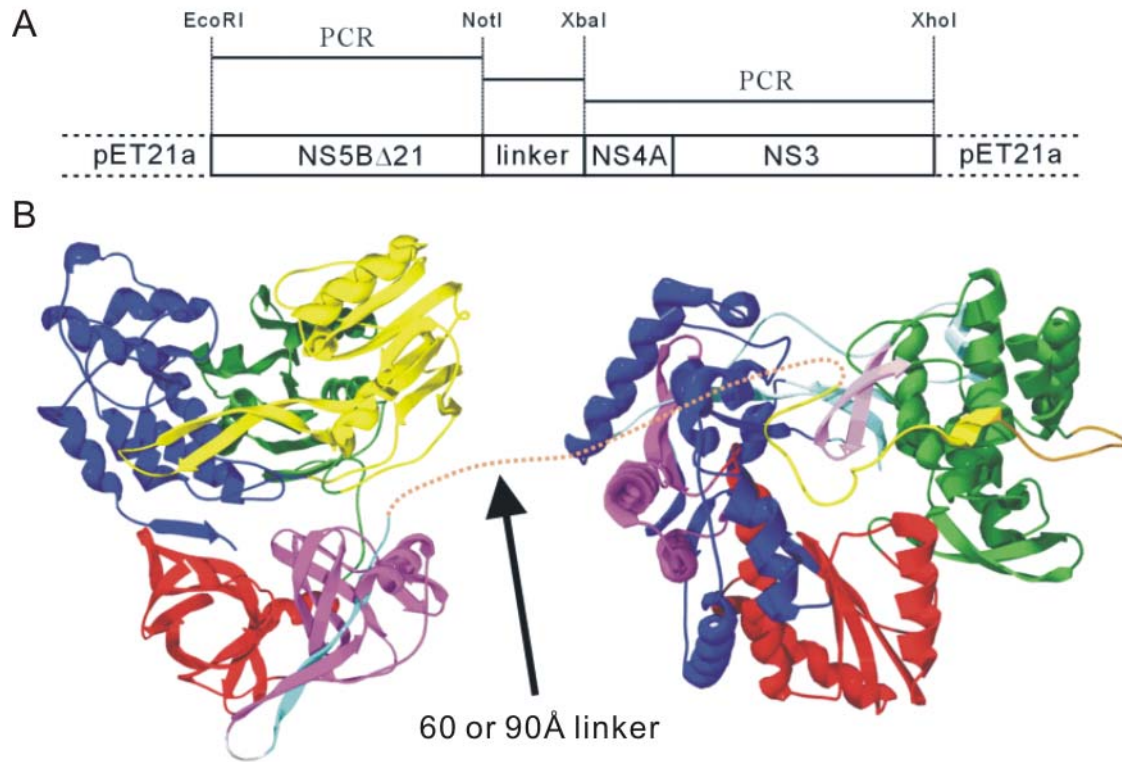
Figure 2.1: The NS5B-NS3 fusion proteins

(A) Restriction enzymes, *Eco*RI, *Not*I, *Xba*I, and *Xho*I, were used to clone DNA fragments encoding NS5BΔ21, linkers, and scNS3-NS4A into *E.coli* expression vector, pET21a. (B) 3D models for scNS3-NS4A and NS5B. The orange dashed line indicates where the linker peptides were added between the C-terminus of scNS3-NS4A and N-terminus of NS5B. The protein structure images are reproduced from pdb files, 1CU1.pdb and 1C2P.pdb.

17

2.2.2   Expression and purification of recombinant enzymes

The plasmid containing full-length NS5B with an N-terminal 6-histidine tag was kindly provided by Dr. Michael M. C. Lai (53). The coding sequence for NS55BΔ21 was amplified from the full-length construct with PCR primers, NS5BΔ21-5' (5'- CGA CGA TAG CTA GCG ACG ATG ACG ATA AGA TGT CCT ATA CGT GGA CA -3') and NS5BΔ21-3' (5'- ATG AGA GGG AAT TCT CAG CGG GGT CGG GCA CGA GA -3'). DNA products were digested with restriction enzymes, *Nhe*I and *Eco*RI, and inserted into the *E. coli* expression vector, pTrcHisB (Invitrogen). An active site mutated version of NS5BΔ21, NS5BΔ21(D317A), was also created by site directed mutagenesis using primers D317AF (5'- GCT CGT GAA CGG AGC AGA CCT TGT CGT TAT CTG TG -3') and D317AR (5'- CAC AGA TAA CGA CAA GGT CTG CTC CGT TCA CGA GC -3'). Plasmids for the expression of C-terminal 6-histidine tagged NS5BΔ21 and NS5BΔ55 were kindly provided by Dr. Joe Jaeger (77). All NS5B constructs including the polymerase-helicase fusion proteins were expressed and purified using the same procedure listed below. Plasmids containing different NS5B constructs were first transformed into the *E. coli* strain, JM109(DE3) (Promega). Plasmid containing colonies were selected on LB plates with 100μg/ml of ampicillin. 5 ml cultures of LB medium containing 100μg/ml of ampicillin were inoculated with single bacteria colony and incubated at 37°C for overnight with vigorous shaking. 2 ml of saturated overnight culture were then used to inoculate a 2-liter Terrific Broth medium and shaken at 37°C until OD_{600} = 0.5 was reached. Expression of NS5B proteins was then induced by adding 0.5mM of IPTG and incubating at 16°C for 16 hours. After induction, cells were harvested by centrifuging at 4000 x g for 10 minutes at 4°C and all solutions were kept at 4°C in subsequent steps. Cell pellets were weighed and resuspended with buffer A

18

(50mM Tris pH 7.5, 300mM NaCl, 10% Glycerol, 0.1% Triton X-100) with a volume of 10ml of buffer per gram of cells. Protease inhibitor cocktail (Sigma) was added according to the company's recommendation and cells were lysed on ice with a Branson Sonifier 450 sonicator at settings of 50% duty cycle and output power at 6 for a total of 2 minutes with 30 seconds on/off cycles. The cell lysates were then subjected to centrifugation at 30,000 rpm for 30 minutes to remove cell debris. $MgCl_2$ (10 mM) and Omnicleave (Epicentre) endonuclease (1U per gram cells) were added to the supernatant. The lysate was then incubated with constant stirring at 4°C for 1 hour to degrade bacterial DNA and RNA. After incubation, the solution was loaded onto a 1 mL Ni-NTA column by gravity and the column was washed with 10 mL of buffer A containing 20 mM imidazole. Bound proteins were eluted with 5 mL of buffer A with 200 mM of imidazole. Fractions containing NS5B were pooled and loaded immediately onto a 1 mL HiTrap Heparin HP column (Amersham Pharmacia Biotech) and washed with buffer A containing 400 mM NaCl. Proteins were eluted with buffer at 800 mM of NaCl. Eluted fractions containing NS5B were pooled and dialyzed two times against 2 liters of storage buffer (50mM Tris pH 7.5, 300 mM potassium acetate, 10% glycerol, 0.1% Triton X-100) for two hours each. DTT was added to the second dialysis buffer and purified NS5B aliquots were quickly frozen in liquid nitrogen and stored at -80°C.

2.2.3   Preparation of RNA templates

For the RdRp assays, five different RNA templates were derived from the 5'UTR and 3'UTR of the HCV-1b strain, pCV-J4L6S (Figure 2.2). The 100 nt X-region sequence includes the 98 nt X-region sequence with two extra guanosines added to the 5'end was obtained using PCR primers 6HCVX4G (5'- ACT GAC TAA TAC GAC TCA CTA TAG GGT GGG CTC CAT CTT AGC CCT AG -3') and 022101HCVXPCR (5'- ACA TGA TCT GCA GAG AGG CC -3'). The 5'UTR template is a 341 nt RNA

19

fragment derived from the 5'UTR sequence of the HCV genome obtained using PCR primers 5'UTR(Fwd) (5'- GCC AGC CCC CTG ATG GG -3') and T7P_341R(5'UTR) (5'- ATA TTA ATA CGA CTC ACT ATA GGG TGC ACG GTC TAC GAG ACC T - 3'). The 220 nt 3'UTR template is derived from the negative sense 3'UTR sequence of HCV genome using PCR primers T7P_9377(3'UTR) (5'- ATA TTA ATA CGA CTC ACT ATA GGG GAG CTA ACC ACT CCA GGC C -3') and 3'UTR(Rev) (5'- ACA TGA TCT GCA GAG AGG CCA GTA TC -3'). In addition, two of 560 nt 5'UTR-3'UTR combined templates were also created for in vitro assays. The 560 nt templates were created as follow. The 5'UTR fragment is the PCR product using primers 5'UTR(forward) (5'- TGA CTT CTA GAG CCA GCC CCC TGA TGG GGG CGA -3') and 5'UTR(reverse) (5'- CTC GAG CGG CCG CGA TCG ACT AGT GCA CGG TCT ACG AGA CCT CCC GGG -3'). The 3'UTR fragment is the PCR product using primers 3'UTR(forward) (5'- ATC GAG CGG CCG CAC GGG GAG CTA ACC ACT CCA GGC C -3') and X-region(reverse) (5'- TGA TGT CTA GAC ATG ATC TGC AGA GAG GCC AGT A -3'). The two DNA fragments were digested with restriction enzymes, *Xba*I and *Not*I, and cloned into a pT7/T3 18U (Ambion) plasmid. The final DNA template for the 5'UTR-3'UTR RNA transcript was obtained using PCR primers T3P_18U (5'- GCC AAG CTA ACC CTC ACT AAA GGG -3') and 3'UTR(Rev). The 3'UTR-5'UTR DNA template for RNA transcription was obtained using PCR with primers 5'UTR(Fwd) and T7P_18U (5'- CCA TGA TTA CGA ATT TAA TAC GAC TCA CTA -3'). Final RNA products were produced using MEGAshortscript T7 or MEGAscript T3 kit (Ambion). RNA products were treated with DNaseI to remove template DNA and purified on a 6% polyacrylamide/7 M urea gel electrophoresis.
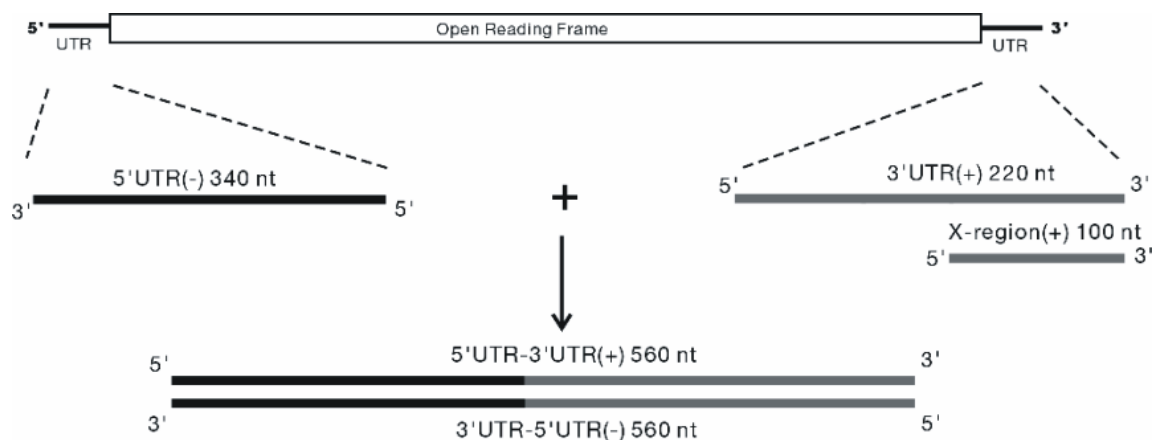
Figure 2.2:   RNA templates derived from 5'UTR and 3'UTR of the HCV genome

Five different RNA templates derived from the 5' and 3' UTR regions of the HCV genome were prepared for the RdRp assays. 5'UTR-3'UTR and 3'UTR-5'UTR templates contain combined sequences from both UTRs and are complementary to each other. Template sequences derived from the positive-sense genome are marked with (+). Templates derived from complementary sequence of HCV genome are marked with (-).

2.2.4 The RNA dependent RNA polymerase assay

The *in vitro* assays for *de novo* initiation and elongation were carried out in a buffer containing, 20 mM HEPES pH 7.3, RNaseOUT RNase inhibitor 0.4 U/µl (Invitrogen), 1 mM DTT, 200 mM potassium acetate, 25 mM $MgCl_2$, 300 nM NS5B enzymes, and 200 nM of RNA template. NS5B enzymes were preincubated with RNA templates in the reaction buffer for 30 minutes at 37°C. A final concentration of 1 mM of each nucleoside triphosphate (NTP) and 33 nM of α-^{32}P-ATP were added to the mixture to initiate the reaction. After incubating at 37°C for the desired length of time, the reaction samples were rapidly quenched with EDTA solution at a final concentration of 150 mM. Each reaction sample was then precipitated by adding 0.5 M ammonium acetate and an equal volume of isopropanol in the presence of 50 ng of yeast tRNA as a co-precipitant. After chilling at -80°C for 30 min and centrifuging for 15 minutes at 13,000 rpm (4°C), RNA pellets were dried by speed-vac (Savant), resuspended in denaturing gel loading buffer (Ambion), and analyzed by 6% polyacrylamide/7M urea gel electrophoresis. Results were visualized and quantified using a phosphorimager (Molecular Dynamics). Steady state rates of nucleotide incorporation were measured with 100 µl of reaction using conditions listed above. Reaction samples of 20 µl were quenched and analyzed at various time points after the reaction started. A serially-diluted α-^{32}P-ATP concentration standard was prepared and exposed to a phosphorimager screen along with the reaction samples. The concentration of nucleotide incorporated was calculated using the ImageQuant software (Molecular Dynamics).

2.2.5 The nucleoside triphosphatase activity assay

We examined the rates of hydrolysis of each nucleoside triphosphate. For each 20 µl of reaction, 5 nM of NS5B recombinant protein was preincubated at 37°C in the

reaction buffer (50 mM HEPES pH7.3, 10 mM KCl, 0.5 mM DTT 100μg/ml BSA, 1 mM MgCl$_2$) for 5 minutes. After the preincubation period, 1 mM final concentration of ribonucleoside triphosphate including 0.5 μCi of corresponding α-^{32}P-NTP was added to initiate the reaction. 4 μl of reaction sample was taken at certain time points and quenched with 4 μl of 0.5 M EDTA solution. The NTP substrate and NDP product were separated on polyethyleneimine-cellulose thin layer chromatography plates with 0.375 M of KH$_2$PO$_4$, pH 3.5 running buffer. The amount of hydrolyzed NTP was quantified using phosphorimager and ImageQuant software. The rate of ATP hydrolysis was also measured in the presence of 500 nM homopolymeric poly(C)$_{40}$ RNA.

2.2.6   Template independent incorporation of ribonucleoside triphosphate and dot blot analysis

In order to investigate the nature of high molecular weight RNA products observed in the NS5B RdRp reactions, single nucleotide incorporation experiments were performed using 300 nM of NS5B enzymes without any RNA templates. In the same reaction buffer as the RdRp assays, 1 mM of one of the four nucleoside triphosphates was added to initiate the reaction. After 90 min of incubation at 37°C, reactions were quenched with 150 mM final concentration of EDTA solution. Reaction products were treated and analyzed according to the procedures listed in section 2.2.4. Dot blot analysis was also performed on the reaction product from ATP incorporation experiments. The reaction products from the single ribonucleotide experiments were first precipitated and resuspended in formamide/formaldehyde denaturing solution. Samples were spotted onto nylon membrane (Micron Separations Inc.) wetted with 10 x SSC solution (1.5 M NaCl, 0.15 M sodium citrate pH 7.0). The membrane was air dried and RNA products cross-linked to the membrane with U.V. light. A 5'end ^{32}P labeled poly(dT)$_{25}$ probe was used

to detect the presence of poly(A) RNA products. Unlabeled poly(dT)$_{25}$ and poly(dA)$_{25}$ were spotted on the same membrane as negative and positive controls.

2.3   RESULTS

2.3.1   Expression and purification of recombinant enzymes

All NS5B constructs were expressed as 6 x His-tagged proteins and purified from *E. coli*. The His-NS5B and His-NS5BΔ21 have the tag on the N-terminus. Fusion enzymes, NS5B60NS3-His and NS5B90NS3-His, are C-terminally tagged. These enzymes were purified in the same manner with two liquid chromatography columns, Ni-NTA followed by heparin-sepharose. The buffer solution used in our purification has 300 mM of NaCl, 10% of glycerol, and 0.1% of Triton X-100 to keep NS5B from aggregating. It has been reported that C-terminal truncated versions of HCV NS5B with histidine tag on the C-terminus have higher RNA polymerization activities than the wild-type protein. Therefore, the C-terminal 6 x His-tagged version of truncated NS5Bs, NS5BΔ21-His and NS5BΔ55-His, were also prepared for comparison with other enzyme constructs in the biochemical assays. The active site mutated His-NS5BΔ21(D317A) was produced by site-directed mutagenesis and was used as a negative control for RdRp enzyme activity assays. Wild-type and truncated versions of NS5B migrate as ~66 kD bands and fusion proteins migrate as ~135 kD protein bands on SDS-PAGE gel stained with Coomassie Blue after purification (Figure 2.3).
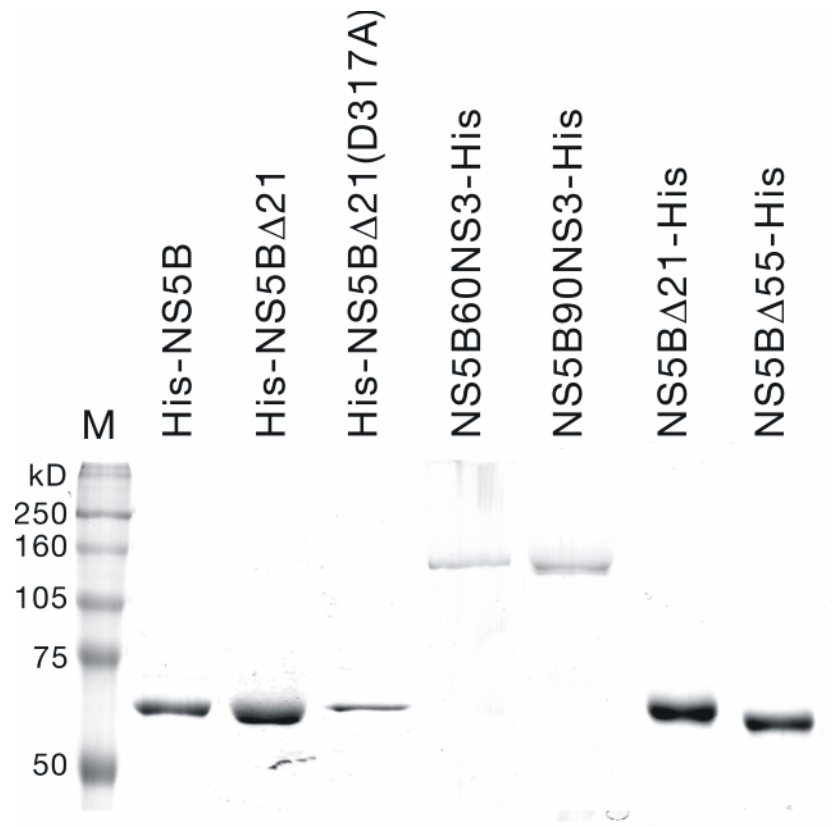
Figure 2.3: Purified NS5B enzymes

Five different constructs of NS5B RdRp enzymes were separated by 8% SDS-PAGE followed by Coomassie Blue staining. His-NS5B, His-NS5BΔ21, and His-NS5BΔ21(D317A) migrate as 66 kD bands on the gel. NS5B60NS3-His and NS5B90NS3-His have a molecular mass around 135 kD. All protein constructs were expressed in the *E.coli* and subsequently purified following the same purification protocol listed in experimental procedures section. Lane M: molecular weight markers.

2.3.2 NS5B RdRps are able to replicate various RNA templates

HCV NS5B was shown to utilize primer-dependent or primer-independent *de novo* initiation mechanisms to replicate different RNA templates (53-56). Five single stranded RNA templates derived from either the 5' or 3' untranslated regions of the HCV genome were studied in the NS5B-catalyzed initiation and elongation reactions. The RNA template sequences named X-region, 3'UTR, and 5'UTR-3'UTR are derived from the positive sense sequence of the HCV genome. In contrast, 5'UTR and 3'UTR-5'UTR templates are derived from the complementary sequence of the HCV genome (Figure 2.2). Although with different efficiencies, results from activity assays showed that the His-NS5B was able to replicate all the RNA templates supplied and produce full-length RNA products in the presence of 25 mM $MgCl_2$ in reaction buffer (Figure 2.4;2.5). For the NS5B60NS3-His fusion construct, very high RNA polymerization activities were detected in the reaction buffer containing 5 mM of $MgCl_2$. However, the reaction products appeared to be template nonspecific and resulted in smeared bands on autoradiograph. The high molecular weight product was greatly reduced when the concentration of $MgCl_2$ was increased to 25 mM. At higher $MgCl_2$ concentration, the NS5B60NS3-His was able to replicate the X-region and 5'UTR templates, but fully replicated RNA products were not observed with the other templates supplied in the same assay (Figure 2.5). Shorter products detected on the polyacrylamide gel indicated that NS5B60NS3-His fusion enzyme was able to utilize these templates as substrates and synthesize RNA products. Overall, the 341 nt 5'UTR RNA template was the most efficient in supporting NS5B replication in our assays. Both wild-type and fusion NS5Bs could replicate this template without the addition of primers and produced full-length RNA products. Other than the template-length RNA products, higher molecular weight

RNA products were also observed in RdRp assays. The amount of template-independent RNA product was more substantial in reactions containing only 5 mM of $MgCl_2$ and was greatly reduced in reactions with 25 mM of $MgCl_2$. This phenomenon was examined more closely in further experiments (see below). Deleting the C-terminal hydrophobic amino acids of NS5B was previously shown to not only improve enzyme solubility but also enzyme activity (50;51). Although the C-terminal deletion mutants expressed in *E. coli* were more soluble and produced higher yields during the purification steps, the full-length enzyme showed higher RdRp activities than the C-terminal truncated versions of the enzyme in our experiments (Figure 2.6).
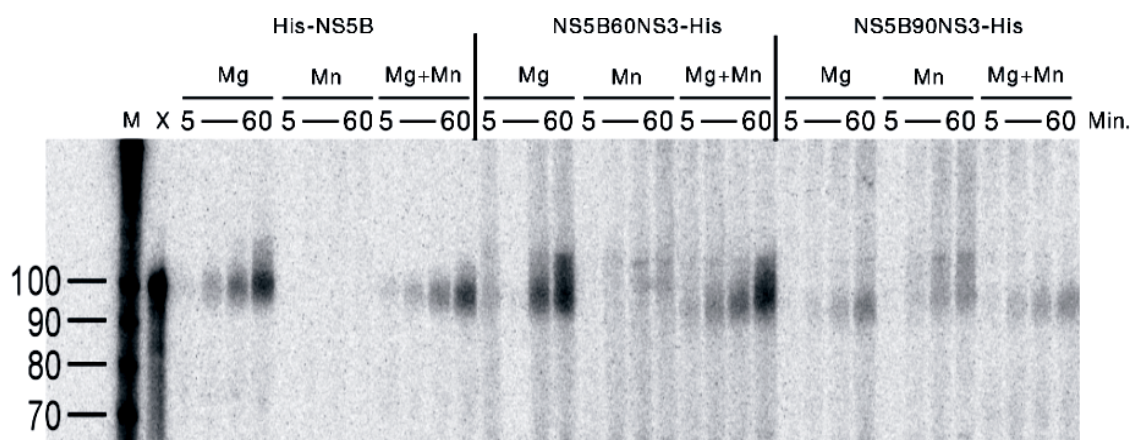
Figure 2.4:   RdRp activities of His-NS5B and NS5B-NS3 fusion enzymes with 100 nt X-region RNA template

NS5B polymerase (300 nM) preincubated with 200 nM of 100 nt X-region RNA template was mixed with a solution of 1 mM NTPs, α-^{32}P-ATP and either 50 mM $MgCl_2$, 7.5 mM $MnCl_2$, or both to start the reactions at room temperature. Reaction samples were taken at 5, 15, 30, and 60 minutes and the reaction quenched by rapidly mixing samples with EDTA solutions. RNA products were analyzed by denaturing polyacrylamide gel electrophoresis and visualized using a phosphorimager. Lanes: M: molecular weight marker. X: 5' end labeled X-region RNA.
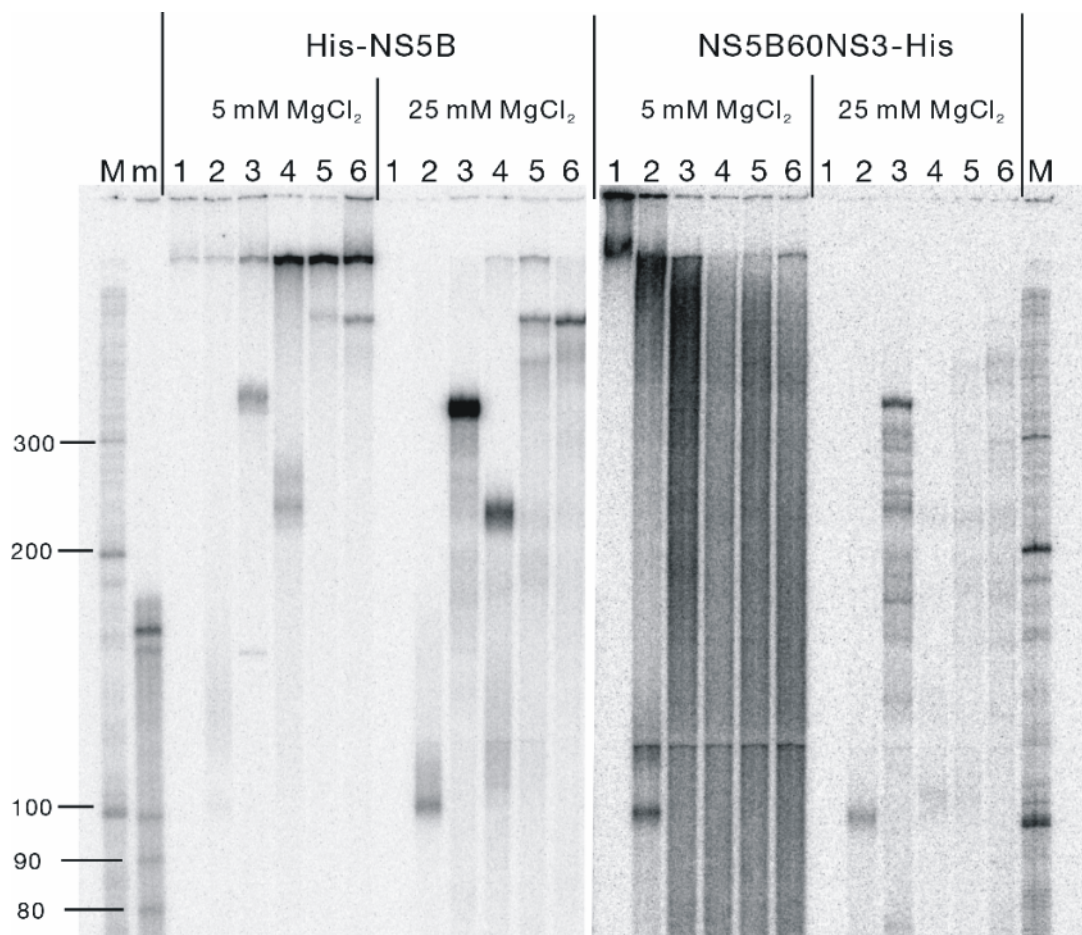
Figure 2.5: NS5B is able to replicate different RNA templates

NS5B enzyme (300 nM) preincubated with 200 nM of each different RNA template was mixed with 1 mM of NTPs, α-^{32}P-ATP, and 5 or 25 mM of MgCl$_2$ to start reactions. After 90 minutes of incubation at 37°C, reactions were quenched by EDTA solutions. Reaction products were separated on 6% denaturing polyacrylamide gels. Lanes: M, m: RNA molecular weight markers. 1: control reaction without RNA template added. 2: 100 nt X-region RNA template. 3: 340 nt 5'UTR RNA template. 4: 220 nt 3'UTR RNA template. 5: 560 nt 5'UTR-3'UTR RNA template. 6: 560 nt 3'UTR-5'UTR RNA template.

29
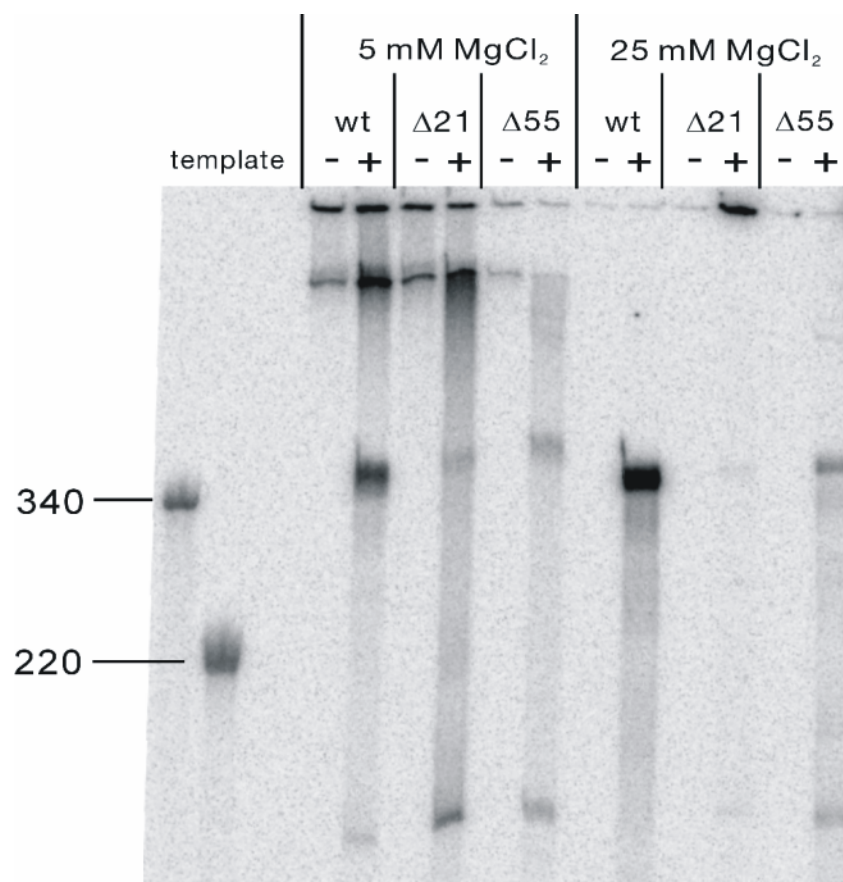
Figure 2.6:   RdRp activities of wild-type and C-terminal truncated forms of NS5B

His-NS5B, NS5BΔ21-His, and NS5BΔ55-His were assayed in reactions containing 300 nM of enzyme and 200 nM of 340 nt 5'UTR RNA template. MgCl$_2$ was added to 5 or 25 mM final concentration. After 120 minutes of incubation at 37°C, reactions were quenched with EDTA and analyzed by 6% polyacrylamide denaturing gel electrophoresis. Reactions without RNA template added were used as negative controls.

2.3.3 Mg^{2+} ion is sufficient to support NS5B RdRp activity

The 5'UTR template was used in experiments with His-NS5B and NS5B60NS3-His proteins to study the effects of different divalent cations on NS5B RdRp activity (Figure 2.7A). The activity of NS5B in the polymerization of template-length RNA product increased with increasing $MgCl_2$ concentration (5 to 25 mM). Although NS5B could replicate the 5'UTR template at 5 mM $MgCl_2$, products longer than the template were also observed. This phenomenon was more obvious in the NS5B60NS3-His catalyzed reactions and appeared as smeared bands on polyacrylamide gels. The optimal $MgCl_2$ concentration for HCV NS5B catalyzed reaction determined in the titration experiments was 25 mM. In the $MnCl_2$ titration experiments with 5 mM $MgCl_2$ present in the solution, His-NS5B incorporated more nucleotides into the RNA products with increasing amounts of $MnCl_2$ added. However, the RNA product of template-length became more undefined and the formation of short (~150 nt) RNA product increased (Figure 2.7B). The same experiment carried out at 25 mM $MgCl_2$ in solution showed the formation of short RNA product when $MnCl_2$ was added, but the amount of major template-length RNA product remained unchanged. In the case of NS5B60NS3-His fusion enzyme, optimal $MgCl_2$ concentration was also 25 mM. High molecular weight RNA product formed at 5 mM of $MgCl_2$ disappeared when 5 mM of $MnCl_2$ was added to the reaction buffer. Adding $MnCl_2$ to NS5B60NS3-His assay at optimum $MgCl_2$ concentration only slightly increased the formation of 341 nt RNA product.

In conclusion, the titration experiments have shown that $MgCl_2$ alone is sufficient to support NS5B RdRp activities. Although adding $MnCl_2$ to the 5 mM $MgCl_2$ reactions could stimulate NS5B activity, it had no significant effect at 25 mM $MgCl_2$ concentration.
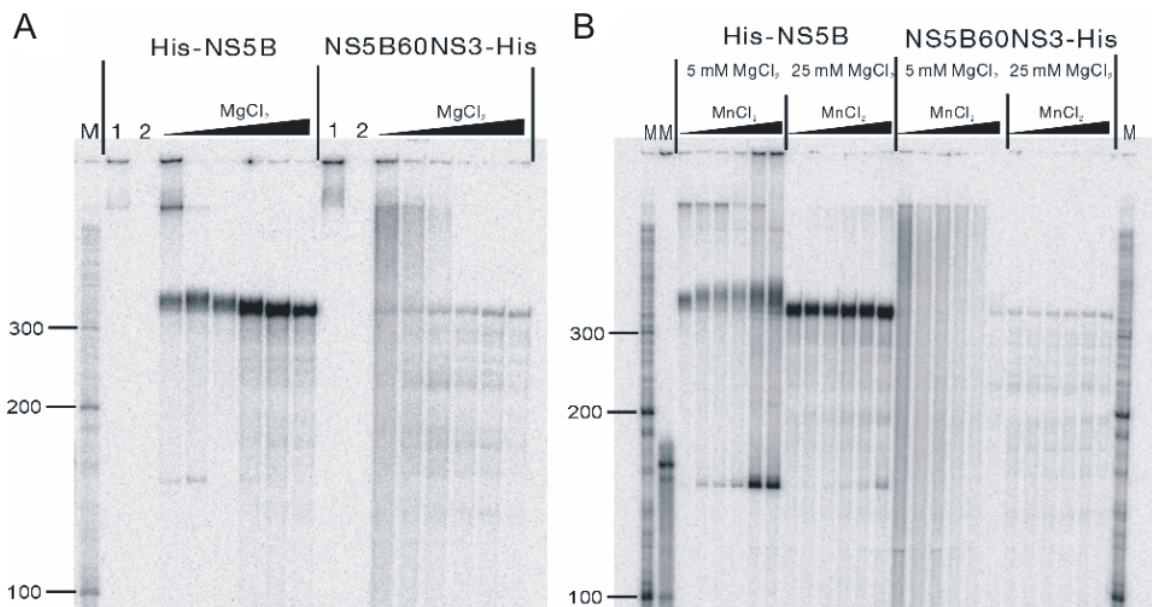
Figure 2.7: Effect of Mg^{2+} and Mn^{2+} ions on NS5B RdRp activity

Titration of NS5B RdRp activity with 5'UTR RNA template was performed in different Mg^{2+} and Mn^{2+} concentrations. Each reaction included 300 nM of enzyme, 200 nM of template, 1 mM of each NTP and α-^{32}P-ATP. Reactions were stopped after 90 minutes of incubation at 37°C. (A) Titration experiments of NS5B RdRp activity with 5, 10, 15, 20, 25, and 30 mM of MgCl$_2$ concentrations. Lane 1: 5 mM MgCl$_2$ control experiment with no RNA template added. 2: 30 mM MgCl$_2$ control experiment with no RNA template added. M: RNA molecular weight markers. (B) Titration experiments of NS5B RdRp activity with 0, 1, 2, 3, 4, and 5 mM of MnCl$_2$ concentrations in the presence of either 5 or 25 mM MgCl$_2$. Lanes M: RNA molecular weight markers.

2.3.4 NS5B-NS3 fusion proteins have NTPase activities

The HCV NS3 protein is a multi-functional enzyme with both helicase and protease activities. Its protease function is required for viral polyprotein processing while the helicase function is believed to assist viral replication. Although the composition of the HCV replication machinery is not defined, the NS3 protein is believed to be a part of the replisome. Experiments described in the previous section have shown that the novel NS5B-NS3 fusions retained the RdRp activity. In order to determine whether the NS3 enzyme was active in the context of NS5B-NS3 fusion, we performed nucleoside triphosphatase activity assays with the fusion proteins. The His-NS5B was used as a negative control in the assay and only showed background levels of ATP hydrolysis after a 60-minute incubation period (Figure 2.8). In contrast, the NS5B-NS3 fusion proteins showed high ATPase activity. The NS5B60NS3-His could hydrolyze ATP at 30 s^{-1}, while NS5B90NS3-His had a rate of 9.0 s^{-1}. The rates for hydrolyzing other ribonucleoside triphosphates were also measured (Table 2.1). Although the fusion enzymes could also hydrolyze GTP, CTP, and UTP, the rate of hydrolyzation was about 10-fold lower when compared with the rate of ATP hydrolysis. The rate of ATP hydrolysis reaction that was catalyzed by the fusion enzymes was also measured in the presence of a poly(C)$_{40}$ homopolymeric RNA. The results showed that adding RNA to the reaction significantly enhanced the ATPase activity of both fusion enzymes. This RNA-stimulated ATPase activity was previously observed with purified NS3 helicase (66). The observation implies that the NS3 helicase of the fusion protein is active in terms of nucleoside triphosphatase activity, but this activity does not seem to be coupled with the NS5B RdRp function in our assay because the addition of poly(C)$_{40}$ RNA enhanced the ATPase activities in the absence of active RNA polymerization.

33
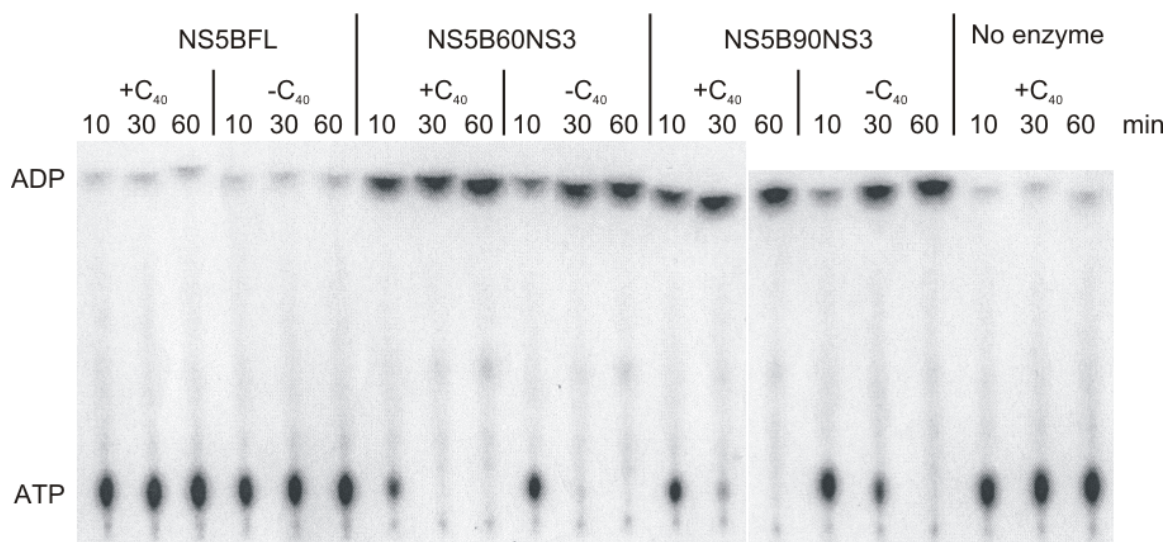
Figure 2.8:   The ATP hydrolysis reaction catalyzed by recombinant NS5B enzymes

The radioautograph shows the separation of ATP from its hydrolysis product, ADP. The full length His-NS5B RdRp shows only background levels of ATPase activity when compared to the control experiments performed without adding purified enzymes. Both NS5B-NS3 fusion constructs show RNA (C_{40}) stimulated ATPase activity.

	NS5B60NS3-His	NS5B90NS3-His
ATP	30 ± 0.7 s^{-1} (40 ± 1.8 s^{-1})	9 ± 0.8 s^{-1} (18 ± 2.1 s^{-1})
GTP	2.3 ± 0.08 s^{-1}	0.99 ± 0.101 s^{-1}
CTP	3.1 ± 0.29 s^{-1}	0.93 ± 0.081 s^{-1}
UTP	2.7 ± 0.06 s^{-1}	1.0 ± 0.19 s^{-1}

( ): Rates measured in the presence of poly(C)$_{40}$ homopolymeric RNA.

Table 2.1:   The nucleotide triphosphotase activity of NS5B-NS3 fusion enzymes

2.3.5 NS5B has a template-independent RNA polymerization activity

In the RdRp assay, unusually high molecular weight reaction products were observed in reactions with 5 mM of $MgCl_2$. The amount of these products was greatly reduced when 25 mM of $MgCl_2$ was added to the reaction buffer. The nature of these very long RNA products was initially thought to be the result from multiple rounds of RNA replication carried out by NS5B using the copy-back mechanism or due to the intrinsic TNTase activity to elongate the RNA template. Control experiments performed in the absence of RNA templates showed the formation of similar RNA products and ruled out those possibilities (Figure 2.6; 2.7). To examine the nature of the template-independent RNA product observed in our assays, RdRp reactions were carried out in the absence of RNA templates and in the presence of a single ribonucleoside triphosphate with the corresponding α-P^{32} labeled nucleotide as substrate. Polyacrylamide gel analysis showed the formation of long RNA products in the reactions where ATP or CTP was added as sole substrate in the presence of 5 mM of $MgCl_2$ (Figure 2.9A). Product in the CTP incorporation experiments disappeared when 25 mM of $MgCl_2$ was added to the reaction buffer while the amount of product was reduced at the higher $MgCl_2$ concentration in the presence of ATP. These high molecular weight products are of similar size to the template nonspecific products formed in our normal RdRp experiments in the presence of all four nucleoside triphosphates with or without the RNA template as substrates. In order to further characterize these long RNA products, dot blot analysis was performed using a radioactively labeled 25mer $poly(dT)_{25}$ probe to detect the formation of poly(A) RNA in the reaction products produced by the fusion enzyme, NS5B60NS3-His (Figure 2.9B). The results showed that the $poly(dT)_{25}$ probe hybridized to reaction products not only from the ATP incorporation experiment but also to products from the

normal RdRp assay. Because RNA templates are not required for the polymerization of long RNA products in the RdRp assays, NS5B is likely to have activity similar to polynucleotide phosphorylase, which catalyzes the polymerization of RNA without the requirement of a RNA template (78;79).
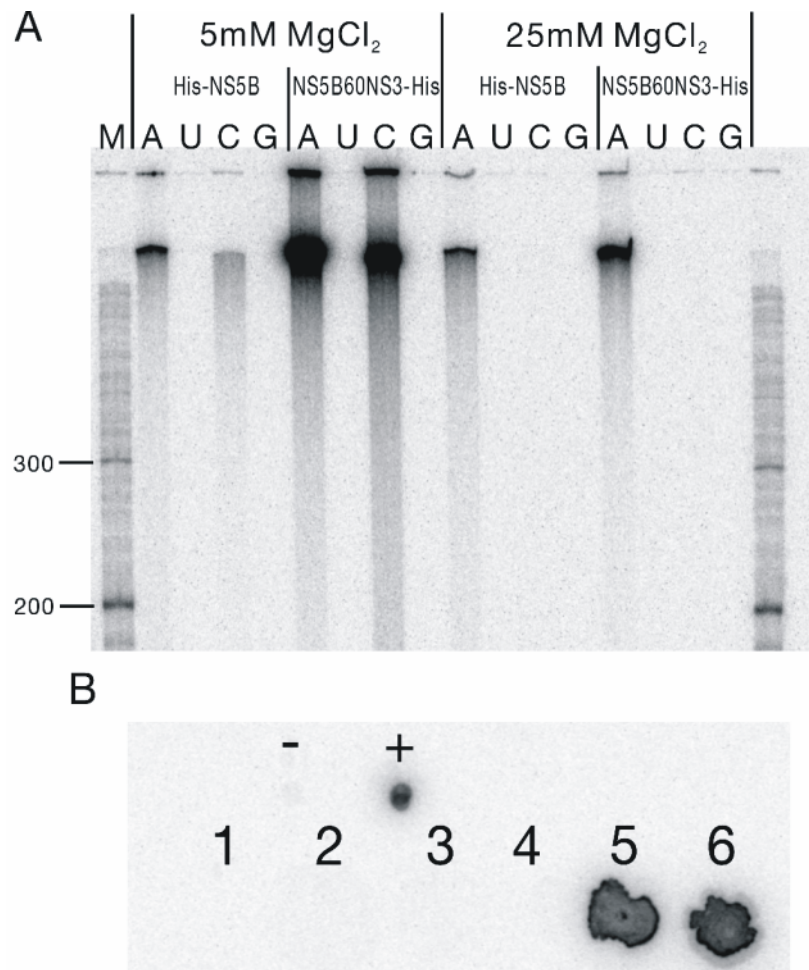
Figure 2.9: The polynucleotide phosphorylase activity of NS5B

Figure 2.9: The polynucleotide phosphorylase activity of NS5B

(A) Separate RdRp assays were performed on His-NS5B and NS5B60NS3-His with each of the four nucleoside triphosphates and the corresponding α-^{32}P labeled nucleotide as substrate. Enzymes (300 nM) were incubated with 1 mM NTP for 90 minutes at 37°C in buffers containing either 5 or 25 mM of $MgCl_2$. No RNA templates were added to these reactions. (B) Dot blot analysis of RdRp reaction products using a 5' end ^{32}P labeled poly(dT)$_{25}$ as hybridization probe. RdRp reactions were performed with fusion enzyme, NS5B60NS3-His, under the following conditions. 1: enzyme alone. 2: enzyme + 3'UTR RNA. 3: enzyme + 3'UTR RNA + NTPs. 4: enzyme + NTPs. 5: enzyme + 3'UTR RNA + ATP. 6: enzyme + ATP. Poly(dA)$_{25}$ and poly(dT)$_{25}$ were used as positive and negative controls for the dot blot analysis.

2.3.6 The high molecular weight RNA products are not associated with NS5B

Another possible explanation for the observation of high molecular weight RNA products on the polyacrylamide gel is that NS5B was covalently attached to the radioactively labeled nucleotide or the newly synthesized RNA. The resulting complexes would have a very high apparent molecular weight and migrate slowly in the sequencing gel relative to the free RNA products. In order to determine whether the radioactivity was associated with NS5B after RdRp reactions, SDS-PAGE was used to separate the reaction products. If the radioactive RNA products were covalently associated with NS5B after the RdRp reaction, the radioactively labeled material would co-migrate with the NS5B enzyme on the polyacrylamide gel. Images from Coomassie Blue staining and the radioautograph showed that NS5B does not migrate with the radioactive products on the gel (Figure 2.10). The result indicates that the observed products are most likely high molecular weight RNA molecules.
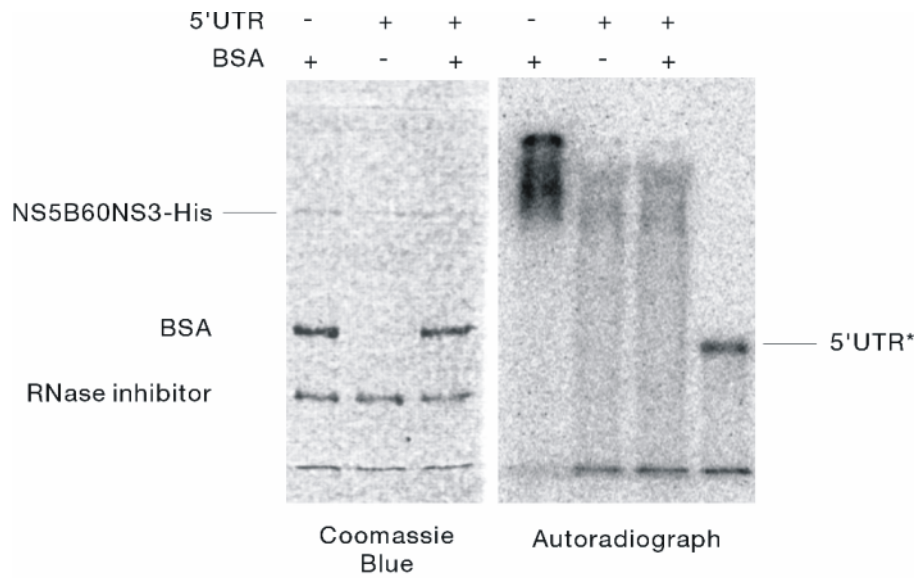
Figure 2.10: The high molecular weight RNA products from NS5B assays are not covalently associated with NS5B

After NS5B RdRp reactions were quenched with EDTA, products were separated by denaturing SDS-PAGE. Coomassie Blue staining shows the protein components from the reactions. The autoradiograph shows the radioactive RNA products from the reactions. The lane on the far right in the autoradiograph is a control reaction using 5' end ^{32}P labeled 5'UTR and unlabeled NTPs as substrates.

2.3.7 NS5B has low RNA polymerization activity

In order to compare RNA polymerization activities between different recombinant enzymes used in this study. Steady state rates of RNA polymerization were measured. Two RNA templates, a homopolymeric poly(C)$_{40}$ RNA (Figure 2.11) and a heteropolymeric 100 nt X-region RNA, were used in the RdRp assays. The rates of nucleotide incorporation for His-NS5B, His-NS5BΔ21, and two NS5B-NS3 fusions were on the order of 10^{-4} s^{-1} (Table 2.2). It would require a year for NS5B to fully replicate a 9500 nt HCV genome at this rate. According to this kinetic measurement, the addition of the NS3 helicase to the NS5B-NS3 fusion proteins did not show any enhancement in terms of the steady state rate of RNA polymerization over both the wild-type and the truncated form of NS5B. However, it is interesting to see that NS5B60NS3-His had higher NTPase activity and a slightly higher rate of RNA polymerization than the NS5B90NS3-His fusion enzyme. In order to answer the question whether higher NTPase activity has a positive effect on RdRp activity in the context of NS5B-NS3 fusion enzymes, more detailed studies are needed.

Figure 2.11: Steady state incorporation of GTP using C_{40} RNA as template

Various NS5B enzymes (400 nM) were incubated with 125 nM of C_{40} homopolymeric RNA and 20 μM of GTP with 20 nM of α-^{32}P-GTP at room temperature for up to 60 min. RNA products resulted from the polymerization reaction were spotted onto a Whatman DE81 paper and exposed to phosphoimager after washing with 5% of NaH_2PO_4 pH 7 solutions. The amount of product formed is quantified with a GTP standard and the steady state rate of product formation is calculated by dividing the rate of product formation by enzyme concentration.

	Templates	
	Poly(C)$_{40}$	X-region
His-NS5B	0.0004 s^{-1}	0.0001 s^{-1}
His-NS5BΔ21	0.0003 s^{-1}	0.0005 s^{-1}
NS5B60NS3-His	0.0002 s^{-1}	0.0003 s^{-1}
NS5B90NS3-His	0.0001 s^{-1}	0.0002 s^{-1}

Table 2.2:   Steady state rates of NS5B catalyzed RNA polymerization

2.4   DISCUSSION

In this study, different constructs of NS5B RdRp were expressed and purified from *E. coli* for kinetic studies. The N-terminal his-tagged NS5B showed more than a four-fold increase in the ability to replicate the 5'UTR RNA template when $MgCl_2$ concentration was increased from 5 to 25 mM. The increase of RdRp activity was also observed in experiments adding 0 to 5 mM of $MnCl_2$ in the presence of 5 mM $MgCl_2$, although the enhancement of NS5B activity was not as dramatic as increasing the $MgCl_2$ concentration to 25 mM. When 25 mM of $MgCl_2$ was included in the reaction buffer, the addition of $MnCl_2$ had a minimal effect on NS5B activity (Figure 2.7). The $MnCl_2$ titration experiment also showed the accumulation of RNA product, which has shorter length than the input template substrates. The formation of short RNA product may suggest early termination of the polymerization reaction. These results are controversial when compared to the previous studies that showed Mn^{2+} ions supported better NS5B enzyme activity than Mg^{2+} ions (50;54;56). In the case of NS5B60NS3-His enzyme, increasing the $MgCl_2$ concentration also improved the enzyme's ability to replicate the 5'UTR RNA. However, when the reaction buffer contained only 5 mM of $MgCl_2$, most of the RNA product was not template-length. By increasing the $MgCl_2$ concentration to more than 20 mM, most of the template-unrelated product disappeared and the amount of template-length products increased. Adding $MnCl_2$ in the presence of 5 mM $MgCl_2$ had similar result, but the effect was not as significant as adding 25 mM $MgCl_2$.

RNA products corresponding to the length of input templates indicated that NS5B enzymes were capable of *de novo* initiation of RNA replication under the assay conditions provided in this study. Experimental data also showed the formation of template irrelevant products. These products were greatly reduced under optimum $MgCl_2$

concentration. We focused on the template-length related products for comparison between enzyme activities and concluded that the wild-type NS5B showed highest RdRP activity in terms of replicating the 5'UTR RNA (Figure 2.6; 2,7). Based on the crystal structure of NS5B, which shows the C-terminal region of NS5B extending into the putative RNA binding groove between the fingers and thumb domains (58), and data suggesting that the truncated forms of NS5B have higher RdRp activity, it has been argued that the C-terminal amino acids of the protein may block the entry of RNA template and nucleotide substrates. Although the structural data and previous observations supported the hypothesis, our observations suggest that the interaction between the C-terminal amino acid residues and those in the substrate binding groove may be destabilized for the entry of substrates by increasing the reaction temperature to 37°C for optimal RdRp activities. In my experiments, deleting the C-terminal region of the enzyme in fact has negative impact on its RdRp activity. Although the full length His-NS5B enzyme showed the highest activity in the RdRp activity assay, it is not possible to conclude that NS5B-NS3 fusion enzyme has lower RdRp activities than NS5B enzyme alone in this study. One of the reasons is that our NS5B-NS3 fusion enzymes expressed high levels of NTPase activity. The nucleotides in the reaction solution were hydrolyzed at very fast rates and decreased the available nucleoside triphosphate substrates available for NS5B-catalyzed polymerization reactions. However, the fully replicated 5'UTR RNA indicated that NS5B60NS3-His has fairly high RdRp activity even under the conditions where available substrates were decreasing at high rates.

It is believed that both UTRs of the HCV genome contain *cis*-acting elements for the initiation of RNA replication. NS5B exhibited different efficiencies for replicating the five RNA templates derived from the 5' and 3' UTRs in our experiments. By comparing the amounts of radioactivity incorporated into the RNA products, the 5'UTR served as

45

the most favorable template for NS5B-catalyzed reaction. However, the efficiency of replication initiation from the different templates could not be compared directly. Because $\alpha\text{-}^{32}$P-ATP was the radioactively labeled nucleotide used in the assays, and the amount of radioactivity incorporated into the RNA products would depend on the number of molecules synthesized and the number of adenosines incorporated per RNA molecule. Taking this factor into consideration, the 5'UTR replication products would have 63 adenosines compared to the 100 adenosines incorporated into the 3'UTR replication product. A conclusion that NS5B initiates RNA polymerization from the RNA template with 5'UTR sequence more efficiently than that with the 3'UTR sequence could be made because more radioactivity was incorporated in reactions using the 5'UTR template (Figure 2.5). This result also supported the observation that NS5B initiates RNA polymerization from the complementary strand of genome more efficiently than from the positive-sense genome (80).

In addition to the RdRp replication products observed in the NS5B assays, longer RNA products were also detected. The formation of these long RNA products was template independent and MgCl$_2$ concentration dependent. Although it has been reported that purified NS5B could have host TNTase contamination (49), this NS5B associated polymerization activity observed in our assays did not require the RNA molecules as substrates. ATP or CTP was the only required substrate for the synthesis of these long RNA molecules (Figure 2.9A). Other than TNTase, *E. coli* PNPase is an enzyme able to catalyze the polymerization of RNA molecules without any RNA template or primer as precursors, but the substrates for PNPase are nucleoside diphosphates (81). Experimental results from this study showed that purified NS5B catalyzed the polymerization of RNA from only ATP or CTP in a RNA-independent manner, with ATP as the preferred substrate. This observation also ruled out the possibility of PNPase contamination in our

46

NS5B enzyme preparation. An interesting observation of NS5B-associated polynucleotide polymerase activity is the dependence of $MgCl_2$ concentration. At 25 mM of $MgCl_2$, NS5B showed optimal RdRp activity and reduced polynucleotide polymerase activity. Whether this activity of NS5B is physiologically important or just an artifact observed *in vitro* is unknown and requires further investigation. However, it is possible that this activity is related to the *de novo* initiation of RNA polymerization.

Chapter 3: Kinetics and Equilibrium of the Nucleotide-induced T7 DNA Polymerase Conformational Change

3.1   INTRODUCTION

Proteins labeled with an environmentally sensitive fluorophore have been used for direct observation of protein conformational changes (82-84). One successful application is monitoring the rapid changes of inorganic phosphate (Pi) concentration in real time using a fluorescently-labeled phosphate binding protein (PBP) from *E. coli* (85). The A197C mutant of PBP was labeled with a coumarin derivative, 7-diethylamino-3-((((2-maleimidyl)ethyl) amino)carbonyl)coumarin (MDCC) (Figure 3.1C) and used as a Pi sensor. This MDCC-PBP conjugate has been used in the study of biological processes involving ATP and GTP hydrolysis. The crystal structure of MDCC-PBP shows that the fluorophore attached at C197 is on the edge of the Pi binding site, close enough to monitor the conformational changes associated with Pi binding without blocking the binding site itself (86). Recently, a high resolution crystal structure of T7 DNA polymerase complex with both primer-template and ddGTP in the polymerase active site shows the fingers domain rotating towards the primer-template to bring the four conserved residues on the O-helix into contact with the incoming nucleotide (Figure 3.1A, 3.1B) (15;25). It is suggested that the movement of this recognition domain is the most critical step to bring all the key residues in the active site into proper positions for phosphoryl transfer to occur. Relating to the induced-fit model for nucleotide selectivity, this movement is thought to represent the rate-limiting conformational change of the polymerase before chemistry. Previous kinetic studies of T7 DNA polymerase indicated that nucleotide binding induces a change in enzyme conformation that is rate-limiting for incorporation (9;22).

48
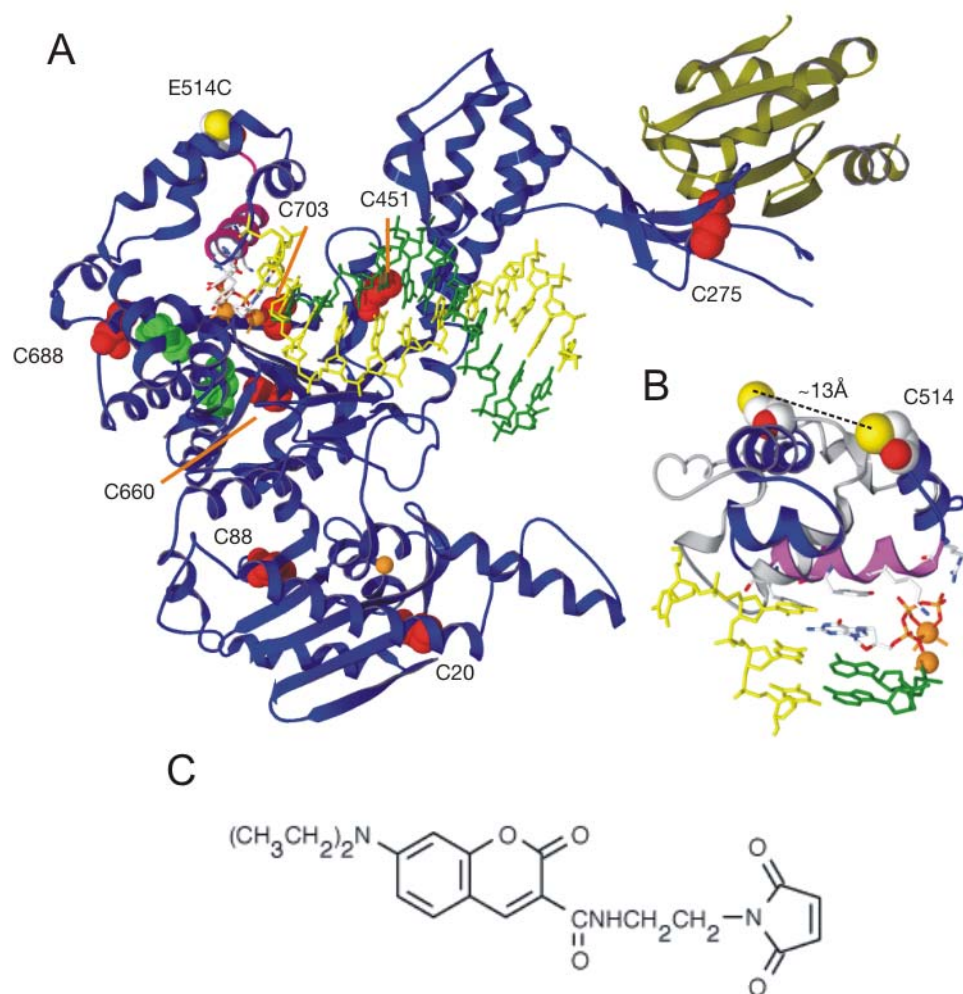
Figure 3.1: The structures of T7 DNA polymerase and MDCC

Figure 3.1:   The structures of T7 DNA polymerase and MDCC

A cys-lite version of T7 DNA polymerase with an E514C mutation was used for site-specific fluorescent labeling. (A) T7 DNA polymerase structure adopted from 1T7P.pdb with the mutated cysteine residues labeled in red. Residue C313 is in a disordered region and not visible in the structure. (B) Overlay of the recognition domain from a ligand-free T7 DNA polymerase structure (gray, unpublished data from Tom Ellenberger) with 1T7P.pdb (color). Amino acid residue 514 shows a 13Å movement when the palm domains from both structures are aligned. (C) The thio-reactive fluorophore used for fluorescent labeling, 7-diethylamino-3-((((2-maleimidyl)ethyl)amino)carbonyl)coumarin (MDCC).

In order to directly study the kinetics of the recognition domain reorganization, we created a novel T7 DNA polymerase construct with an E514C mutation for site-specific labeling with MDCC. The fluorescence from the MDCC label showed significant signal change upon nucleotide binding, and correct nucleotide binding was distinguishable from mismatched nucleotide binding. The kinetic analysis revealed that the conformational change step is at least partially rate-limiting during the incorporation of correct nucleotide. Also, the reaction pathway that leads to misincorporation is not only thermodynamically unfavorable but also proceeds through a reaction pathway that is different from correct incorporation. Based on the kinetic data, a new model is proposed to explain how polymerases utilize the small free energy differences in the ground-state binding step to achieve high fidelity DNA polymerization.

3.2   MATERIALS AND METHODS

3.2.1   Mutagenesis of the T7 DNA polymerase

The plasmid encoding an exonuclease-deficient (exo⁻) mutant of T7 DNA polymerase, pG5X (87), was used for constructing a cys-lite enzyme. Eight of the ten cysteines in the native enzyme were mutated using PCR with synthetic DNA primers. PCR segments separated by each mutated residue were synthesized using DNA primer pairs: T7P-43F (GCT CGC CCG GGG ATC CGG) and C20S_R (GAT AAC CCC GCT GTG GAA CTT AGT GAC GCT CTC), C20S_F (TAA GTT CCA CAG CGG GGT TAT CTA CGA CTA CTC CAC C) and C88A_R (GGG TGT CAA TAG CGT TCT CAC GAG GAA GGT GG), C88A_F (CGT GAG AAC GCT ATT GAC ACC CTT GTG TTG TCA CG) and C275A_R (TCG CGG ATG GGC GAA CAT CTC AGT GCC ACC TTT AGG), C275A_F (CTG AGA TGT TCC CCC ATC CGC GAA CAG GTA AGC) and C313A_R (GGT ATC AAG TTC GGC AGG CTC ACG GCC TTC TCG C),

51

C313A_F (CGT GAG CCT GCC GAA CTT GAT ACC CGC GAG TAC GTT GC) and C451S_R (GCA GCG CGA CTC TGC TCT CCA TAA GGA GAA CGT AC), C451S_F (GGA GAG CAG AGT CGC GCT GCT TTT GGC GC) and C660A_R (GGT ACG GGC GCC TAC TTG GAT TTC ATC ATG TAC CC), C660A_F (CAA GTA GGC GCC CGT ACC GAA GAG ATT GCT C) and C688A_R (ATC CAG AAG AGC CCG GAA GTT CCA GTG GTC TCC), C688A_F (GAA CTT CCG GGC TCT TCT GGA TAC CGA AGG TAA GAT) and C703A_R (ATC AGT GGG CAA TCG CCC AAT TAG GAC CCA TC), C703A_F (TTG GGC GAT TGC CCA CTG ATA CAG GAG GCT ACT C) and 2435*Hind*III_R (TAG AGC CAA GCT TGC ATG CC). The complete coding sequence for the cys-lite version of exo⁻ T7 DNA polymerase was created by sequentially joining the DNA segment and amplified by PCR with T7P-43F and 2435*Hind*III_R primers. The E514C mutation was introduced using primers, E514C_R (GGT AGG TAG GCA AGC AGC TAT CTG GTT CTT AGT GTG G) and E514C_F (CCA GAT AGC TGC TTG CCT ACC TAC CCG AGA TAA CGC TAA G). The amplified DNA product was cloned back into *Bam*HI-*Hind*III-digested pG5X plasmid and confirmed with sequencing analysis. The final plasmid construct expresses an exo⁻ T7 DNA polymerase with eight native cysteines removed and one cysteine added at residue number 514 (E514C-8C mutant of T7 DNA polymerase).

3.2.2 Expression, purification, and MDCC labeling of the T7 DNA polymerase

The plasmid encoding the E514C-8C mutant of the T7 DNA polymerase was 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 with 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 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).

Before the MDCC labeling step, the purified T7 DNA polymerase was dialyzed against the labeling reaction buffer (40 mM HEPES, pH 8, 0.1 mM EDTA, 10% glycerol, 50 mM NaCl, and 1 mM Tris(2-carboxyethyl)phosphine hydrochloride). 10 mM stock solution of MDCC (Molecular Probes) was prepared by dissolving solid MDCC with DMSO and stored at -80°C prior to use. Protein labeling was carried out in the labeling reaction buffer. 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 DTT solution. Labeled mutant protein (MDCC-E514C-8C) was purified through a ssDNA-cellulose column to remove excess MDCC. Eluted T7 DNA polymerase was dialyzed against the T7 storage buffer and stored at -80°C. The purity of the final enzyme preparation was analyzed using SDS-PAGE followed by Coomassie Blue staining (Figure 3.2A).

3.2.3 Expression and purification of *E. coli* thioredoxin

E. coli strain SK3981 containing a plasmid pBHK8 was obtained from Dr. Peter Guengerich (Vanderbilt University School of Medicine) as a frozen glycerol culture. The pBHK8 carries a 3000 bp *E. coli* thioredoxin (*Trx*A) sequence cloned into a derivative of plasmid pBR325 (87). Frozen *E. coli* were revived by plating on an LB agar plate with 100 μg/ml of ampicillin and incubated overnight at 37°C. Four colonies were inoculated

into 50 ml of LB with 100 μg/ml ampicillin and incubated with shaking at 37°C for 14 h. The saturated culture (30 ml) was transferred into 6 L of KB medium (0.5% yeast extract, 1% casamino acids, 1.2% glycerol, and 20 ml/L of a balanced salt solution: 40.6 mM $MgSO_4$, 476 mM citrate, 2.9 M KPO_4, and 840 mM sodium ammonium phosphate with 20 μg/ml ampicillin) and grown for 24 hours at 37°C. *E. coli* cells were harvested by centrifugation at 5,000 x g for 15 min and resuspended with lysis buffer (10 mM Tris-HCl, pH 8.5, 3 mM EDTA with 1 mM each PMSF, TLCK, and TPCK) at a concentration of 0.5 g cells/ml. The cell suspension was sonicated for 3 min on ice and centrifuged 20 min at 12,000 x g at 4°C to clear the lysate. Pellets were resuspended with a small volume of lysis buffer and re-centrifuged. Combined supernatant was incubated with constant stirring in a water bath while gradually increasing the temperature from 37 to 63°C. After 10 min of heat denaturation, the lysate was cooled down to 4°C on ice and centrifuged for 20 min at 40,000 x g to remove denatured cellular proteins. The pH and conductivity of the supernatant were adjusted by adding the Tris base and ddH_2O to equal the DEAE loading buffer (20 mM Tris-HCl pH 8.5 and 3 mM EDTA). The lysate was loaded onto a 200 ml DEAE-sepharose column at 2 ml/min and washed with loading buffer until the UV absorbance returned to the base level. Thioredoxin was eluted with a 2 L 0-200 mM NaCl gradient in the same buffer. Fractions containing thioredoxin were pooled and concentrated to less than 10 ml by ultrafiltration through YM-10 membrane. Concentrated protein solution was then loaded onto a 100 ml Sephadex G-50 (Sigma) column with buffer containing Tris-HCl, pH 8.5, and 3 mM EDTA at 0.2 ml/min. Thioredoxin was eluted with the same buffer. Aliquots of purified thioredoxin were quickly frozen in liquid nitrogen and stored at -80°C. The purity of the final protein preperation was analyzed using SDS-PAGE followed by Coomassie Blue staining (Figure 3.2B).
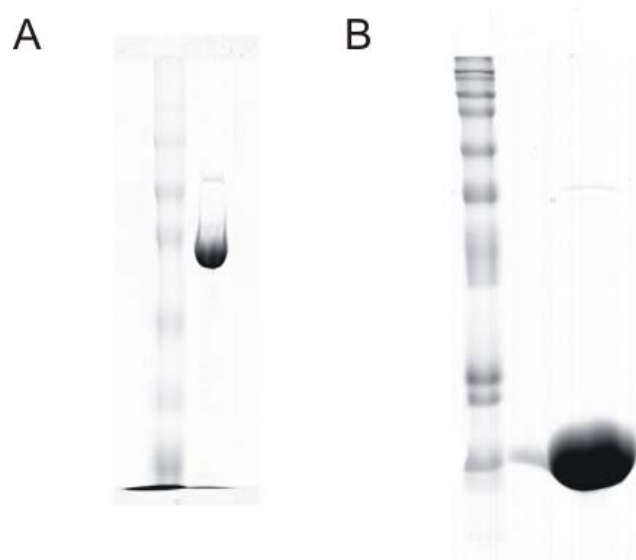
Figure 3.2:   Purified MDCC-E514C-8C and thioredoxin on Coomassie Blue stained
SDS-PAGE gels

(A) The purified MDCC-E514C-8C enzyme migrates as an 80 kD band on the SDS-
PAGE. The protein is analyzed to be 95% pure using the ImageQuant software. (B) The
purified *E. coli* thioredoxin migrates as a 12 kD band on the SDS-PAGE. The thioredoxin
is analyzed to be 98% pure using the ImageQuant software.

3.2.4 DNA substrates used in the kinetic studies

A 27/45-mer primer/template DNA is used in this study. Sequences of the 27mer and 45-mer synthetic DNA oligonucleotides for the T7 DNA polymerase assays were adopted and modified from a previous study conducted on the HIV reverse transcriptase (27-mer: GCC TCG CAG CCG TCC AAC CAA CTC AAC, 45-mer: GGA CGG CAT TGG ATC GAG GTT GAG TTG GTT GGA CGG CTG CGA GGC) (23). The duplex DNA was formed by mixing 27-mer primer and 45-mer template at 1:1 molar ratio, heated to 95°C, and cooled slowly to room temperature for annealing. All DNA oligomers used in the study were custom synthesized by Integrated DNA Technologies and purified by 15% polyacrylamide/7M urea denaturing gel electrophoresis.

3.2.5 Rapid chemical quench-flow experiments

The rapid quench-flow method was used to study the time dependence of nucleotide incorporation under pre-steady state kinetic conditions. A KinTek RQF-3 Rapid Quench Flow apparatus was used to perform the experiments. Typically, one syringe was loaded with 200 nM of T7 DNA polymerase preincubated with 4 μM of thioredoxin and 600 nM of 5'end ^{32}P labeled 27-mer primer/45-mer template substrate DNA in the T7 reaction buffer (40 mM Tris-HCl pH, 7.5, 1 mM EDTA, and 1 mM DTT). Another syringe was loaded with deoxyribonucleoside triphosphate solution at the desired concentration in the same reaction buffer with 25 mM of $MgCl_2$. The reaction was started by the rapid mixing of the reactants from both syringes. The DNA polymerization reaction was then quenched after a preset reaction time with a second mixing of 500 mM EDTA to remove Mg^{2+} ions from the solution and stop catalysis. DNA products from the nucleotide incorporation experiments were analyzed by denaturing polyacrylamide gel electrophoresis. The amounts of product formed were

quantified with phosphorimager and ImageQuant software. Kinetic parameters were derived by fitting to the appropriate equation by non-linear regression.

3.2.6 Stopped-flow experiments

In order to directly measure the rate of nucleotide-induced conformational change of the T7 DNA polymerase. Fluorescent transients following the nucleotide binding to the MDCC-E514C-8C T7 DNA polymerase were recorded with a stopped-flow apparatus (KinTek Model SF-2003). A DNA duplex with dideoxy-CMP (ddC)-terminated primer was used in stopped-flow experiments to prevent the polymerization reaction following the nucleotide binding steps. The fluorescent signal following enzyme isomerization was recorded beginning immediately after mixing enzyme-DNA solution (400 nM of enzyme, 8 μM of thioredoxin, 600 nM of 27-ddC/45-mer, and 12.5 mM of $MgCl_2$ in T7 reaction buffer) and dNTP solutions (varying nucleotide concentration and 12.5 mM of $MgCl_2$ in T7 reaction buffer). The MDCC-E514C-8C enzyme was excited at 425 nm wavelength, and fluorescent emission was monitored with a photomultiplier tube equipped with a 450 nm high-pass optical filter (Corion). Data traces collected in stopped-flow experiments were analyzed by non-linear regression to obtain kinetic parameters.

3.3   RESULTS

3.3.1 MDCC-E514C-8C T7 DNA polymerase has enzyme activity similar to the wild-type enzyme

To specifically label the single cysteine residue introduced into the T7 DNA polymerase, eight of the ten cysteines close to the surface of the enzyme were mutated to create a cys-lite enzyme (Figure 3.1A). These cysteine residues were selected by examining the available ternary complex structure of the T7 DNA polymerase (15). Two of the cysteines, C20 and C451, were mutated to serine residues to preserve the potential hydrogen-bonding capability with adjacent amino acid residues, N171 and Y447. The

58

remaining cysteines were mutated to alanine residues. Amino acid residue E514 of the T7 DNA polymerase is located in the recognition domain on the tip of the fingers domain. It shows a significant movement of 13Å when the ligand-free polymerase binds its DNA and nucleotide substrates (Figure 3.1B) and was selected for site-specific labeling with an environmentally sensitive coumarin fluorophore, MDCC (Figure 3.1C). Prior to the use of this novel DNA polymerase construct for studying the kinetics of enzyme conformational change upon nucleotide binding, the activity of DNA polymerization was compared with the wild-type exo$^-$ enzyme. A single nucleotide incorporation reaction under pre-steady state conditions was performed in the quench-flow apparatus (Figure 3.3). The wild-type exo$^-$ enzyme catalyzes DNA polymerization at a rate of 310 s^{-1}, which is similar to native T7 DNA polymerase (9), while the MDCC-E514C-8C construct catalyzes the same reaction at a rate of 180 s^{-1}. Although this novel construct consists of nine mutations and a MDCC-label, the result showed that the heavily modified polymerase still retains high enzyme activity and could be used in our kinetic studies.
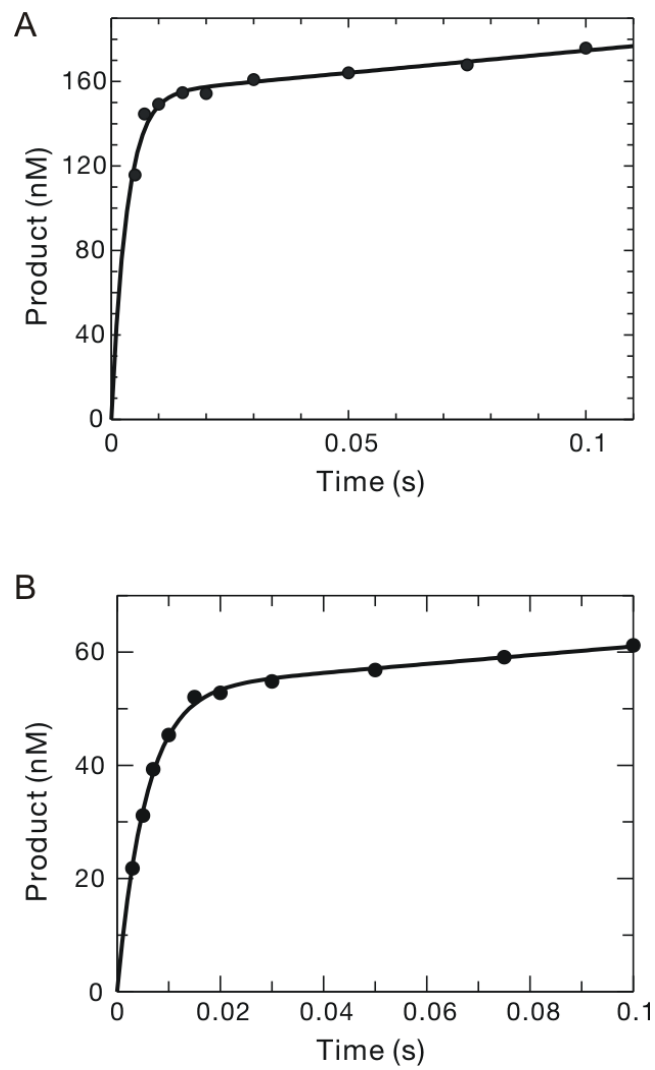
Figure 3.3:   The MDCC-E514C-8C mutant enzyme has similar nucleotide incorporation activity to the wild-type exo⁻ T7 DNA polymerase

Figure 3.3: The MDCC-E514C-8C mutant enzyme has similar nucleotide incorporation activity to the wild-type exo⁻ T7 DNA polymerase

The time-dependent nucleotide incorporation experiment was performed on a quench-flow apparatus to compare the enzyme activities of the wild-type exo⁻ and the novel MDCC-E514C-8C mutant enzymes. Results were fitted to a burst equation ([product] = $A*(1-\exp(-k_1*t))+k_2*t$) by non-linear regression to derive the rate of product formation under pre-steady state conditions. (A) The wild-type exo⁻ enzyme has a rate of 310 ± 27 s^{-1}. (B) The MDCC-E514C-8C enzyme has a rate of 180 ± 6 s^{-1}. Reactions were carried out with 100 nM of enzyme, 300 nM of 27/45-mer DNA duplex, and 50 μM of dCTP at 20°C.

3.3.2 Fluorescent signal from the MDCC-E514C-8C T7 DNA polymerase responds specifically to the binding of correct and mismatched nucleotides

The most critical requirement for studying the kinetics of the isomerization of T7 DNA polymerase during DNA synthesis is to establish a signal that is sensitive to the movement of the recognition domain. In order to test whether the signal from the MDCC label at amino acid position 514 reports conformational changes of the polymerase, the fluorescence emission spectrum of the enzyme was compared at different substrate binding states (Figure 3.4). Starting with free enzyme (E) in solution, adding duplex DNA substrate with a ddC-terminated primer (E-DNA$_{dd}$) results in a small decrease in fluorescence (5% at emission maximum of 460 nm). The small signal change indicates either that the labeling is not reporting the subtle structural changes upon DNA binding or that there are no significant environmental changes surrounding the recognition domain upon DNA binding. This phenomenon was reflected in the fact that the fluorescence signal could not be used to measure the dissociation constant for DNA substrate binding because the level of noise is greater than the substrate-induced decrease of fluorescent signal. In contrast, a significant fluorescent change was observed after adding deoxynucleoside triphosphate substrates into the E-DNA$_{dd}$ complex. With the addition of correctly base-paired nucleotide, dCTP, the fluorescent intensity dropped and reached a minimal level with increasing amount of dCTP added. A 30% drop of the emitted fluorescence was observed (Figure 3.4). If dCTP is replaced with a mismatched deoxyribonucleoside triphosphate, dGTP for example, a 40% increase in the fluorescent intensity at saturating nucleotide concentration was recorded (Figure 3.4). Binding of other mismatched nucleotides also produced ternary complexes with higher fluorescence states. The fluorescence profiles of the MDCC-E514C-8C T7 DNA polymerase at different substrate-bound states show that the signal from MDCC labeling gives a 30 to

40% change from the E-DNA$_{dd}$ state at the emission maximum. In addition, decreasing fluorescence with correct nucleotide binding and increasing fluorescence with mismatched nucleotide binding indicates that the observed signal changes were not artifacts. The signal observed with this novel enzyme reports specific conformational states when the polymerase is complexed with different substrates.
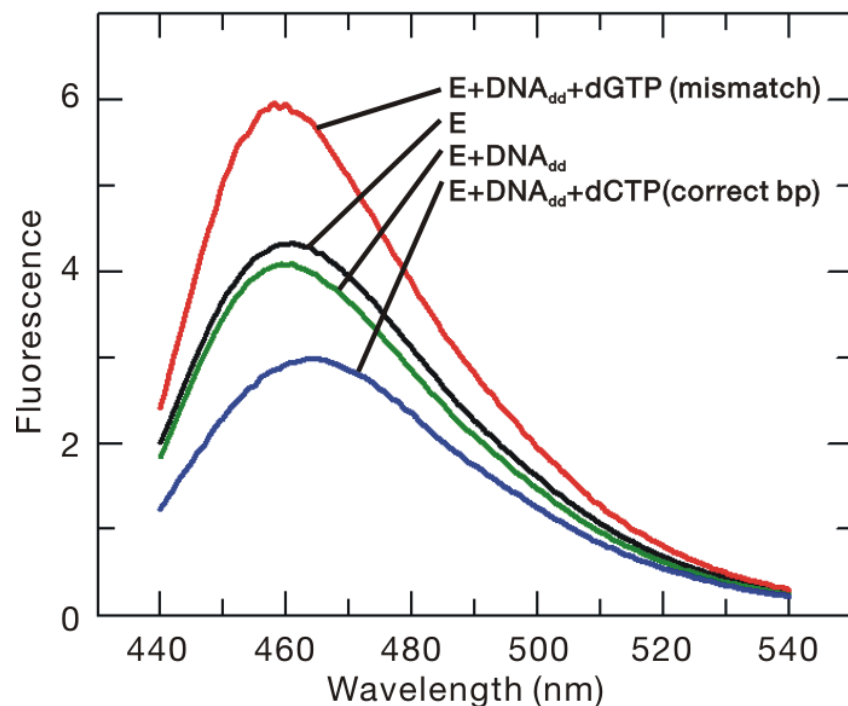
Figure 3.4:   The emission profile of the MDCC-E514C-8C

The fluorescence of MDCC-E514C-8C responds specifically to the binding of different deoxyribonucleoside triphosphate molecules. When the added nucleotide is correctly base-paired with the template base, e.g. dCTP, the fluorescent intensity decreases; while adding a mismatched nucleotide, e.g. dGTP, causes the fluorescence to increase. The fluorescent signals were measured with 200 nM of enzyme and 300 nM of DNA_{dd} duplex at room temperature. 1 µM of dCTP or 1mM of dGTP was added for the measurement of ternary complexes.

3.3.3 Kinetics of the nucleotide-induced polymerase isomerization

The fluorescence emitted from the MDCC-E514C-8C enzyme showed 30 to 40 % changes in intensity depending on the properties of bound nucleotides. The change of the fluorescence indicates that the enzyme isomerization occurs when the nucleotide substrates bind to the E-DNA complex. The significant changes in the fluorescence enable the kinetic studies of the nucleotide-induced enzyme conformational change during DNA synthesis. We set forth to characterize the kinetic properties of the T7 DNA polymerase isomerization during nucleotide binding. A series of fluorescent transients corresponding to the decrease of fluorescence when dCTP binds to an E-DNA$_{dd}$ complex was collected with the stopped-flow apparatus at increasing concentrations of dCTP (Figure 3.5A). The stopped-flow traces were analyzed using non-linear regression with a double exponential equation to determine the rate of enzyme isomerization following nucleotide binding. Previous studies predicted a two-step binding mechanism for nucleotide binding and selection for T7 DNA polymerase (Scheme 3.1)(9;22). If the fluorescence data supports the same mechanism for nucleotide binding, the relationship between the rate of enzyme isomerization and dCTP concentration should follow a hyperbolic equation: $rate = \dfrac{K_{max} \times [dCTP]}{k_d + [dCTP]}$

The non-linear regression analysis shows that the data follows the trend with a fast initial ground-state binding, K_d of 28 ± 6.2 µM, and a slower conformational change step with a maximum rate of 660 ± 52 s^{-1} (k_2) (Figure 3.5B). Although a double exponential equation has to be used to obtain the best fitting results by non-linear regression, the rate of the fast phase is assigned to the actual conformational change step based on the following arguments. First, the amplitude of the fast phase accounts for most of the observed amplitude (86% at 1 µM dCTP) and the maximum rate of the slow phase

is only 20 s^{-1}. Second, only a single phase of fluorescent change is observed when the same experiment is carried out at 10°C (Figure 3.6). The maximum observed rate of enzyme isomerization is reduced to 358 ± 10.3 s^{-1}, which is approximately half of the rate measured at 20°C.

In order to complete the two-step binding kinetic pathway for the correct nucleotide, the rate of enzyme isomerization in the reverse direction was also measured. The experiment was carried out by mixing E-DNA$_{dd}$-dCTP complex with a 10-fold excess of unlabeled E-DNA complex. When dCTP is released from the ternary complex, it is trapped in the unlabeled enzyme complex and polymerized into DNA products. The single exponential increase of fluorescence has a rate of 1.6 ± 0.01 s^{-1} (k_{-2}) (Figure 3.5C). The overall dissociation constant for correct nucleotide binding can be calculated from measured kinetic parameters as 0.07 ± 0.016 μM using the equation $K_d*(k_{-2}/k_2)$.

$$E \cdot D_n \underset{\text{open}}{\overset{K_d}{\rightleftharpoons}} ED_n \cdot dCTP \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} FD_n \cdot dCTP \underset{\text{Closed}}{\rightleftharpoons} FD_{n+1} \cdot PPi$$

Ground-state binding   Conformational change   Chemical step
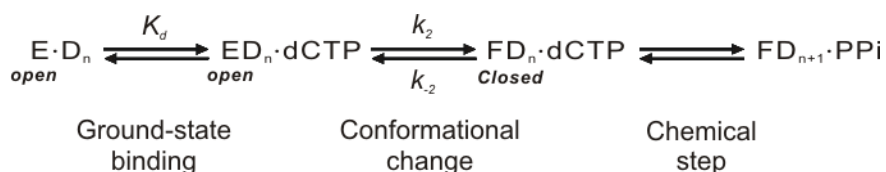
Scheme 3.1: The two-step binding mechanism

66
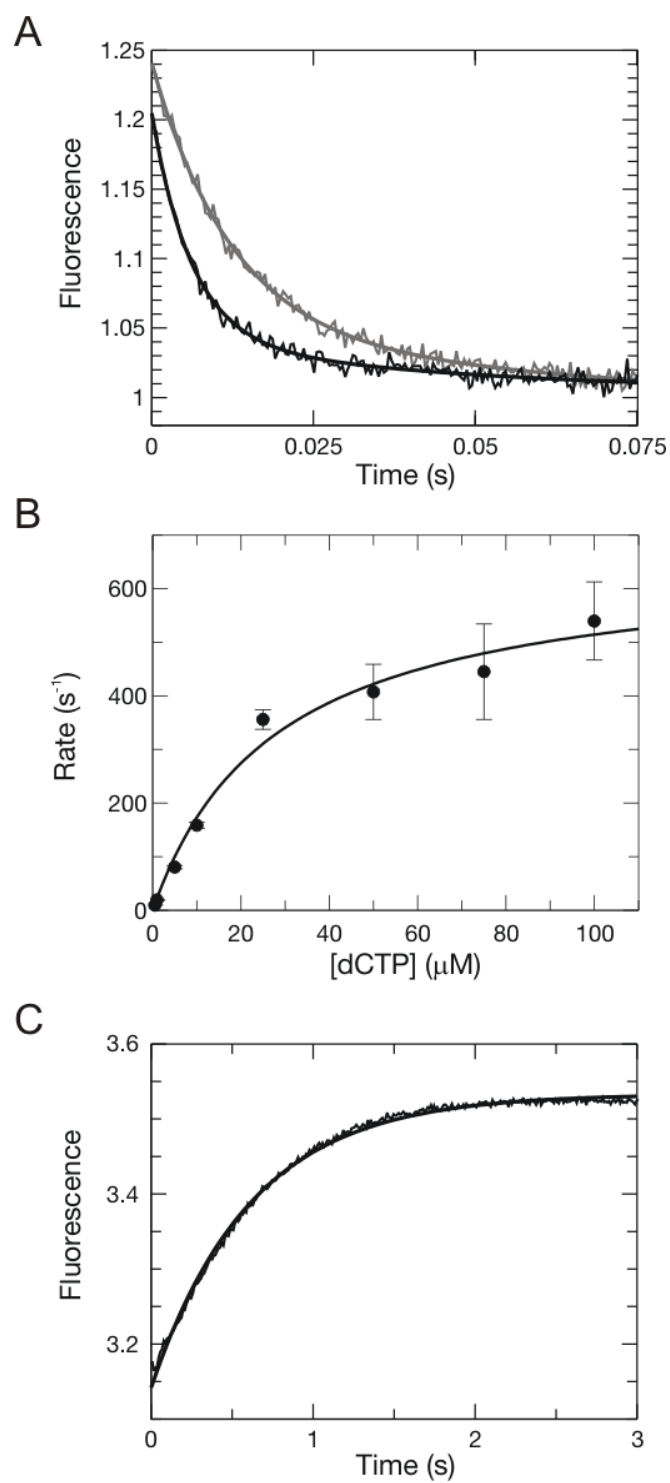
Figure 3.5:   The binding of dCTP to the E-DNA$_{dd}$ complex followed a two-step binding
mechanism

Figure 3.5:   The binding of dCTP to the E-DNA$_{dd}$ complex follows a two-step binding mechanism

(A) Stopped-flow fluorescent transients corresponding to the binding of 5 μM (gray) and 10 μM (black) dCTP to the E-DNA$_{dd}$ complex. The data was best fit to a double exponential equation. The rates are 80 ± 3.1 s^{-1} (k_1) and 17 ± 2.7 s^{-1} (k_2) at 5 μM dCTP, and 159 ± 5.3 s^{-1} (k_1) and 17 ± 1.7 s^{-1} (k_2) at 10 μM dCTP. (B) The dCTP concentration-dependence of the fast-phase rate. The hyperbolic equation was used in non-linear regressions to derive a ground-state binding K_d of 28 ± 6.2 μM and a maximum rate of 660 ± 52 s^{-1}, k_2, for enzyme isomerization following the initial ground-state binding. (C) The release of dCTP from the E-DNA$_{dd}$-dCTP ternary complex was measured in a stopped-flow experiment. The signal corresponding to the release of dCTP follows a single exponential transition with a rate of 1.6 ± 0.01 s^{-1} (k_{-2}). Enzyme 200 nM and 300 nM of DNA$_{dd}$ duplex were used in the measurements carried out at 20°C.
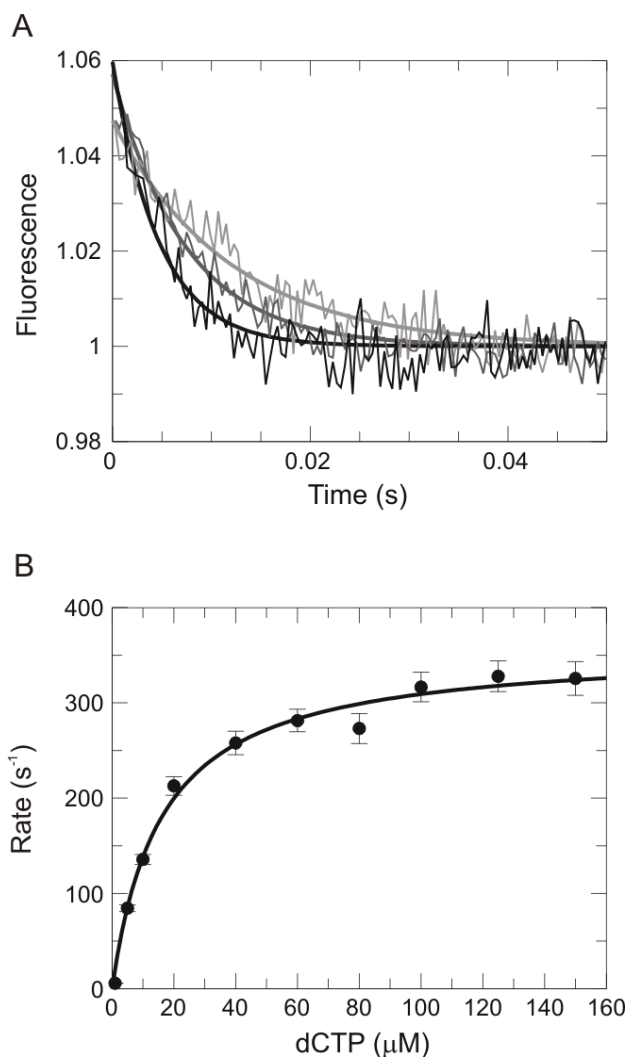
Figure 3.6:   Kinetics of the dCTP-induced conformational change at 10°C

(A) Stopped-flow fluorescent transients corresponding to the binding of 5 μM (gray), 10 μM (dark gray), and 20 μM (black) dCTP to the E-DNA$_{dd}$ complex. The data was best fit to a single exponential equation with rates, 84 ± 3.6 s^{-1}, 136 ± 5.2 s^{-1}, and 213 ± 9.8 s^{-1}. (B) The dCTP concentration-dependence of the conformational change rate. The hyperbolic equation was used in non-linear regression to derive a ground-state binding K_d of 16 ± 1.9 μM and a maximum rate of 358 ± 10.3 s^{-1}, k_2, for enzyme isomerization.

69

3.3.4 The equilibrium titration of fluorescent signals with dCTP

In the previous section, measurement of kinetic parameters for the two-step binding of correct nucleotide, dCTP, is presented. Although the overall dissociation constant for dCTP could be calculated from the stopped-flow experiments, the constant is also measured directly in an equilibrium titration experiment (Figure 3.7). A 200 nM solution of MDCC-E514C-8C T7 DNA polymerase in the T7 reaction buffer and 12.5 mM $MgCl_2$ was preincubated with 300 nM 27-ddC/45-mer DNA duplex. Solutions containing dCTP with 12.5 mM of $MgCl_2$ were used to titrate the enzyme-DNA complex. Fluorescent intensities at equilibration after each addition of nucleotide solution were monitored with a fluorimeter (Photon Technology International) set at 425 nm for excitation and 460 nm for emission detection. The overall dissociation constant obtained by fitting to the quadratic equation is 0.07 ± 0.028 µM. The results from the stopped-flow and equilibrium titration experiments give the same K_d.
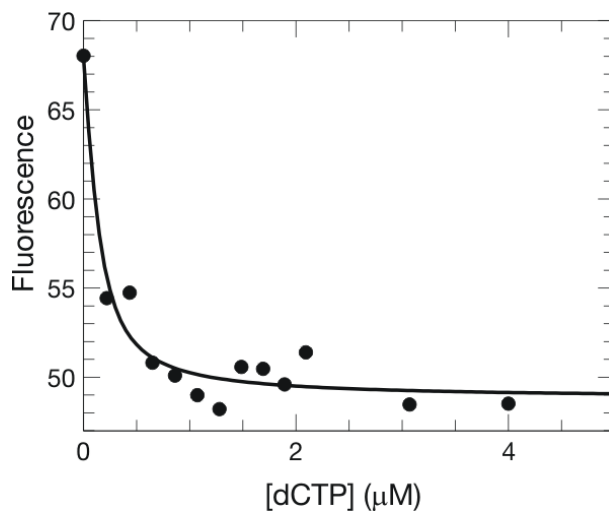
Figure 3.7:　The equilibrium titration of dCTP binding

The overall dissociation constant for the two-step dCTP binding to the E-DNA$_{dd}$ complex was measured by an equilibrium titration experiment. Fluorescent intensities at different dCTP concentrations were fit to a quadratic equation to derive a K_d value of 0.07 ± 0.028 μM.

$$F = F_{initial} - \Delta F \times \frac{(E_0 + K_d + S_0) - \sqrt{(E_0 + K_d + S_0)^2 - 4 \times E_0 \times S_0}}{2 \times E_0}$$

where E$_0$ = 200 nM (concentration of E-DNA$_{dd}$ complex)

3.3.5 The α phosphothioate elemental effect of nucleotide incorporation and enzyme conformational change

The stopped-flow experiment revealed a maximum rate of 660 s^{-1} for the closing isomerization step of the enzyme following ground-state binding of a correct nucleotide. A corresponding rapid chemical quench study showed that the MDCC-labeled enzyme catalyzes the polymerization reaction of dCTP at a maximum rate of 234 ± 9.4 s^{-1} under pre-steady state condition (Figure 3.8A). To get an observed rate of 234 s^{-1} following a 660 s^{-1} step, the actual rate of chemical reaction must be 360 s^{-1}. The next question to ask is whether the 360 s^{-1} is a direct measurement of phosphodiester bond formation or it is limited by steps that occur before the chemistry along the reaction pathway. In order to investigate this problem, the rate of DNA polymerization was measured with an α phosphothioate analog of dCTP (dCTPαS). The pre-steady state rate of nucleotide incorporation is decreased to 4 ± 0.1 s^{-1} when dCTP was substituted with dCTPαS (Figure 3.8B). Comparing the rates of incorporation from the two experiments yields an elemental effect of 90 (360/4 = 90). The results from this study suggest a rate-limiting chemistry step during nucleotide incorporation catalyzed by the MDCC-E514C-8C enzyme.

The dCTPαS was also used in the stopped-flow experiments to determine whether the α phosphothioate analog of dCTP affects the kinetics of nucleotide-induced isomerization preceding the chemistry step. Again, fluorescent transients collected with the stopped-flow apparatus were fit to a single exponential equation to obtain the rate of enzyme conformational change at different nucleotide concentrations (Figure 3.8C). The rate increase corresponding to a higher dCTPαS concentration follows a hyperbola defined by the two-step binding mechanism. The ground-state K_d of 16 ± 2.0 μM and a maximum rate of 680 ± 34 s^{-1} for the conformational change step were derived from the

data (Figure 3.8D). Results show that substituting normal dCTP substrate with dCTPαS reduces the rate of nucleotide incorporation by 90-fold while the rate of nucleotide induced conformational change remains the same.
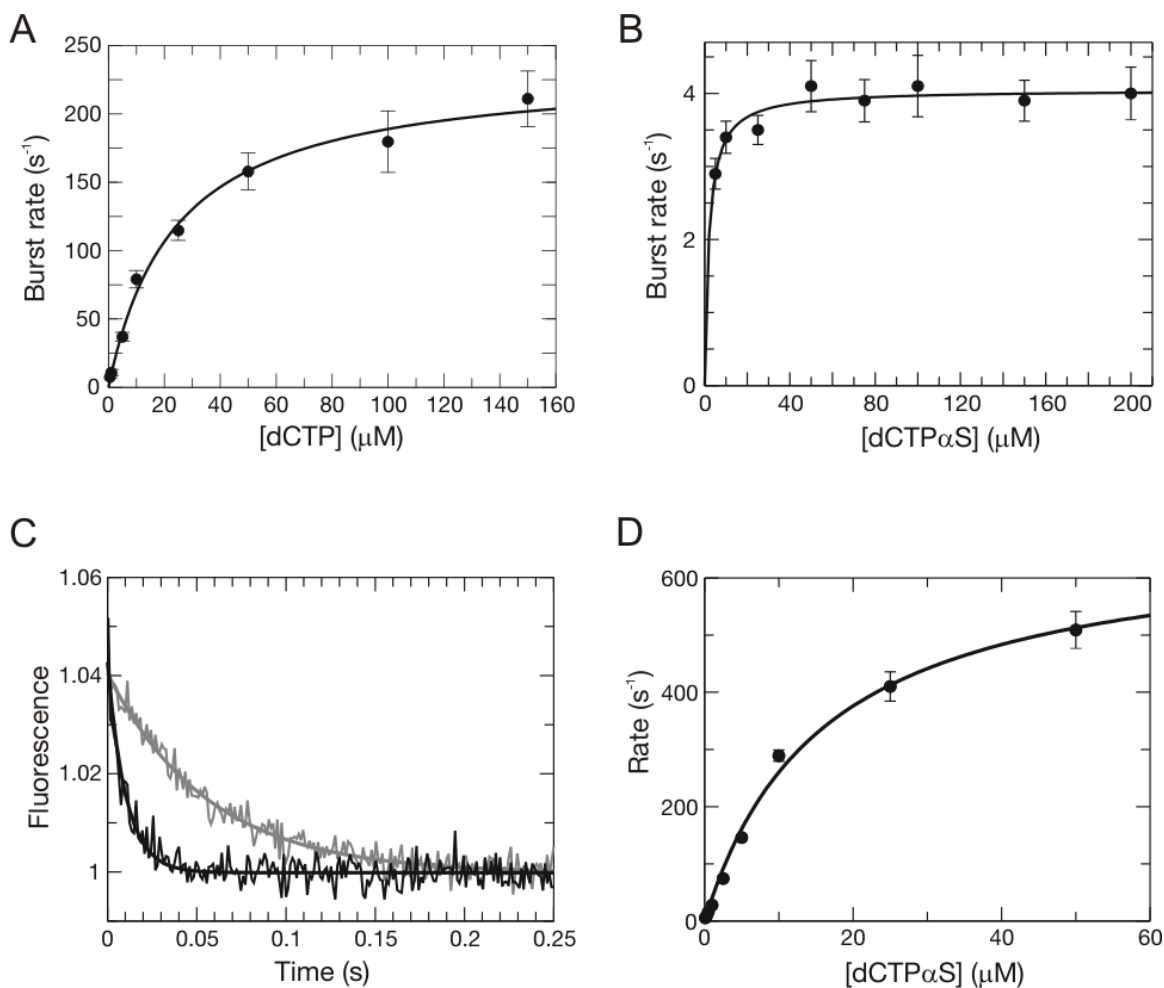
Figure 3.8: The α phophothioate elemental effect on the MDCC-E14C-8C catalyzed nucleotide incorporation reaction

Figure 3.8:   The α phophothioate elemental effect on the MDCC-E14C-8C catalyzed nucleotide incorporation reaction

(A) The maximum rate of dCTP incorporation derived from a concentration-dependent pre-steady state polymerization assay is 234 ± 9.4 s^{-1} with an apparent K_d of 24 ± 3.1 μM for dCTP. (B) Substituting dCTP with dCTPαS in the concentration dependent incorporation assays gave an apparent K_d of 2 ± 0.4 μM for nucleotide binding and a maximum rate of 4 ± 0.1 s^{-1} for dCTPαS incorporation. (C) The fluorescent signal change from the binding of dCTPαS to the E-DNA$_{dd}$ complex follows a single exponential decay. Stopped-flow traces from 1 μM (gray) and 5 μM (black) nucleotide concentrations are shown. (D) The rate of fluorescence decay is plotted against dCTPαS concentrations to extract kinetic parameters defined by the two-step binding mechanism. Fitting the data to a hyperbolic equation gave a ground-state K_d of 16 ± 2.0 μM and a maximum rate of 680 ± 34 s^{-1} for the dCTPαS-induced enzyme conformational change. The pre-steady state rates of DNA polymerization were measured under single turnover conditions, 300 nM of enzyme and 100 nM of DNA substrate. Stopped-flow experiments were carried out with 200 nM of enzyme and 300 nM of DNA$_{dd}$ substrate.

3.3.6 The kinetics of dGTP misincorporation

Previous kinetic studies showed that the T7 DNA polymerase catalyzes the misincorporation of dGTP onto an A template base with a k_{cat}/K_m value of 4.2 $M^{-1}s^{-1}$ comparing to 1.5×10^7 $M^{-1}s^{-1}$ for correct incorporation (9;22). The measurement suggested that the enzyme has a 10^6-fold ($1.5 \times 10^7/4.2$) selectivity for the correct nucleotide over mismatched nucleotide. In order to compare the kinetics of misincorporation between our MDCC-E514C-8C mutant with the wild-type exo$^-$ enzyme, the same parameter for dGTP misincorporation was measured using the 27/45-mer template (Section 3.2.4). Time courses of product formation at different dGTP concentrations were fitted to a single exponential equation to extract the rate of polymerization. The k_{cat} of 0.12 s^{-1} and K_m of 82 μM were determined by fitting the rates of misincorporation as a function of dGTP concentration to a hyperbolic equation (Figure 3.9). The result shows that the MDCC-E514C-8C enzyme has a k_{cat}/K_m value of 1500 $M^{-1}s^{-1}$ for dGTP misincorporation compared with 1.5×10^7 $M^{-1}s^{-1}$ for correct incorporation. The differences in k_{cat}/K_m values suggest that the fluorescently labeled enzyme has a somewhat lower nucleotide incorporation fidelity (10^4-fold discrimination for correct versus incorrect nucleotide) than wild-type enzyme. The lower selectivity may due to the combined effect of protein mutations, fluorescent labeling, and the differences in the DNA substrates between our current study and the published study (22).
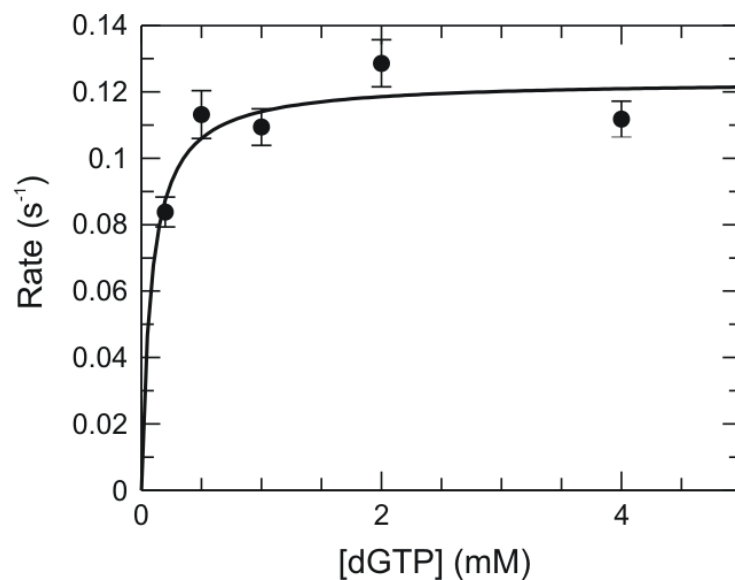
Figure 3.9:   Rates of misincorporation as a function of dGTP concentration

The rate of dGTP misincorporation was measured at five different nucleotide concentrations (0.2, 0.5, 1, 2, and 4 mM). A hyperbolic equation was used to extract the kinetic constants k_{cat} (0.12 ± 0.007 s^{-1}) and K_m (82 ± 37.0 µM) from the data. The reaction was carried out using 300 nM of enzyme and 100 nM of 27/45-mer DNA.

3.3.7 The conformational change induced by the mismatched nucleotide binding

The preliminary emission scan data revealed that the binding of mismatched nucleotide leads to a higher fluorescent state of the enzyme, which was different from the binding of correct nucleotide (Figure 3.4). We proposed that the differences in fluorescent states induced by nucleotide binding are translated from the differences in conformational states when the ternary complexes are formed. In addition, whether the bound deoxyribonucleoside triphosphate is correctly base-paired with the template base or not decides the final conformational state of the ternary complexes. However, the more interesting question to be addressed is whether the incorporation of correct and mismatched nucleotides follows the same reaction pathway. In order to answer this question, we took advantage of the fact that the MDCC-labeled polymerase has a signal reporting the differences in enzyme conformations induced by incorrect nucleotides. Fluorescent transients collected at different dGTP concentrations in the stopped-flow apparatus showed increasing fluorescence upon the binding of the nucleotide to E-DNA$_{dd}$ complex (Figure 3.10A). When the total fluorescent increase from each of the stopped-flow traces was plotted against the concentration of added dGTP, an overall K_d of 60 ± 17 μM could be derived from a hyperbolic equation (Figure 3.10B). The same dissociation constant was also measured in an equilibrium titration experiment and a similar value of 80 ± 11 μM was obtained (Figure 3.11). The higher K_d value indicates that the binding of mismatched nucleotide (dGTP) is about 1000-fold weaker than the binding of correct nucleotide (dCTP). In order to derive a kinetic model that accounts for the data and reflects the fact that each fluorescence curve follows a triple exponential function, a kinetic pathway with three isomerization steps following ground-state binding is suggested. The interrelationships between observed rates and amplitudes are lost in

conventional data fitting to multiple exponential kinetic. Therefore, we fit the data globally to a three-step reaction pathway using computer simulation. All kinetic parameters in the pathway were derived from a global fitting using stopped-flow data and the KinTekSim software (Scheme 3.2). The result from the fitting appears to closely resemble the data we observed in the stopped-flow experiments (Figure 3.10A). Although the rate constants may not represent a unique fit to the data, the model accounts for several unusual aspects of the kinetic data using a minimal model.

Correct nucleotide - dCTP

$$\underset{open}{E \cdot D_n} \xrightarrow{28\ \mu M} \underset{open}{ED_n \cdot dCTP} \underset{1.6\ s^{-1}\ Closed}{\overset{660\ s^{-1}}{\rightleftharpoons}} FD_n \cdot dCTP \xrightarrow{360\ s^{-1}} FD_{n+1} \cdot PPi$$

Mis-matched nucleotide - dGTP

$$\underset{open}{E \cdot D_n} \xrightarrow{86\ \mu M} \underset{open}{ED_n \cdot dGTP} \underset{665\ s^{-1}}{\overset{220\ s^{-1}}{\rightleftharpoons}} FD_n \cdot dGTP \underset{105\ s^{-1}}{\overset{30\ s^{-1}}{\rightleftharpoons}} GD_n \cdot dGTP \underset{7\ s^{-1}\ Low\text{-}catalytic}{\overset{12\ s^{-1}}{\rightleftharpoons}} HD_n \cdot dGTP \overset{0.12\ s^{-1}}{\rightleftharpoons} HD_{n+1} \cdot PPi$$

α phosphothioate analog - dCTPαS

$$\underset{open}{E \cdot D_n} \xrightarrow{16\ \mu M} \underset{open}{ED_n \cdot dCTP\alpha S} \underset{2\ s^{-1}\ Closed}{\overset{680\ s^{-1}}{\rightleftharpoons}} FD_n \cdot dCTP\alpha S \overset{4\ s^{-1}}{\rightleftharpoons} FD_{n+1} \cdot PPi$$
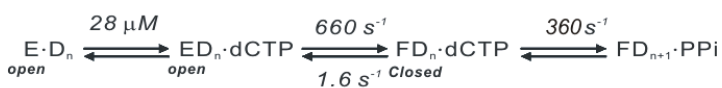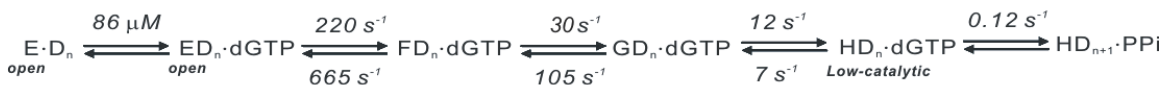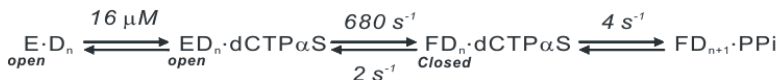
Scheme 3.2:   Pathways of nucleotide incorporation catalyzed by the MDCC-E514C-8C
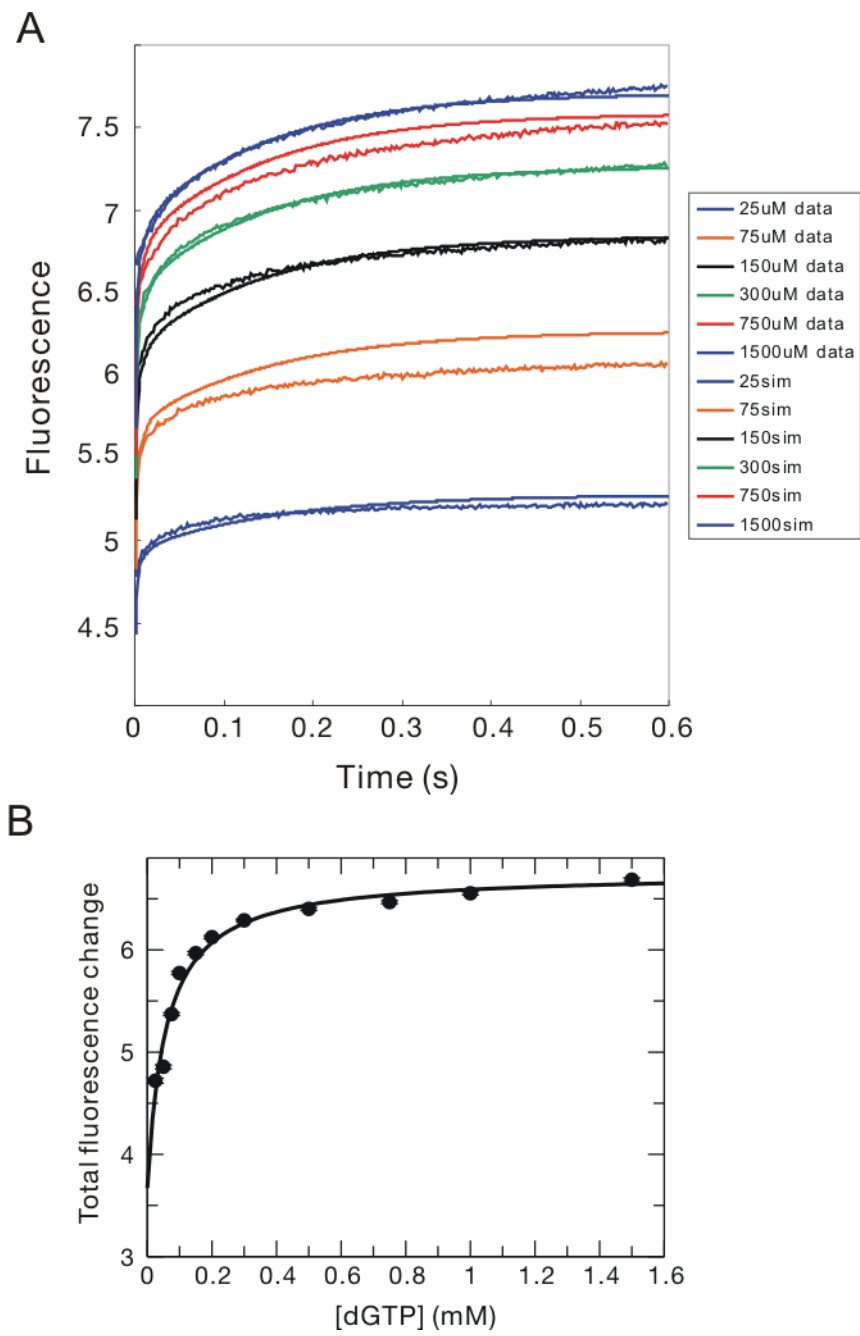
78

Figure 3.10: Kinetics of the conformational change following mismatched nucleotide binding

Figure 3.10: Kinetics of the conformational change following mismatched nucleotide
binding

The binding of the mismatched nucleotide (dGTP) did not follow the two-step binding pathway observed with the correctly base-paired nucleotide. (A) Stopped-flow fluorescent transients of mismatched nucleotide binding with 25 to 1500 μM of dGTP. The fluorescence increase did not follow a single exponential transition; thus, a more complicated mechanism (Scheme 3.2) was used to simulate the observed data (smooth curves). (B) The overall dissociation constant, 60 ± 17 μM, was derived from the stopped-flow data by fitting the total fluorescent change against the added dGTP concentration to a hyperbolic equation. The dGTP-binding experiments were carried out with 200 nM of enzyme with 300 nM of DNA_{dd} at 20°C.
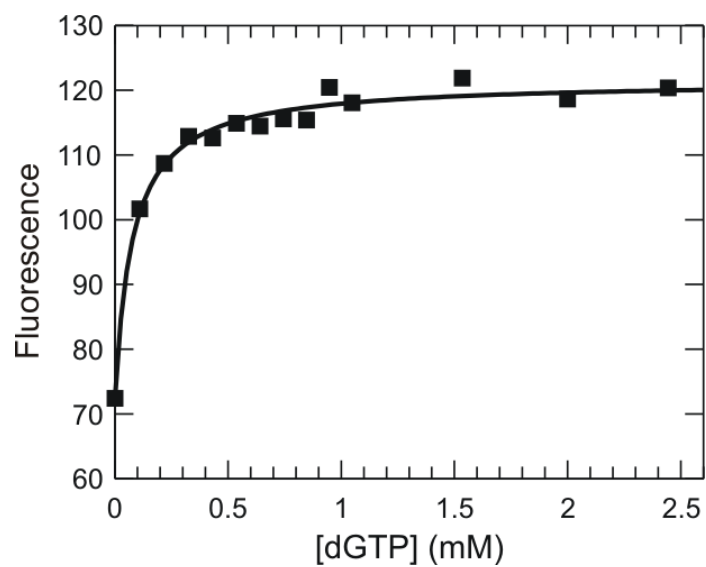
Figure 3.11: Equilibrium titration of the mismatched nucleotide binding, dGTP

The overall dissociation constant for the dGTP binding to the E-DNA$_{dd}$ complex was measured by an equilibrium titration experiment. Fluorescent intensities at different dGTP concentrations were fit to a quadratic equation to derive a K_d value of 80 ± 11 μM.

The most unusual aspect of the data is in the observation that the amplitudes of each phase increase with increasing nucleotide concentration. This can be explained quantitatively by proposing that each isomerization following nucleotide binding is unfavorable thermodynamically. This unique feature of the data demonstrates that the polymerase responds to a mismatched nucleotide in which nucleotide binding energy is not utilized for tighter binding. Rather, we propose it is used to misalign catalytic residues to reduce the rate of catalysis. We refer to the placement of a fluorophore at a key position to sense subtle changes in structure by the term "conformationally sensitive fluorescence".

3.4   DISCUSSION

In this study, we directly measured the kinetics of conformational changes in the recognition domain during DNA polymerization using a fluorescently-labeled T7 DNA polymerase. A novel construct with a single MDCC molecule attached to amino acid C514 was created. This mutant T7 polymerase retained most of its polymerase activity, and the fluorescent signal was sufficient for kinetic studies. The fluorescent signal allows us to observe subtle conformational changes resulting from the binding of either correct or the mismatched nucleotide. The data suggests that the three-dimensional structure of the ternary complex with a mismatched nucleotide must at least have an altered recognition domain when compared to the available crystal structure showing T7 DNA polymerase bound to DNA and a correct nucleotide (15). The markedly reduced rate of catalysis (0.12/360) implies that the key catalytic residues and the substrates are misaligned.

Several kinetic studies done on the pol I family enzymes suggested that the conformational change step limits the net rate of nucleotide incorporation. In addition, this rate-limiting conformational change step follows the ground-state binding of

nucleotides and serves as the most critical step for the selection of correct nucleotides for DNA polymerization. Our data showed that the conformational change step for the correct nucleotide binding occurs before the chemistry at a rate of 660 s^{-1}. The observed rate of dCTP polymerization is 234 s^{-1} for the MDCC-E514C-8C enzyme, which is comparable to the polymerization rate of wild-type enzyme (9;22). To get an observed rate of 234 s^{-1} following 660 s^{-1}, the actual rate of chemical reaction must be 360 s^{-1}. Thus, the conformational change step is partially rate-limiting. The same assay using dCTPαS implied that the mutant protein catalyzes phosphoryl transfer reaction at a much slower rate than the native enzyme. As a result, the observed rate for conformational change does not limit the rate of nucleotide incorporation in the case of the MDCC-E514C-8C enzyme, but a rate of 660 s^{-1} would be at least partially rate-limiting for the wild-type if we hypothesize that the observed elemental effect affects the rate of the chemistry step equally for both T7 enzymes. Another factor may account for the fast rate of conformational change observed in our experiments is the use of a ddNMP-terminated primer as the substrate. The lack of a 3'-OH group on the primer strand could potentially introduce an artifact into our measurement and result in a faster observed rate of conformational change. In conclusion, results from our kinetic analysis suggest that the conformational change step is at least partially rate-limiting in the polymerization reaction catalyzed by native T7 DNA polymerase.

In the substrate binding assays using a mismatched nucleotide (dGTP), a dramatic change in the fluorescent signal was observed. The increase in fluorescence demonstrated that the enzyme adopts a different state from the one resulting from binding of the correct nucleotide. The differences in the kinetics of observed fluorescent transients in the stopped-flow experiments not only suggested an altered conformation induced by the binding of mismatched nucleotide but also revealed a multi-step pathway that leads to

83

misincorporation. The misincorporation fluorescent traces showed a triple exponential increase in signal implying that at least three conformational change steps were occurring before the chemistry step. A reaction pathway consisting of an initial ground-state binding and three consecutive isomerization steps was proposed to describe the observed fluorescent transitions. Kinetic parameters for each of the steps were derived from global fitting (Scheme 3.2). Results show that T7 polymerase binds the mismatched nucleotide with a higher (weaker) ground-state binding K_d, and the conformational change steps leading to polymerization become largely unfavorable. The calculated equilibrium constant for the conformational change step is also reduced from 400 for correct nucleotide binding to 0.2 (product of $K_2K_3K_4$) for mismatched nucleotide binding. In addition, our kinetic analysis revealed that the MDCC-E514C-8C has lower DNA replication fidelity (10^4-fold selectivity compared to 10^6-fold for the wild-type enzyme) which may be a result of enzyme modification and the use of a different DNA template from previous studies.

In order to illustrate the mechanism of nucleotide selection, free energy profiles for correct and mismatched incorporation were calculated using kinetic parameters listed in scheme 3.2, (Figure 3.12A). The traditional thermodynamic free energy profile shows that the conformational change steps leading to misincorporation have higher activation energy barriers than the correct incorporation. Therefore, the reaction pathway for incorrect incorporation is largely unfavorable. However, how the small free energy differences between the formation and the loss of a few hydrogen bonds leads to the exceptional nucleotide selectivity of DNA polymerases can not be explained based upon an oversimplified analysis of the reaction scheme.

Crystal structures of DNA polymerases at different ligand-bound states and previous kinetic studies led to the suggestion of the induced-fit mechanism for enzyme

selectivity (9;15-18;22;23;25). No direct measurement of the enzyme isomerization step was ever achieved. For the first time, conclusions made from the kinetic data collecting from the MDCC-E514C-8C experiments are able to provide clues in solving the polymerase selectivity puzzle. First, thermodynamic calculations showed that the conformational change leading to correct incorporation is favorable (Figure 3.12A). Second, the binding of a mismatched nucleotide leads to a different conformational state of the E-DNA-nucleotide complex which is revealed in the differences in fluorescent signal. Third, the rate for the chemistry step is reduced by 3000-fold for misincorporation. The second and third conclusions combined suggest a sub-optimal alignment of the catalytic residues in the enzyme active site for mismatch incorporation to occur. Based on conclusions from kinetic studies, a novel mechanism is proposed to explain how polymerases achieve high selectivity for nucleotide substrates.

In the case of correct nucleotide incorporation, the weak ground-state binding (K_d = 28 μM) induces a conformational change leading to tighter substrate binding with an equilibrium constant of 400. The conformational change leads to a closed active site and aligns the catalytic residues properly for catalysis (Scheme 3.2). In contrast, if a mismatched nucleotide enters the open active site of the polymerase, the inability to form a correct Watson-Crick base pair would result in a nonproductive collision and lower the ground-state binding affinity (K_d = 86 μM). Moreover, if the mismatched nucleotide stays in the active site and enzyme isomerization starts to occur, not only does the rate of conformational change slow down significantly but it is also unfavorable. These changes shift the equilibrium of the conformational change steps toward the open state and facilitate the dissociation of mismatched nucleotide from the active site. In addition, the rearrangement of the recognition domain is now different from that induced by correct nucleotide binding. This altered pathway for isomerization eventually results in a

85

misalignment of the catalytic residues which is evident by the increase in fluorescent signal. This misalignment slows down the chemistry step for misincoroporation considerably. With this new mechanistic model explaining how DNA polymerase selects correct nucleotides for incorporation, we expand the linear free energy profile to a two-dimensional free energy contour (Fig. 3.12B) and three-dimensional free energy surface (Fig. 3.12C) to include the kinetic and fluorescent information that was not able to provide by traditional free energy profiles.
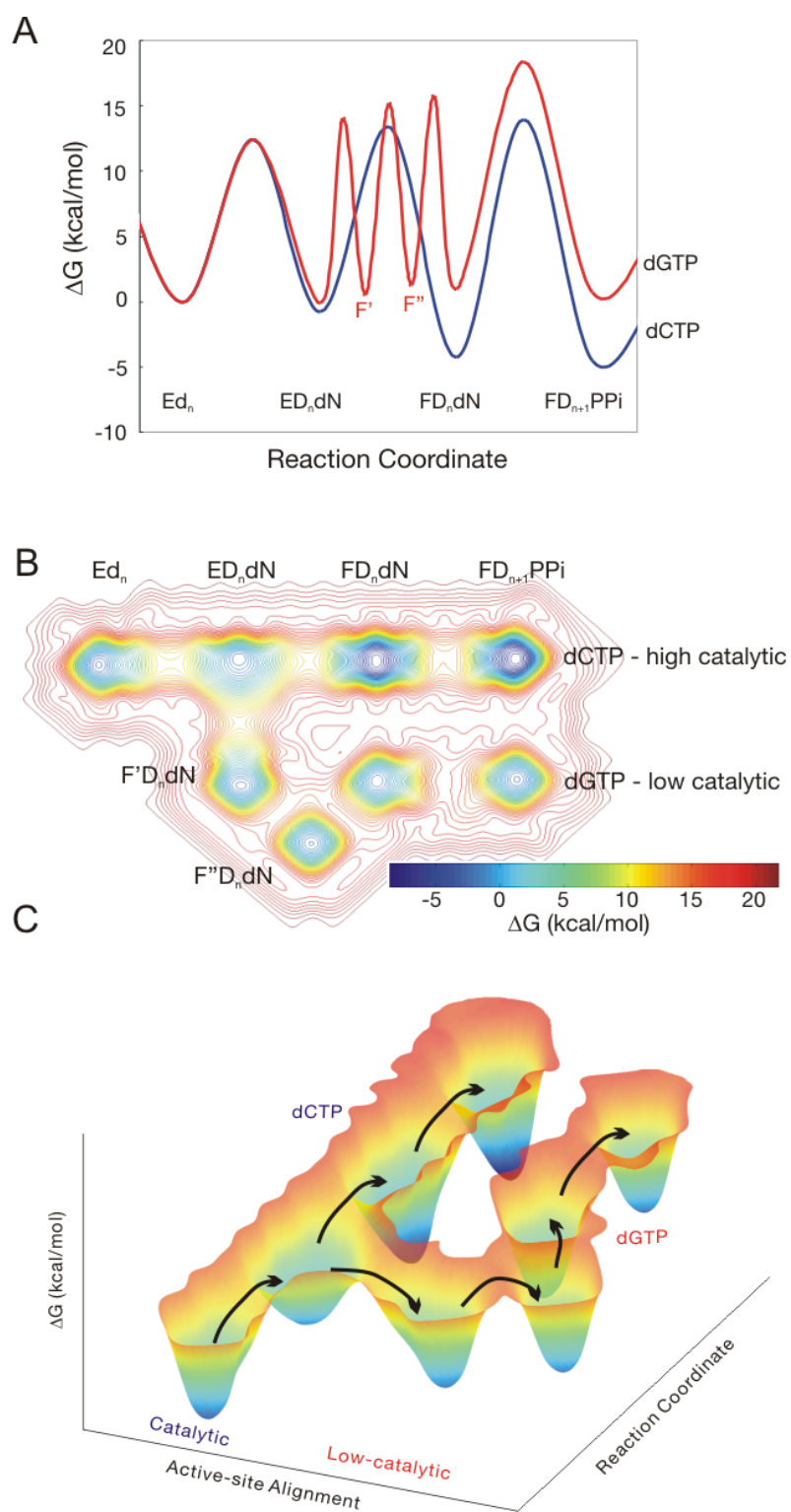
Figure 3.12: The alternative pathway for misincorporation revealed in the kinetic studies

Figure 3.12: The alternative pathway for misincorporation revealed in the kinetic studies

The kinetic analysis of the nucleotide-induced conformational change during DNA polymerization revealed an alternative pathway for the binding and incorporation of mismatched nucleotides. (A) A conventional free energy profile calculated from the kinetic constants is given in scheme 3.2. The pathway that leads to the incorporation of the mismatched nucleotide dGTP is unfavorable compared to the incorporation of the correct nucleotide dCTP. (B) A two-dimensional free energy contour is proposed to include the conformational information from our fluorescent studies. The free energy contour shows that the binding of correct and incorrect nucleotides results in different conformational states of the E-DNA$_{dd}$-dNTP ternary complex through different pathways. (C) A hypothetical three-dimensional free energy surface for T7 DNA polymerase catalyzed reactions. The binding of mismatched nucleotide dGTP may result in a misalignment of catalytic residues in the enzyme active site which greatly reduces the rate of DNA polymerization.

The fluorescence kinetic data obtained in our experiments using the novel MDCC-E514C-8C T7 DNA polymerase provides information for us to further understand the mechanism of nucleotide selectivity during DNA replication. The proposed model suggests that the binding free energy from mismatched nucleotides is used to misalign the key catalytic residues in the enzyme active site and greatly reduces the rate of misincorporation. It is this misalignment of the key catalytic residues that provides the largest contribution to the polymerase fidelity. This model represents a new paradigm for how polymerases achieve high selectivity for correct substrates. By proposing that the binding free energy could be used to align the active site residues differently, a small difference in the ground-state binding free energy can result in high fidelity DNA replication. This model is compatible with the previously proposed induced-fit mechanism but transcends the thermodynamic formalism that is traditionally used to explain enzyme selectivity.

Chapter 4: Kinetics of the T7 DNA Polymerase Conformational Change during DNA Polymerization

4.1 INTRODUCTION

The kinetic data obtained from experiments using the MDCC-E514C-8C construct of T7 DNA polymerase has suggested a mechanism that accounts for how DNA polymerases achieve high-fidelity DNA replication. The key feature of the proposed mechanism is that the binding free energy can be utilized by the polymerase to misalign catalytic residues and reduce the rate of misincorporation by approximately 2000-fold. However, the kinetics of the nucleotide-induced conformational change was measured using a DNA template with a ddCMP terminated primer, and the result may not be applicable to the situation when a primer has a normal 3'-OH group. In order to address this problem, fluorescent signal changes following normal nucleotide incorporation were studied using the stopped-flow method.

The complete DNA polymerization pathway resolved using transient kinetic analysis has suggested that the T7 DNA polymerase cycles through its open and closed conformations along the pathway (9). When the polymerase binds to DNA, it is in the open conformation waiting for a nucleotide substrate. Binding of the correct nucleotide induces an isomerization of the enzyme leading to its closed conformation, which allows optimal substrate binding and facilitates the phosphoryl transfer reaction to occur. The polymerase returns to an open conformation after the chemistry step to allow pyrophosphate release, DNA translocation, and the binding of the next nucleotide. For the MDCC-E514C-8C enzyme, we expect at least two fluorescent transitions, one following the closing of the enzyme after nucleotide binding and a return to the starting state following the opening of the enzyme after chemistry. According to the available

fluorescent data, the stopped-flow traces should first show a decrease and then an increase if the prediction is correct. The nucleotide binding studies carried out using the MDCC-E514C-8C enzyme have provided valuable information that helps us understand the nucleotide recognition mechanism of DNA polymerases. Examination of nucleotide incorporation with the same signal may provide a complete picture of the polymerase conformational changes along its reaction pathway.

4.2   MATERIALS AND METHODS

4.2.1   Single nucleotide incorporation reaction in the stopped-flow apparatus

One of the syringes on the stopped-flow apparatus was loaded with a solution containing 400 nM of MDCC-E514C-8C enzyme, 600 nM of 27/45-mer DNA (sequence listed in Section 3.2.4), and 12.5 mM of $MgCl_2$ in T7 reaction buffer. Another syringe was loaded with a solution containing either dCTP or dCTPαS and 12.5 mM $MgCl_2$ in T7 reaction buffer. The reaction was initiated by mixing contents from both syringes, and fluorescent signals were recorded using a photomultiplier tube equipped with a 450 nm high pass filter. The excitation wavelength for the experiment was 425 nm.

4.2.2   The double-mixing stopped-flow experiment

In order to perform a double nucleotide incorporation experiment, the stopped-flow instrument was set to double-mixing mode (Figure 4.1). The buffer, concentration of enzyme, and DNA used in the experiment were the same as the single nucleotide incorporation reaction (Section 4.2.1). The double-mixing experiment began with the mixing of E-DNA solution with the first nucleotide to be incorporated (e.g. dCTPαS). This first reaction was allowed to proceed for certain amount of time (t1) before the second mixing occurred. The second mixing added the next nucleotide (e.g. dTTP) to the

91

reaction mixture, and the fluorescent transition corresponding to the incorporation reaction of the second nucleotide was recorded.
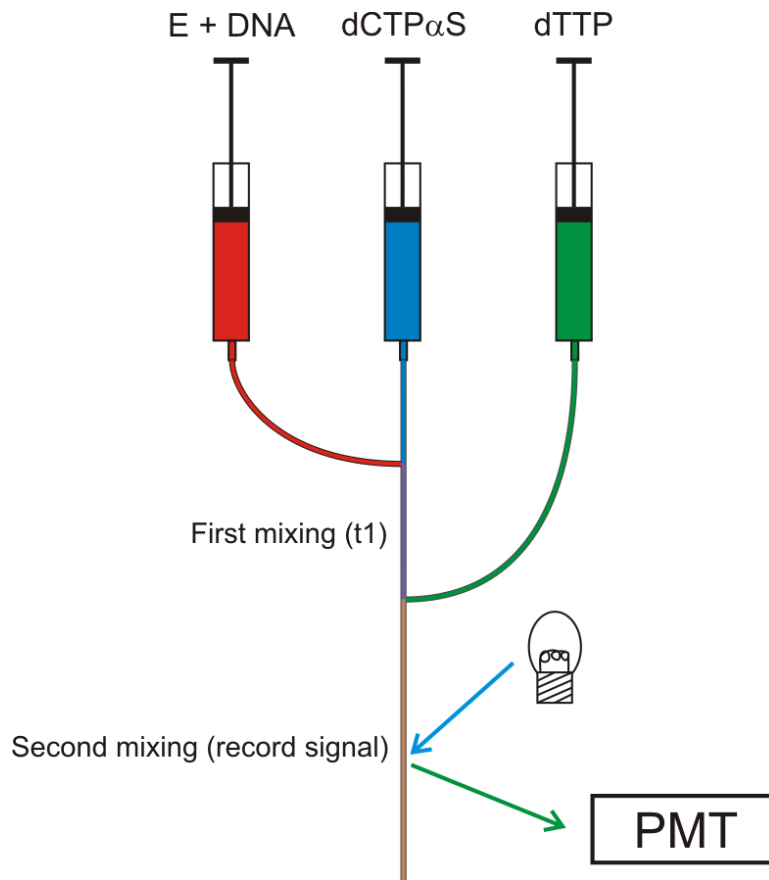
Figure 4.1:   The instrument set-up for double-mixing stopped-flow experiment

The double-mixing mode of the stopped-flow instrument allows two different nucleotide substrates to be sequentially added to the E-DNA solution. At the start of the experiment, E-DNA is mixed with dCTPαS (or dCTP) solution and reacts for a certain length of time (t1). Following the second mixing with dTTP, fluorescent transitions corresponding to the binding and polymerization of the second nucleotide are recorded.

4.2.3 The double nucleotide incorporation reaction

Double nucleotide incorporation reactions were carried out in reaction mixtures containing 100 nM of MDCC-E514C-8C enzyme, 300 nM of DNA template (Figure 4.3A), 100 μM of dCTP, 100 μM dTTP, and 12.5 mM $MgCl_2$ in T7 reaction buffer at 20°C. The quench-flow experiment used a DNA substrate with a ^{32}P-label on the 5' end of the 27-mer primer. Reaction products were separated using denaturing polyacrylamide gel electrophoresis and quantified with a phosphorimager. Stopped-flow experiments were carried out under the same reaction conditions with an unlabeled DNA substrate.

4.3   RESULTS

4.3.1   The fluorescent change during the single nucleotide incorporation reaction

The first experiment to study the conformational change steps in the complete polymerase pathway was to resolve the fluorescent states of the polymerase during a single nucleotide incorporation reaction. Although at least two fluorescent transitions were predicted, the stopped-flow experiment showed that most of the observed signal followed a single exponential increase due to relaxation of the enzyme to the open state after the dCTP incorporation reaction (Figure 4.2A). A small lag observed at the beginning of the 5 μM dCTP trace suggested the existence of a missing phase and that the rate of this initial fluorescence change may be too fast to be fully resolved by the stopped-flow instrument. In an attempt to reveal this missing step, the same experiment was carried out using dCTPαS instead of dCTP. By reducing the rate of nucleotide incorporation, the decrease in fluorescence following the binding of nucleotide became partially observable (Figure 4.2B). However, the amplitude of the fluorescent change corresponding to this initial fast phase was relatively small.
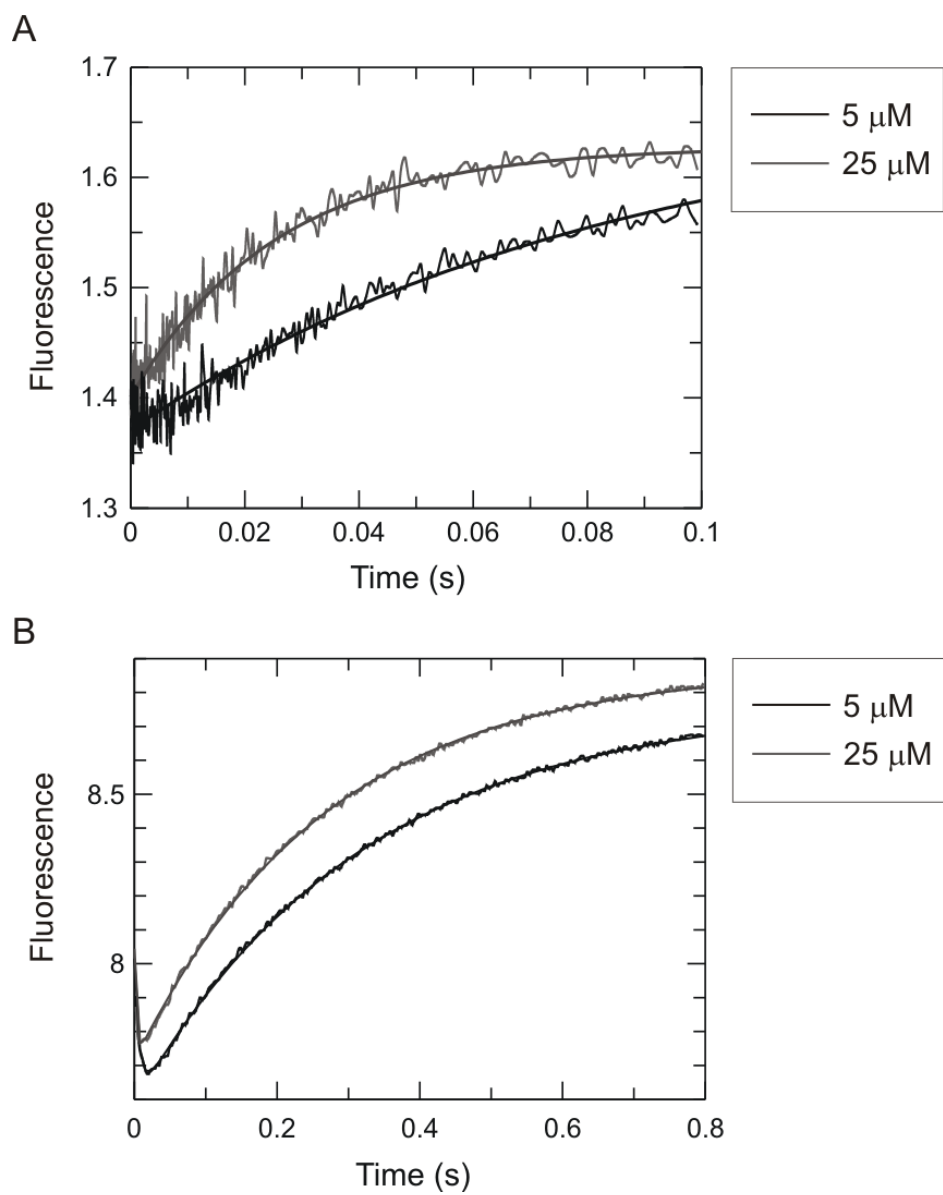
Figure 4.2: Fluorescent transients in dCTP and dCTPαS incorporation reactions

Figure 4.2:   Fluorescent transients in dCTP and dCTPαS incorporation reactions

(A) Fluorescent transient shows a single exponential increase of signal when 25 μM dCTP was added to the dCTP incorporation reaction. The 5 μM dCTP experiment shows a small lag in the beginning of the reaction while the rest of the signal follows a single exponential increase. The rates of fluorescence change are 12 ± 1.4 s^{-1} at 5 μM of dCTP and 39 ± 1.7 s^{-1} at 25 μM of dCTP determined by fitting data to a single exponential equation. (B) Part of the fluorescence change corresponding to the closing conformational change (decrease in fluorescence) upon nucleotide binding is resolved by replacing dCTP with dCTPαS. The fluorescence traces were fitted to a double exponential equation to derive the rates of fluorescence changes. Rates are 148 ± 5.2 s^{-1} and 3.2 ± 0.02 s^{-1} at 5 μM of dCTPαS. The same experiment performed with 25 μM of dCTPαS has rates of 349 ± 35.4 s^{-1} and 3.7 ± 0.02 s^{-1}.

Stopped-flow experiments using dCTPαS have shown that reducing the rate of nucleotide incorporation allows the observation of both the closing and opening conformational changes of T7 DNA polymerase. Another method to slow down the rate of polymerization reaction is to incorporate a normal nucleotide after the α phosphothioate analog has been incorporated into the 3'-end of the DNA primer. A redesigned DNA primer/template substrate allows the incorporation of dTTP after dCTPαS is incorporated (Figure 4.3A). The experiment began with the mixing of E-DNA complex with dCTPαS. When dTTP was added to the reaction, the signal from the incorporation was recorded (Figure 4.3B). In this double incorporation experiment, the MDCC fluorescent transition showed the closing and opening conformational changes in the dTTP incorporation reaction. Because the fluorescent transition predicted by our kinetic data was fully resolved in this experiment, the double incorporation method was used for kinetic studies of enzyme conformational changes in the nucleotide incorporation pathway.

A

5' GCCTCGCAGCCGTCCAACCAACTCAAC
3' CGGAGCGTCGGCAGGTTGGTTGAGTTGGATCTAGGTTACGGCAGG 5'

B
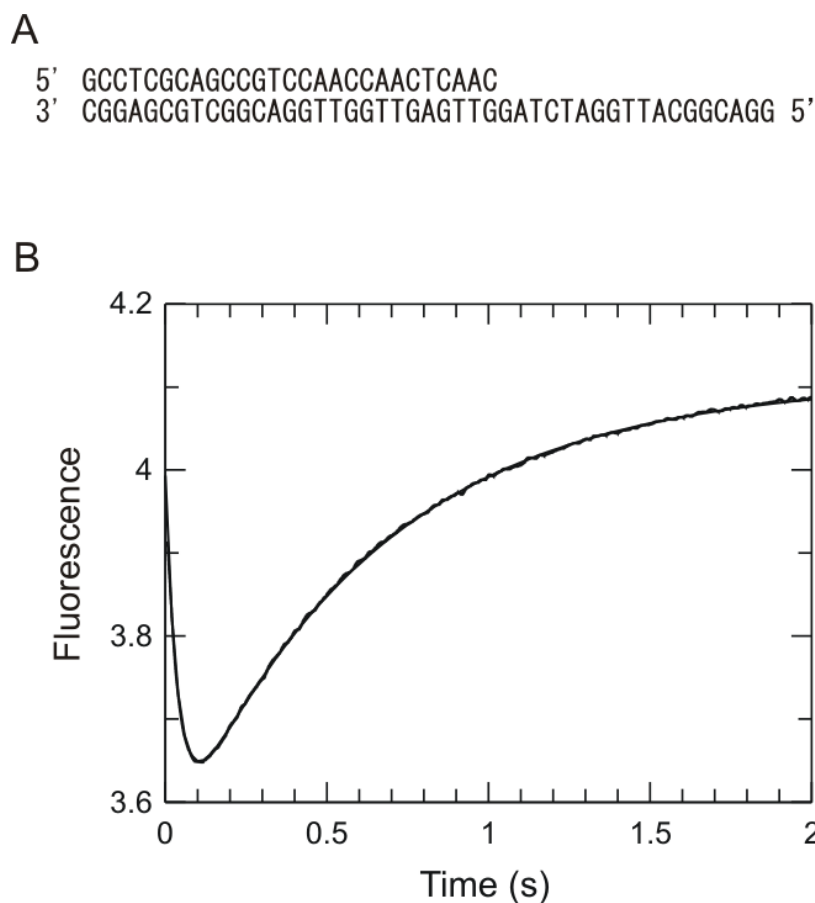
Figure 4.3:   The fluorescent transition following the dTTP incorporation after dCTPαS

(A) The DNA primer/template substrate used in the double incorporation studies. (B) The incorporation of dTTP on top of dCTPαS produced a fluorescent transition with two distinct phases. The decrease in signal represents the closing conformational change induced by dTTP binding. The increase in signal represents the opening conformational change after the chemistry step. Rates of 26 ± 0.2 s^{-1} and 1.6 ± 0.01 s^{-1} were determined by fitting the data to a double exponential equation.

4.3.2 The kinetics of the polymerase conformational changes in a double incorporation reaction

Results presented in the previous section show that a double nucleotide incorporation experiment allows us to observe the fluorescent transitions monitoring the sequential closing and opening conformational changes of the T7 DNA polymerase during a single turnover. In order to establish the relationship between conformational change and a reaction that has two consecutive nucleotide incorporation events, double-mixing experiments were carried out in the stopped-flow instrument.

The first question to be addressed was how the incorporation of the first nucleotide affects the observable conformational changes during the incorporation of the second nucleotide. The experiment was performed by first adding 1 μM of dCTPαS and letting the incorporation reaction proceed for various lengths of time before the addition of dTTP. Immediately after the mixing of dTTP, the fluorescent transition was recorded (Figure 4.4A). The results from the experiment show that as the incorporation of dCTPαS was allowed to proceed for a longer time, the decrease in fluorescence following the dTTP-induced conformational change showed larger amplitude. Furthermore, although the signal from the dCTPαS incorporation reaction was not observed directly, the rate of enzyme opening after the chemistry step can still be estimated from the stopped-flow data using the initial fluorescent intensities from each trace. By fitting the initial intensities to a single exponential equation, a rate constant of 2.4 ± 0.25 s^{-1} represents the opening conformational change was derived. Conformational changes during the dTTP incorporation reaction were also observed in the experiment where dCTPαS was replaced with normal dCTP (Figure 4.4B). The rate of enzyme re-opening after the incorporation of dCTP was 16 ± 3.1 s^{-1}. Although the estimated rates for the opening isomerization are different for dCTP and dCTPαS reactions, the numbers are within two-fold differences

from the actual polymerization rates for each nucleotide at the concentration used (1µM) (estimated from experiments presented in Figure 3.8). The comparison suggests that the observed rate of enzyme re-opening is limited by the chemistry step preceding the conformational change.
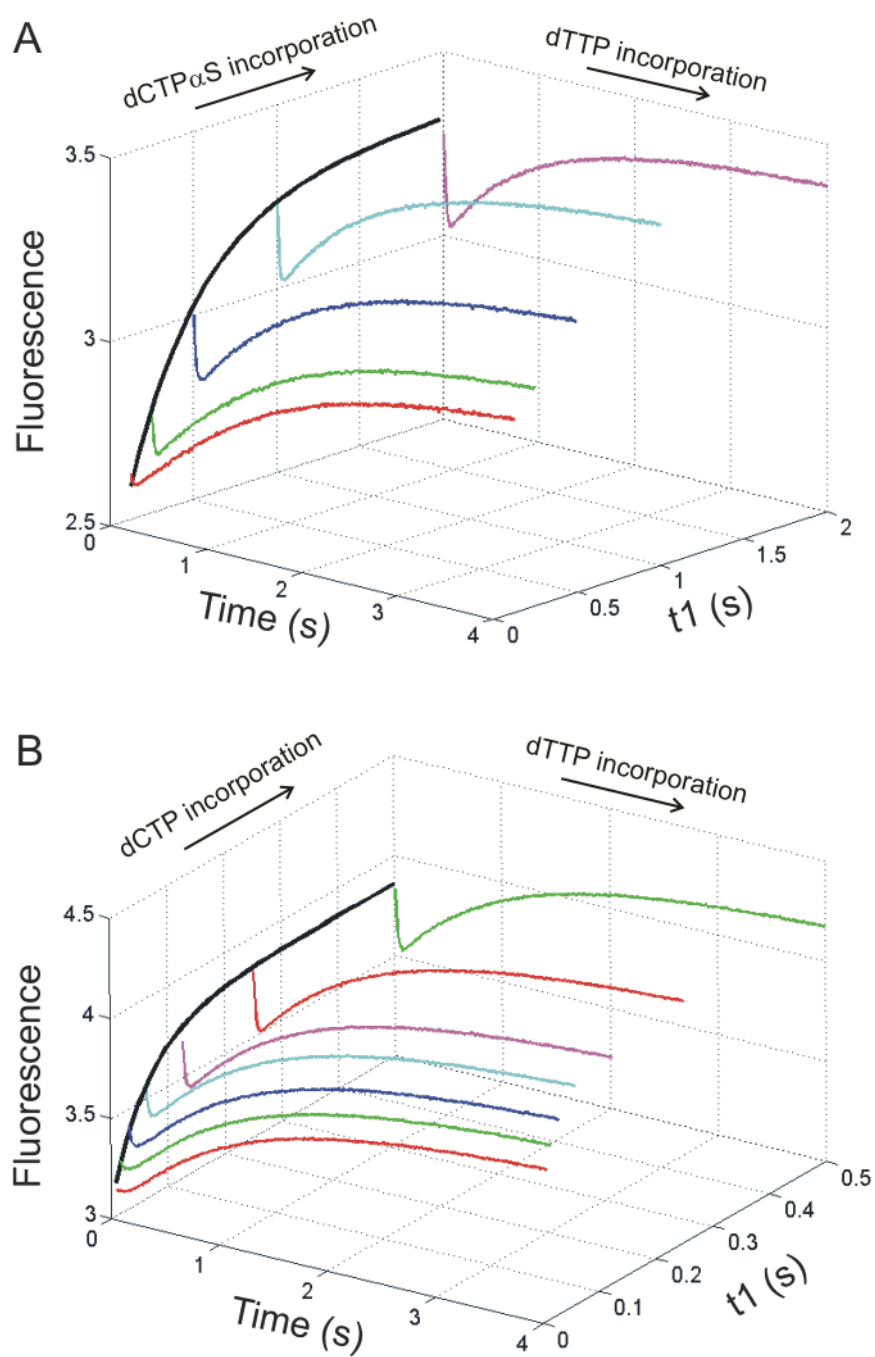
Figure 4.4: Double-mixing experiments with two nucleotides incorporated sequentially into DNA products

Figure 4.4:   Double-mixing experiments with two nucleotides incorporated sequentially
               into DNA products

The conformational changes of the T7 DNA polymerase during dTTP incorporation can be followed with the double-mixing experiments. The addition of dTTP resulted in a fluorescent change that corresponds to the closing (decreasing signal) and opening (increasing signal) steps of the polymerase as the nucleotide incorporation reaction proceeded. (A) Experiments with dCTPαS as the first substrate. (B) Experiments with dCTP as the first substrate. The reaction was carried out with 200 nM of MDCC-E514C-8C, 300 nM of DNA, 1 μM of dCTPαS or dCTP, and 75 μM of dTTP. Single exponential fits of the initial fluorescent data were plotted as black curves with a rate of 2.4 ± 0.25 s^{-1} for dCTPαS and 16 ± 3.1 s^{-1} for dCTP.

4.3.3 The kinetics of the conformational changes during the incorporation of dTTP

In order to establish detailed conformation kinetics of the T7 DNA polymerase in a nucleotide incorporation reaction, fluorescent transition was examined in double-mixing experiments at various dTTP concentrations. The result shown in Figure 4.4B suggests that the dCTP incorporation reaction approaches completion in 0.5 s reaction time (t1). Also, an optimal signal change is available for monitoring the conformational changes during dTTP incorporation when t1 is greater than 0.5 s. Therefore, the double-mixing experiment was carried out using 1 μM of dCTP for 0.5 s in the first incorporation reaction. After dCTP was incorporated into DNA, different concentrations of dTTP were added and the fluorescent changes resulting from dTTP incorporation were recorded (Figure 4.5).

According to previous discussions, the decrease in fluorescence represents the closing conformational change and the increase in fluorescence represents the opening conformational change. The stopped-flow data revealed that the closing isomerization step leading to dTTP incorporation has a maximum rate of 82 ± 3.0 s^{-1} at saturating dTTP concentration. Following the chemistry step, the enzyme returns to its DNA-bound open state at a maximum rate of 0.6 ± 0.01 s^{-1} (Figure 4.6).
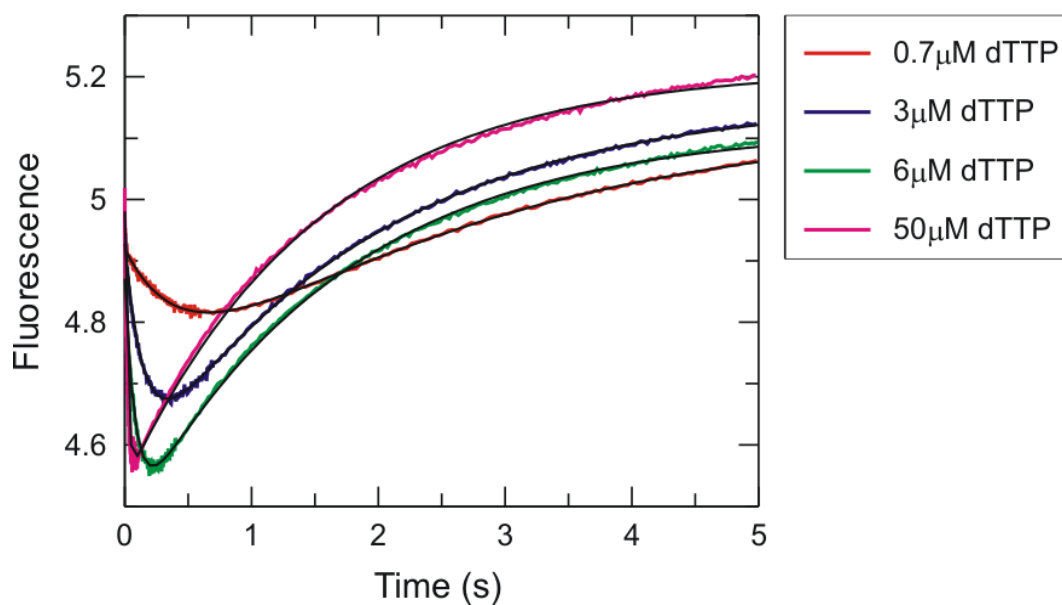
Figure 4.5: The concentration dependence of fluorescent signal changes corresponding to the incorporation of dTTP

The fluorescent signal changes during dTTP incorporation catalyzed by the MDCC-E514C-8C enzyme have two exponential phases. Fluorescent traces were fit to a double exponential equation to derive the rates of enzyme conformational changes. Fitted curves were plotted in black.
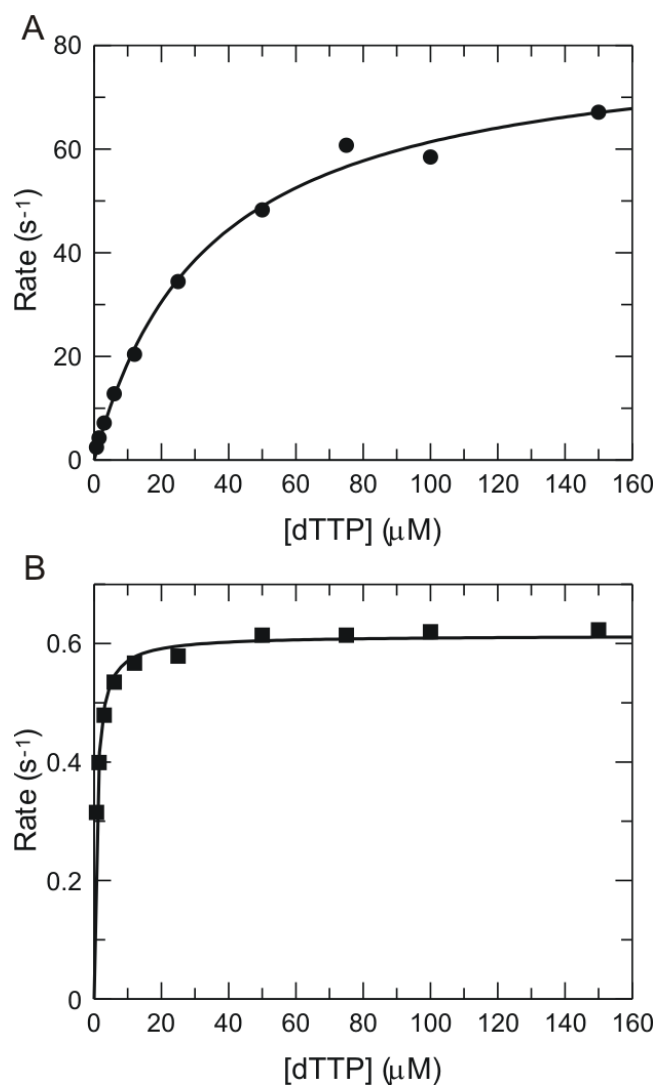
Figure 4.6:   The dTTP concentration dependence of the conformational changes

The rate of conformational changes during dTTP incorporation increases with the concentration of dTTP. The maximum rate of each phase was derived by fitting the data to a hyperbolic equation. (A) The maximum rate corresponding to the closing isomerization is 82 ± 3.0 s^{-1} with an apparent K_d of 34 ± 3.5 μM. (B) The maximum rate corresponding to the re-opening of enzyme is 0.6 ± 0.01 s^{-1} with an apparent K_d of 0.8 ± 0.05 μM.

4.3.4 Conformational changes of the T7 DNA polymerase during misincorporation

In order to study the conformational changes of T7 DNA polymerase in a misincorporation reaction, fluorescent transitions corresponding to the misincorporation of dGTP following dCTP incorporation were recorded with the stopped-flow apparatus (Figure 4.7). The double-mixing experiment was carried out with 100 nM of MDCC-E514C-8C, 300 nM of DNA substrate, 1 μM dCTP, and 50-1500 μM of dGTP. The reaction time for dCTP incorporation was 0.5 s. The results showed that a fast drop in fluorescence occurred upon the addition of dGTP and could not be resolved in the stopped-flow experiments. This phenomenon suggested that the ground-state binding of dGTP was fast and reached equilibrium in a short time. Although some of the fluorescent traces showed a small drop at the beginning of the reaction, e.g. the 50 μM, 150 μM, and 200 μM dGTP traces, the fluorescent transitions corresponding to ground-state binding were not observed in most of the reactions. However, the initial fluorescence was used to estimate the K_d of the ground-state binding of dGTP by fitting the fluorescent intensity to a quadratic equation. The dissociation constant obtained by this method is 280 ± 88 μM (Figure 4.8A). Most of the fluorescent transitions observed in this experiment showed a double exponential increase, which was similar to the fluorescent transitions observed in the mismatched nucleotide binding studies (Figure 3.9). In order to study the kinetics of conformational changes during dGTP misincorporation, fluorescent traces at each of the dGTP concentrations were fit to a double exponential equation to obtain the rates of the fluorescent changes. When the rate obtained from the double exponential fit was plotted against the dGTP concentration, the maximum rates of 4.7 ± 0.63 s^{-1} and 1.6 ± 0.34 s^{-1} for each conformational change step was derived by fitting the data to a hyperbolic equation (Figure 4.8B).
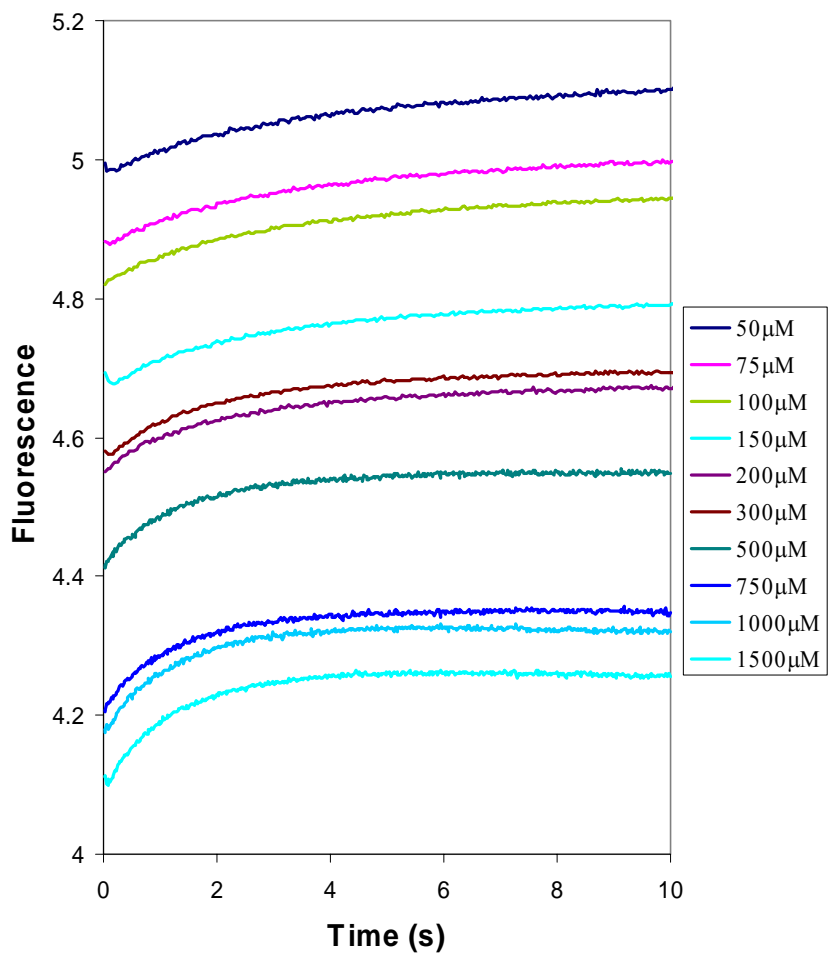
Figure 4.7: Fluorescent transients corresponding to dGTP misincorporation

The fluorescent transients of dGTP misincorporation were recorded in a double-mixing experiment. The loss of fluorescent signal due to the ground-state binding of dGTP was not resolved in the experiment. Most of the fluorescent transitions follow a double exponential increase.
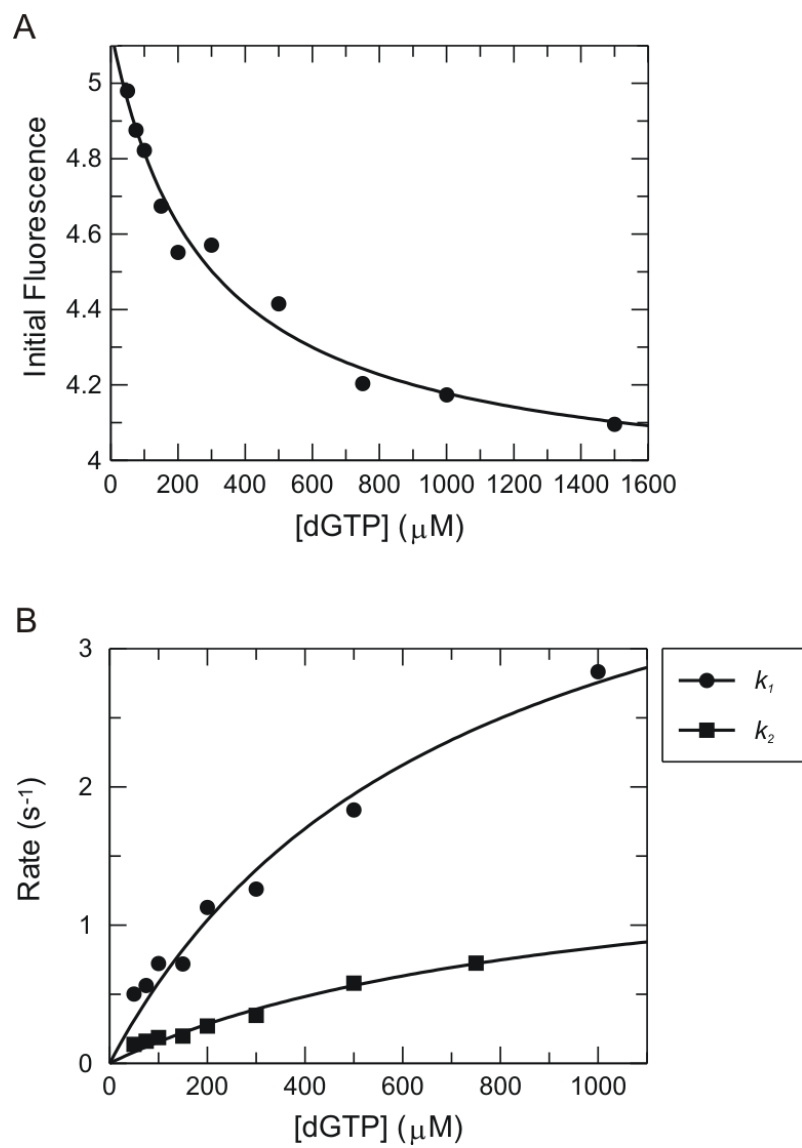
Figure 4.8: The kinetics of the conformational changes during dGTP misincorporation

Figure 4.8: The kinetics of the conformational changes during dGTP misincorporation

(A) The fast equilibrium of ground-state binding allows us to estimate a K_d of 280 ± 88 µM for the ground-state binding step by fitting the initial fluorescent intensities to a quadratic equation. (B) The rates of conformational changes following the ground-state binding were obtained from double exponential fits of the traces. By fitting the rate of each phase to a hyperbolic equation, maximum rates for the two conformational change steps were derived. The maximum rate of the fast phase k_1 is 4.7 ± 0.63 s^{-1} with an apparent K_d of 710 ± 166 µM. The maximum rate of the slow phase k_2 is 1.6 ± 0.34 s^{-1} with an apparent K_d of 970 ± 294 µM.

4.3.5 The kinetics of the enzyme conformational changes during continuous incorporation of two nucleotides

Fluorescent studies have revealed the dynamics of the T7 DNA polymerase conformations throughout the nucleotide incorporation pathway. However, the relationship between the kinetics of actual DNA elongation and observed conformational changes was not established in the stopped-flow experiments. In order to directly relate the enzyme conformations to its catalytic activities, a double nucleotide incorporation reaction was performed on both the stopped-flow and quench-flow instruments. This approach allowed us to observe the conformational changes along with the processive nucleotide incorporation on the same time scale.

Results from the quench-flow studies showed a burst of 28-mer DNA formation, which corresponds to the fast incorporation of dCTP at the initial stage of the reaction. As the reaction progresses, the amount of 28-mer product reached a maximum level and then started to decline due to dTTP incorporation, which converts 28-mer into 29-mer DNA (Figure 4.9, upper panel). By performing the same experiment in a stopped-flow apparatus, the fluorescent data could be compared directly with the time course of chemical reactions (Figure 4.9 lower panel). Because the kinetic parameters for the reaction pathway of incorporating both dCTP and dTTP are not governed by simple equations, kinetic parameters for each of the steps cannot be derived by conventional data fitting using quench-flow data. In order to establish a reaction pathway that accounts for the experimental results from the quench-flow study, the global fitting method was used to derive each of the parameters in a proposed reaction mechanism (Scheme 4.1). Computer simulation using the derived kinetic parameters shows that the proposed mechanism for the two nucleotide incorporation reaction can be used to describe the observed product formation patterns closely (Figure 4.9, upper panel).

109
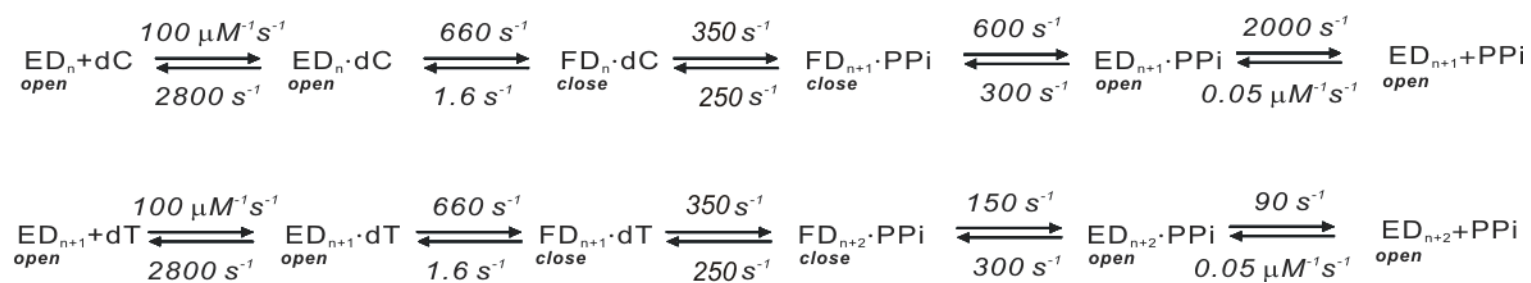
Scheme 4.1: The reaction pathway for the sequential incorporation of dCTP and dTTP
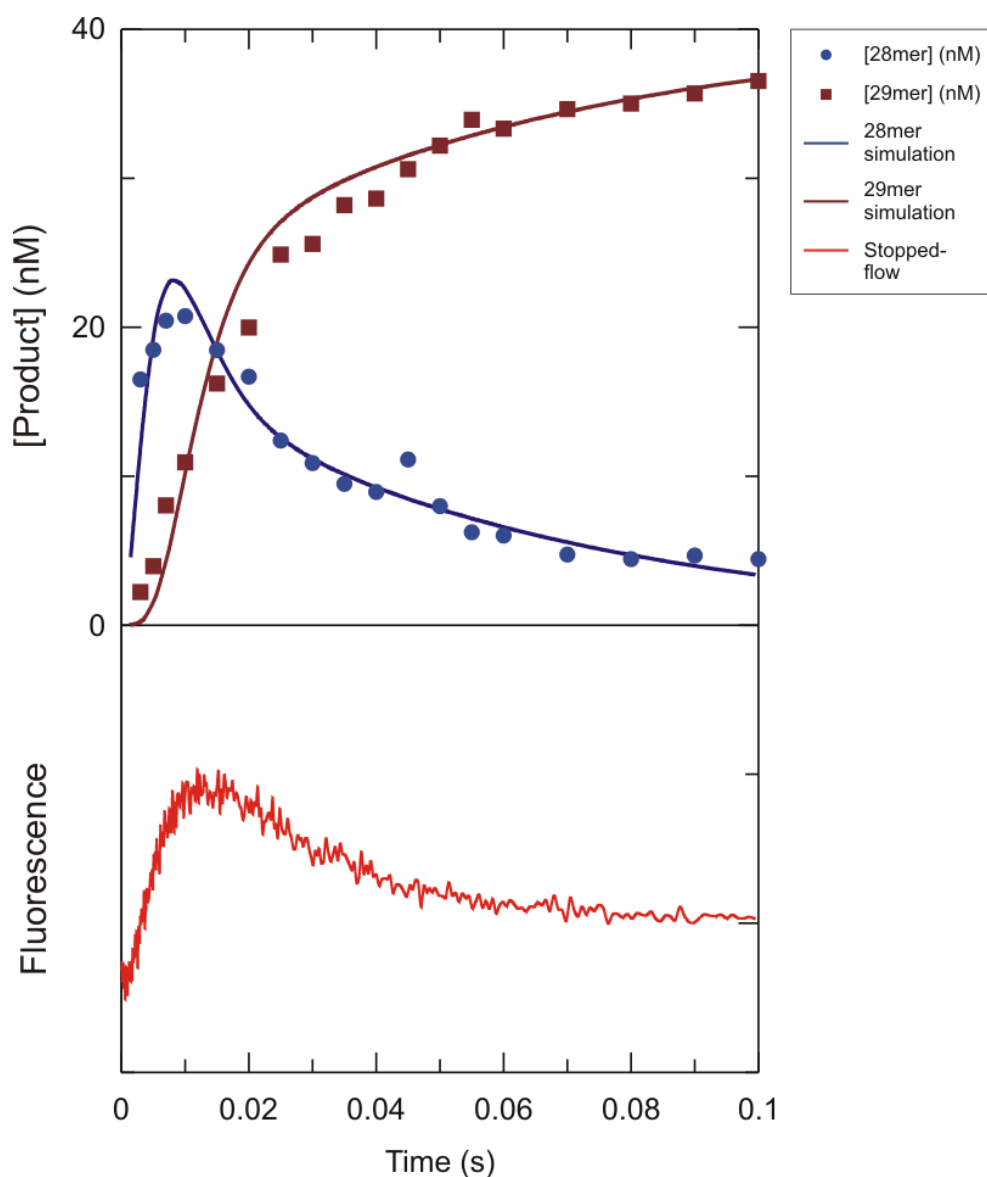
Figure 4.9:   Conformational changes during a two-nucleotide incorporation reaction

(A) Quench-flow experiment showing the relative amount of 28-mer and 29-mer products during a 0.1 s reaction time course. Solid lines represent the computer simulated results using the kinetic mechanism in Scheme 4.1. (B) The fluorescent transient recorded in a corresponding stopped-flow experiment.

4.4   DISCUSSION

Although prior kinetic studies of the T7 DNA polymerase have detailed the enzyme mechanism, the kinetics of the conformational changes during DNA polymerization was not measured directly. Results presented in the previous chapter have shown that the fluorescent signal from the novel MDCC-E514C-8C T7 DNA polymerase can be used to study the polymerase conformational changes. The fluorescent study also provided information to understand how correct nucleotides are selected during DNA polymerization. In order to further understand the kinetics of enzyme conformational changes during nucleotide incorporation, the fluorescent changes of the polymerase were examined in the double-mixing stopped-flow experiments.

Our first attempt to directly observe conformational changes in a dCTP incorporation reaction did not show a decrease in fluorescence due to the closing conformational change step induced by nucleotide binding. Instead, the result suggested that the closing isomerization induced by dCTP binding was too fast to be resolved in a normal incorporation reaction. Alternatively, the change in enzyme state upon dNTP binding may not produce a fluorescence change with DNA that is not terminated by a dideoxynucleotide. Although slowing down the polymerization reaction by replacing dCTP with dCTPαS permitted the observation of the missing conformational change step, the amplitude of the fluorescent change was relatively small. The second approach was to perform a double nucleotide incorporation reaction in the stopped-flow apparatus. The double-mixing experiments clearly revealed two conformational change steps during the incorporation reaction of dTTP. The fluorescent transition corresponding to dTTP binding was fully resolved as the incorporation of the first nucleotide, dCTPαS or dCTP, approached completion. The experiment also showed a shifting of stopped-flow

112

transients toward higher fluorescent states as the time for first incorporation lengthened. Initial fluorescent states from each double-mixing reaction trace allowed the reconstitution of fluorescent changes during the polymerization of both dCTP and dCTPαS (black curves shown in Figure 4.4). The increase in the initial fluorescent intensity is presumably the re-opening of the polymerase after the chemistry step. As more enzyme molecules returned to the open state and were ready for the incorporation of the next nucleotide, the amplitude of the signal changes due to dTTP binding also increased. The rates of the opening conformational changes following the incorporation of dCTP and dCTPαS are within two-fold differences from the rates of the chemical reactions. The small difference between the rate of fluorescent change and the chemical reaction suggests that the rate of opening isomerization is limited by the chemistry step. Moreover, the lack of fluorescent signals showing the reopening of enzyme at short delay times (t1) in double-mixing experiments may signify that the polymerase does not necessarily return to its fully open state under processive incorporation conditions.

The double-mixing experiment with various dTTP concentrations showed that the rate of closing isomerization increases with the dTTP concentration and has a maximum rate of 82 ± 3 s^{-1} with an apparent K_d of 34 ± 3.5 μM for dTTP binding (Figure 4.6A). Although the rate is slower than that measured in the dCTP binding experiments with a ddCMP-terminated primer, it is comparable to the product formation rate of 29-mer in the double nucleotide incorporation reaction (Figure 4.9), which has a rate of approximately 50 s^{-1} at the beginning of the reaction when the data was fit to an exponential equation. This result suggests that the kinetics of nucleotide-induced isomerization of T7 DNA polymerase is slower during processive incorporation conditions. In contrast, the rate of this same step is too fast to be resolved by the stopped-flow instrument during the incorporation of the first nucleotide. In conclusion, the results suggest that the T7 DNA

113

polymerase may adopt different mechanisms for replication initiation and processive nucleotide incorporation. In particular, it is possible that during processive synthesis some of the nucleotide binding energy is used to drive translocation of the DNA.

When dTTP was replaced with dGTP (mismatch) in the stopped-flow studies, only an increase in fluorescence following the fast ground-state binding was observed. The measured apparent dissociation constant for the ground-state binding was 280 μM. The fluorescent changes corresponding to the nucleotide-induced isomerization followed a double exponential with a maximum rate 4.7 s^{-1} for the fast phase and 1.6 s^{-1} for the slow phase. The apparent dissociation constants for the corresponding steps were 710 μM and 970 μM (Figure 4.8). Kinetic parameters derived from this experiment show weaker ground-state binding and slower conformational changes that lead to misincorporation. The kinetic consequences of the conformational change steps in the misincorporation reaction pathway are similar to that observed in the nucleotide binding studies using a ddCMP-terminiated primer in the DNA substrate. Therefore, the proposed mechanism of misaligning catalytic residues to slow down the rate of the chemistry step during misincorporation is also observed in the double incorporation reaction.

The double-mixing experiments have provided a signal to study the kinetics of closing and opening conformational changes during the incorporation of a single nucleotide. In order to assess the dynamics of T7 DNA polymerase conformations under processive polymerization conditions, the product formation pattern was compared directly with the fluorescent signal in a double-nucleotide incorporation reaction (Figure 4.9). The product formation patterns for the 28-mer and 29-mer can be simulated using the mechanism and parameters shown in Scheme 4.1. The series of conformational change events during the reaction can be derived using the fluorescent data acquired in the corresponding stopped-flow experiment as follows. At the beginning of the reaction,

binding of dCTP induces a very fast conformational change of the enzyme leading to a burst of 28-mer product formation. Although a small drop in fluorescence was observed in the stopped-flow experiment, most of the signal due to the dCTP-induced conformational change was too fast to be resolved. As the reaction proceeds, the rate of 28-mer formation declines and the 29-mer DNA starts to accumulate. At this stage, enzyme molecules are finishing the incorporation of dCTP and starting to return to its open DNA-bound state. The fluorescent intensity shows an increase during this phase. When the amount of 28-mer product starts to decrease, the opening of the enzyme after dCTP incorporation and the closing of the enzyme after dTTP binding reaches a steady state and the enzyme fluorescence remains constant for a short period. Afterward, a slow decrease in fluorescence indicates that more enzyme molecules are binding to dTTP as dCTP incorporation reaches completion, and dTTP incorporation becomes the major reaction in the solution. Although the discussion is only a qualitative analysis of how T7 DNA polymerase cycles through its different conformations during the two-nucleotide incorporation reaction, our data represents the first experimental evidence that allows the establishment of a relationship between polymerase conformations and its catalytic activities. Experiments to obtain kinetic parameters that govern these isomerization steps are currently underway.

Chapter 5: Utilizing the MDCC-E514C-8C Fluorescence to Detect Single Nucleotide Polymorphisms

5.1   INTRODUCTION

The MDCC-E514C-8C T7 DNA polymerase provides a fluorescent signal that is able to distinguish between correct and incorrect nucleotide binding. Our kinetic studies of the polymerase conformational changes using MDCC fluorescence also provide mechanistic information that helps us understand how T7 DNA polymerase achieves high fidelity DNA replication. Besides using the fluorescent signal to examine DNA polymerase mechanisms, we have developed a procedure to utilize this signal for detecting point mutations in DNA sequences.

The single nucleotide polymorphism (SNP) is a small genetic change that is found in an organism's genomic DNA sequence. The SNP variation occurs when a single nucleotide such as A is changed into one of the other three nucleotides – G, T, or C. Although it is estimated that 1 % of the human DNA sequence contains SNPs, most of the variations are located in the non-coding regions of the human genome and do not produce any physical changes. However, it is believed that some of these changes may make certain people more susceptible to some diseases (e.g. cancers) or affect their response to drugs. In order to associate certain SNP patterns with the susceptibility of disease or the effectiveness of drug treatments, large scale screening for all SNPs in the human genome is currently underway (88). The completion of the human genome project has provided the most critical information for this research, but developing a high throughput and cost-efficient method to search for SNPs is still needed (89). In this chapter, we present experiments using well-characterized SNP sequences from the human

F8 gene to demonstrate how point mutations could be detected using the fluorescent signal from the MDCC-E514C-8C T7 DNA polymerase.

The human factor VIII (*F8*) gene on the X-chromosome encodes a protein which is involved in the blood coagulation pathway. It has 26 exons spanning 186 kb of DNA and encodes a 9 kb mRNA. Hemophilia A is a bleeding disorder occurring in patients who carry deleterious mutations on their *F8* gene. This X-linked disease affects approximately 0.02 % of males (90). Causes of mutations on the *F8* gene include inversions, point mutations, deletions, and insertions (91).

5.2   MATERIALS AND METHODS

In order to investigate whether the signal from MDCC-E514C-8C T7 DNA polymerase could be used to detect point mutations, well-characterized point mutations at four positions of the *F8* gene were selected for synthesizing primer/template DNAs (Table 5.1). The primers were designed to have their 3' ends sit at a position located two bases before the positions that contain point mutations. The mutation detection assays began with a solution containing 50 nM of MDCC-E514C-8C enzyme, 100 nM of DNA substrates, and 100 nM of dideoxynucleoside triphosphate which correctly base-paired with the first template base, e.g. ddTTP for 712G and 712A DNA substrates. The solution was incubated at room temperature for at least 1 min to complete the incorporation reaction, and the emission profile of the protein-DNA complex was recorded from 440 to 540 nm with the excitation wavelength set at 425 nm. Next, 1 µM of dATP was added to the solution and the emission profile was recorded again. The procedure was repeated for each of the other three nucleotides (dTTP, dGTP, and dCTP) added to 1 µM at the last step.

117

Nucleotide	Synthetic DNA Substrate
712G	5' ACCTACTCATATCTTTCTCA 3' TGGATGAGTATAGAAAGAGTACACCTGGACCATTTTCTG 5'
712A	5' ACCTACTCATATCTTTCTCA 3' TGGATGAGTATAGAAAGAGTATACCTGGACCATTTTCTG 5'
1807C	5' GGGCCAACTAAATCAGATCC 3' CCCGGTTGATTTAGTCTAGGAGCCACGGACTGGGCGATA 5'
1807T	5' GGGCCAACTAAATCAGATCC 3' CCCGGTTGATTTAGTCTAGGAACCACGGACTGGGCGATA 5'
6049C	5' ATGGCTCAGGATCAAAGGAT 3' TACCGAGTCCTAGTTTCCTAAGCTACCATAGACGAGTCG 5'
6049T	5' ATGGCTCAGGATCAAAGGAT 3' TACCGAGTCCTAGTTTCCTAAACTACCATAGACGAGTCG 5'
6050G	5' ATGGCTCAGGATCAAAGGATT 3' TACCGAGTCCTAGTTTCCTAAGCTACCATAGACGAGTCG 5'
6050T	5' ATGGCTCAGGATCAAAGGATT 3' TACCGAGTCCTAGTTTCCTAAGATACCATAGACGAGTCG 5'
6050A	5' ATGGCTCAGGATCAAAGGATT 3' TACCGAGTCCTAGTTTCCTAAGTTACCATAGACGAGTCG 5'

Table 5.1:   DNA primer/template sequences used in mutation detection experiments

The sequences of the DNA substrates are derived from the *F8* gene sequence. The position of the template bases containing point mutations are marked in red.

118

5.3   RESULTS

When the 712G DNA substrate was used in the assay, the addition of 1 μM of dGTP resulted in the largest drop in fluorescence at around 460 nm. The result indicated that dGTP was correctly base-paired with the template base after the incorporation of ddTTP (Figure 5.1A). In the case of the 712A DNA substrate, dATP was the correct nucleotide and caused the largest decrease in signal instead of dGTP (Figure 5.1B). In order to present the data in a clear and concise manner, a relative fluorescence plot was created by dividing the final fluorescent intensity (second scan after the addition of dNTPs) by the initial intensity (first scan before the addition of dNTPs) at 448 nm wavelength (Figure 5.2). By plotting only the relative fluorescence data, the decrease in fluorescence upon correct nucleotide binding could be easily identified. Other than the 712G and 712A DNA substrates, seven other primer/template combinations listed in Table 5.1 were also tested in the same assay and yielded similar results.
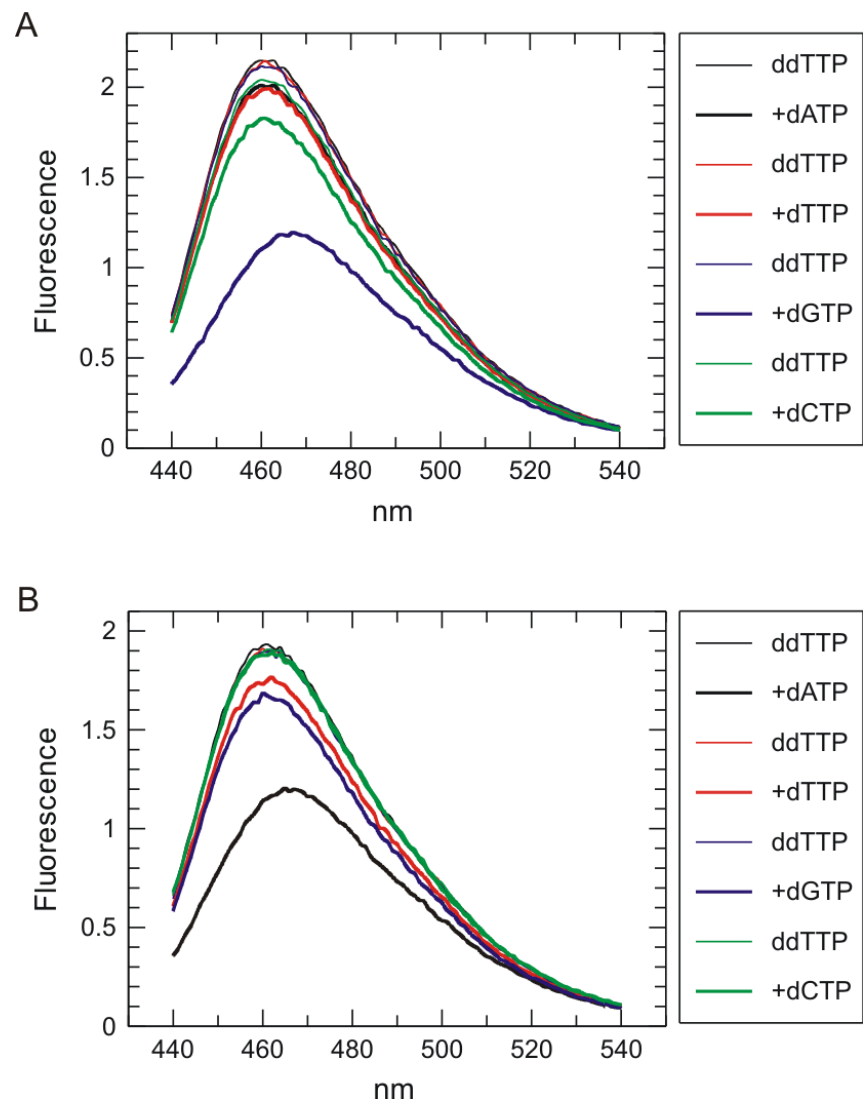
Figure 5.1:   Emission spectra of the MDCC-E514C-8C protein after binding of different nucleotides

Figure 5.1: Emission spectra of the MDCC-E514C-8C protein after binding of different nucleotides

The emission spectrum of the MDCC-labeled T7 DNA polymerase was recorded after the incorporation of ddTTP (thin traces) and after the addition of each nucleotide (thick traces). (A) Fluorescent emission scans showed that the addition of 1 μM dGTP caused the most significant decrease in fluorescence among experiments using the 712G DNA substrate. (B) Fluorescent emission scans from experiments using the 712A DNA substrate showed that dATP became the nucleotide that caused the largest decrease in enzyme fluorescence.
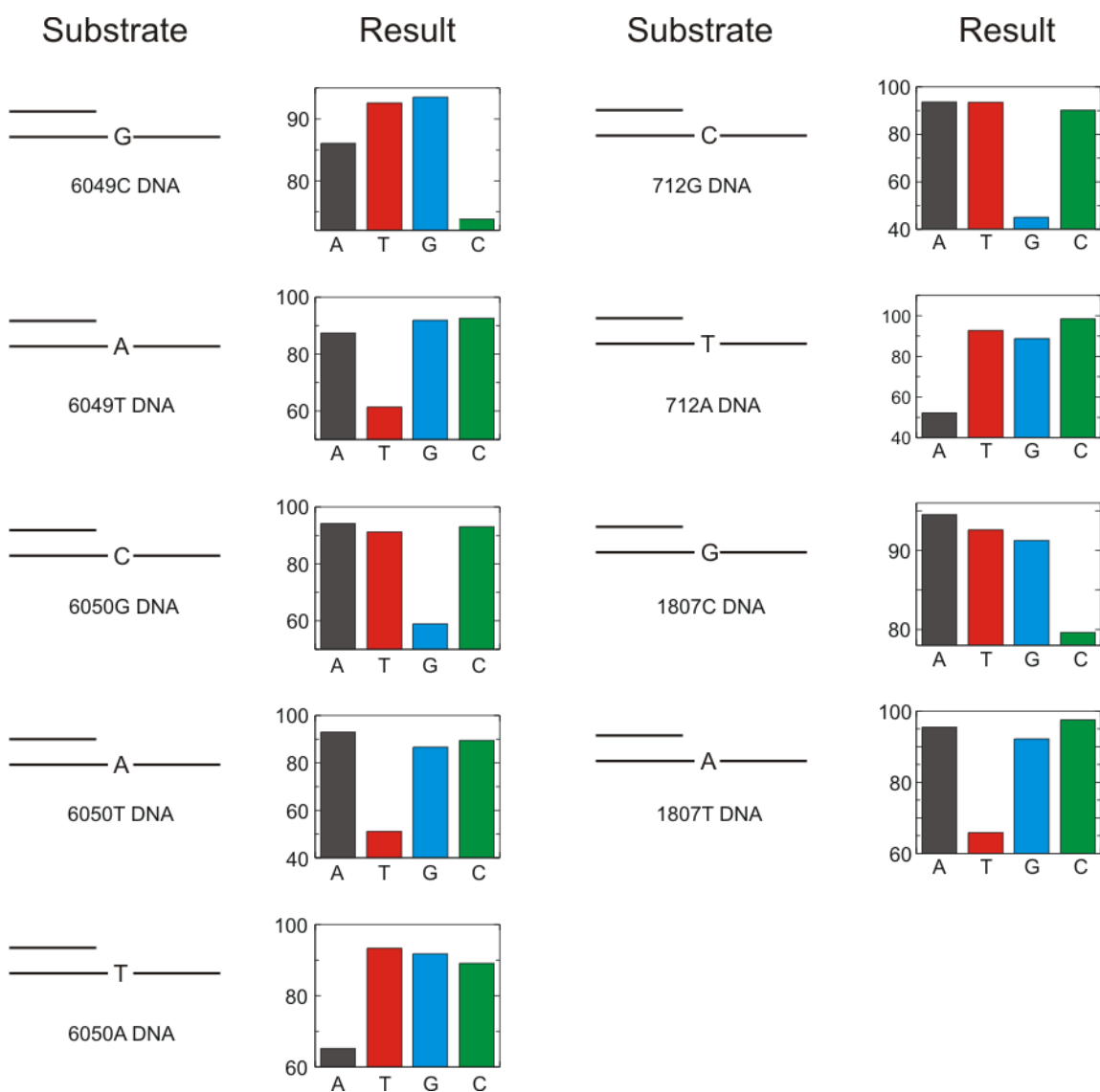
Figure 5.2:   Relative fluorescent changes at 448 nm calculated from the emission scan experiment

The plot of relative fluorescent changes in percentage simplifies the data interpretation in Figure 5.1. Signal from the MDCC-E514C-8C enzyme is able to distinguish every point mutation in DNA substrates listed in Table 5.1. Addition of the correct nucleotide results in the lowest relative fluorescence. Template bases being detected are shown in the substrate columns.

122

5.4 DISCUSSION

Experiments using seven synthetic DNA primer/template sequences have shown that the fluorescent intensity of the MDCC-E514C-8C enzyme is dependent on the identity of the bound nucleotide. When 1 μM of dNTP was added to the E-DNA solution with a ddNMP-terminated primer, a significant decrease in fluorescence around 460 nm was only observed if the added nucleotide was correctly base-paired with the template base. Other variables in these SNP detection experiments, such as different flanking sequences for the template base and different ddNMP at the 3' end of the primers, did not affect the fluorescence response to nucleotide binding. In conclusion, signals from the novel MDCC-E514C-8C construct may provide a tool not only to detect a point mutation in a DNA sequence but also to reveal the nature of that mutation without the need of sequencing analysis.

The experiments presented in this section also serve as control experiments for our kinetic studies using pre-defined 27/45-mer DNA substrate. Although the observed decrease in fluorescent intensity did not hold constant for each of the DNA substrate being tested, only the correct nucleotide binding caused a significant decrease in MDCC fluorescence. The experiment supports the conclusion that the MDCC signal is reporting subtle structural differences in the recognition domain following the conformational changes induced by the binding of nucleotides. Results from the mutation detection experiments also suggest that the selection mechanism is similar for all four deoxynucleoside triphosphates, and the kinetic mechanisms derived from our experiments are valid. Furthermore, the differences in fluorescent levels resulting from the binding of different nucleotides may be utilized in applications that require higher sensitivities than SNP detection. For example, if the relationship between the fluorescent

123

levels and the properties of the bound nucleotides during DNA polymerization can be established, the MDCC-E514C-8C signal may be used to carry out DNA sequencing analysis in real-time.

References

1. Delarue, M., Poch, O., Tordo, N., Moras, D., and Argos, P. (1990) *Protein Engineering* **3,** 461-467

2. Joyce, C. M. and Steitz, T. A. (1994) *Annual Review of Biochemistry* **63,** 777-822

3. Steitz, T. A. (1999) *Journal of Biological Chemistry* **274,** 17395-17398

4. Wang, J., Sattar, A. K. M. A., Wang, C. C., Karam, J. D., Konigsberg, W. H., and Steitz, T. A. (1997) *Cell* **89,** 1087-1099

5. Sawaya, M. R., Pelletier, H., Kumar, A., Wilson, S. H., and Kraut, J. (1994) *Science* **264,** 1930-1935

6. Brautigam, C. A. and Steitz, T. A. (1998) *Current Opinion in Structural Biology* **8,** 54-63

7. Steitz, T. A. (1993) *Current Opinion in Structural Biology* **3,** 31-38

8. Steitz, T. A., Smerdon, S. J., Jager, J., and Joyce, C. M. (1994) *Science* **266,** 2022-2025

9. Patel, S. S., Wong, I., and Johnson, K. A. (1991) *Biochemistry* **30,** 511-525

10. Donlin, M. J., Patel, S. S., and Johnson, K. A. (1991) *Biochemistry* **30,** 538-546

11. Loeb, L. A. and Kunkel, T. A. (1982) *Annual Review of Biochemistry* **51,** 429-457

12. Goodman, M. F. (1997) *Proceedings of the National Academy of Sciences of the United States of America* **94,** 10493-10495

13. Kornberg, A. and Baker, T. A. (1992) *DNA replication*, 2nd Ed., W. H. Freeman and Company, New York

14. Beckman, R. A. and Loeb, L. A. (1993) *Quarterly Reviews of Biophysics* **26,** 225-331

15. Doublie, S., Tabor, S., Long, A. M., Richardson, C. C., and Ellenberger, T. (1998) *Nature* **391,** 251-258

16. Pelletier, H., Sawaya, M. R., Kumar, A., Wilson, S. H., and Kraut, J. (1994) *Science* **264,** 1891-1903

17. Sawaya, M. R., Prasad, R., Wilson, S. H., Kraut, J., and Pelletier, H. (1997) *Biochemistry* **36,** 11205-11215

18. Huang, H. F., Chopra, R., Verdine, G. L., and Harrison, S. C. (1998) *Science* **282,** 1669-1675

19. Herschlag, D. and Jencks, W. P. (1989) *Journal of the American Chemical Society* **111,** 7587-7596

20. Herschlag, D. and Jencks, W. P. (1990) *Biochemistry* **29,** 5172-5179

21. Herschlag, D., Piccirilli, J. A., and Cech, T. R. (1991) *Biochemistry* **30,** 4844-4854

22. Wong, I., Patel, S. S., and Johnson, K. A. (1991) *Biochemistry* **30,** 526-537

23. Kati, W. M., Johnson, K. A., Jerva, L. F., and Anderson, K. S. (1992) *Journal of Biological Chemistry* **267,** 25988-25997

24. Johnson, K. A. (1993) *Annual Review of Biochemistry* **62,** 685-713

25. Doublie, S. and Ellenberger, T. (1998) *Current Opinion in Structural Biology* **8,** 704-712

26. Johnson, S. J. and Beese, L. S. (2004) *Cell* **116,** 803-816

27. Patel, P. H., Suzuki, M., Adman, E., Shinkai, A., and Loeb, L. A. (2001) *J.Mol.Biol.* **308,** 823-837

28. Showalter, A. K. and Tsai, M. D. (2002) *Biochemistry* **41,** 10571-10576

29. Fersht, A. (1998) *Structure and mechanism in protein science: a guide to enzyme catalysis and protein folding*, W. H. Freeman and Company, New York

30. Dunlap, C. A. and Tsai, M. D. (2002) *Biochemistry* **41,** 11226-11235

31. Arndt, J. W., Gong, W. M., Zhong, X. J., Showalter, A. K., Liu, J., Dunlap, C. A., Lin, Z., Paxson, C., Tsai, M. D., and Chan, M. K. (2001) *Biochemistry* **40,** 5368-5375

32. Purohit, V., Grindley, N. D. F., and Joyce, C. M. (2003) *Biochemistry* **42,** 10200-10211

33. Poynard, T., Ratziu, V., Benhamou, Y., Opolon, P., Cacoub, P., and Bedossa, P. (2000) *Baillieres Best Pract.Res.Clin.Gastroenterol.* **14,** 211-228

34. (1999) *J.Viral Hepat.* **6,** 35-47

35. Shimotohno, K. (2000) *Semin.Cancer Biol.* **10,** 233-240

36. Koff, R. S. (2003) *Int.J.Parasitol.* **33,** 517-523

37. Simmonds, P., Holmes, E. C., Cha, T. A., Chan, S. W., McOmish, F., Irvine, B., Beall, E., Yap, P. L., Kolberg, J., and Urdea, M. S. (1993) *J.Gen.Virol.* **74 (Pt 11),** 2391-2399

38. McHutchison, J. G., Gordon, S. C., Schiff, E. R., Shiffman, M. L., Lee, W. M., Rustgi, V. K., Goodman, Z. D., Ling, M. H., Cort, S., and Albrecht, J. K. (1998) *New England Journal of Medicine* **339,** 1485-1492

39. Poynard, T., Marcellin, P., Lee, S. S., Niederau, C., Minuk, G. S., Ideo, G., Bain, V., Heathcote, J., Zeuzem, S., Trepo, C., and Albrecht, J. (1998) *Lancet* **352,** 1426-1432

40. Ohba, K., Mizokami, M., Lau, J. Y., Orito, E., Ikeo, K., and Gojobori, T. (1996) *FEBS Lett.* **378,** 232-234

41. Choo, Q. L., Kuo, G., Weiner, A. J., Overby, L. R., Bradley, D. W., and Houghton, M. (1989) *Science* **244,** 359-362

42. Kato, N., Hijikata, M., Ootsuyama, Y., Nakagawa, M., Ohkoshi, S., Sugimura, T., and Shimotonho, K. (1990) *Proceedings of the National Academy of Sciences of the United States of America* **87,** 9524-9528

43. Takamizawa, A., Mori, C., Fuke, I., Manabe, S., Murakami, S., Fujita, J., Onishi, E., Andoh, T., Yoshida, I., and Okayama, H. (1991) *Journal of Virology* **65,** 1105-1113

44. Brown, E. A., Zhang, H. C., Ping, L. H., and Lemon, S. M. (1992) *Nucleic Acids Research* **20,** 5041-5045

45. Bukh, J., Purcell, R. H., and Miller, R. H. (1992) *Proceedings of the National Academy of Sciences of the United States of America* **89,** 4942-4946

46. Kolykhalov, A. A., Feinstone, S. M., and Rice, C. M. (1996) *Journal of Virology* **70,** 3363-3371

47. Tanaka, T., Kato, N., Cho, M. J., Sugiyama, K., and Shimotohno, K. (1996) *Journal of Virology* **70,** 3307-3312

48. Yamada, N., Tanihara, K., Takada, A., Yorihuzi, T., Tsutsumi, M., Shimomura, H., Tsuji, T., and Date, T. (1996) *Virology* **223,** 255-261

127

49. Lohmann, V., Korner, F., Herian, U., and Bartenschlager, R. (1997) *Journal of Virology* **71,** 8416-8428

50. Ferrari, E., Wright-Minogue, J., Fang, J. W. S., Baroudy, B. M., Lau, J. Y. N., and Hong, Z. (1999) *Journal of Virology* **73,** 1649-1654

51. Tomei, L., Vitale, R. L., Incitti, I., Serafini, S., Altamura, S., Vitelli, A., and De Francesco, R. (2000) *Journal of General Virology* **81,** 759-767

52. Yuan, Z. H., Kumar, U., Thomas, H. C., Wen, Y. M., and Monjardino, J. (1997) *Biochemical and Biophysical Research Communications* **232,** 231-235

53. Oh, J. W., Ito, T., and Lai, M. M. C. (1999) *Journal of Virology* **73,** 7694-7702

54. Luo, G. X., Hamatake, R. K., Mathis, D. M., Racela, J., Rigat, K. L., Lemm, J., and Colonno, R. J. (2000) *Journal of Virology* **74,** 851-863

55. Kao, C. C., Yang, X. Y., Kline, A., Wang, Q. M., Barket, D., and Heinz, B. A. (2000) *Journal of Virology* **74,** 11121-11128

56. Zhong, W. D., Uss, A. S., Ferrari, E., Lau, J. Y. N., and Hong, Z. (2000) *Journal of Virology* **74,** 2017-2022

57. Uchiyama, Y., Huang, Y., Kanamori, H., Uchida, M., Doi, T., Takamizawa, A., Hamakubo, T., and Kodama, T. (2002) *Hepatol.Res.* **23,** 90-97

58. Lesburg, C. A., Cable, M. B., Ferrari, E., Hong, Z., Mannarino, A. F., and Weber, P. C. (1999) *Nature Structural Biology* **6,** 937-943

59. Adachi, T., Ago, H., Habuka, N., Okuda, K., Komatsu, M., Ikeda, S., and Yatsunami, K. (2002) *Biochimica et Biophysica Acta-Proteins and Proteomics* **1601,** 38-48

60. Blight, K. J. and Rice, C. M. (1997) *J.Virol.* **71,** 7345-7352

61. Ishido, S., Fujita, T., and Hotta, H. (1998) *Biochem.Biophys.Res.Commun.* **244,** 35-40

62. Piccininni, S., Varaklioti, A., Nardelli, M., Dave, B., Raney, K. D., and McCarthy, J. E. (2002) *J.Biol.Chem.* **277,** 45670-45679

63. Gallinari, P., Brennan, D., Nardi, C., Brunetti, M., Tomei, L., Steinkuhler, C., and De Francesco, R. (1998) *Journal of Virology* **72,** 6758-6769

64. Back, S. H., Kim, J. E., Rho, J., Hahm, B., Lee, T. G., Kim, E. E., Cho, J. M., and Jang, S. K. (2000) *Protein Expr.Purif.* **20,** 196-206

128

65. Bartenschlager, R., Ahlborn-Laake, L., Mous, J., and Jacobsen, H. (1993) *J.Virol.* **67,** 3835-3844

66. Suzich, J. A., Tamura, J. K., Palmerhill, F., Warrener, P., Grakoui, A., Rice, C. M., Feinstone, S. M., and Collett, M. S. (1993) *Journal of Virology* **67,** 6152-6158

67. Howe, A. Y. M., Chase, R., Taremi, S. S., Risano, C., Beyer, B., Malcolm, B., and Lau, J. Y. N. (1999) *Protein Science* **8,** 1332-1341

68. Taremi, S. S., Beyer, B., Maher, M., Yao, N. H., Prosise, W., Weber, P. C., and Malcolm, B. A. (1998) *Protein Science* **7,** 2143-2149

69. Yao, N., Reichert, P., Taremi, S. S., Prosise, W. W., and Weber, P. C. (1999) *Structure Fold.Des* **7,** 1353-1363

70. Modrich, P. and Richardson, C. C. (1975) *Journal of Biological Chemistry* **250,** 5515-5522

71. Mark, D. F. and Richardson, C. C. (1976) *Proceedings of the National Academy of Sciences of the United States of America* **73,** 780-784

72. Singha, N. C., Vlamis-Gardikas, A., and Holmgren, A. (2003) *Journal of Biological Chemistry* **278,** 21421-21428

73. Tabor, S., Huber, H. E., and Richardson, C. C. (1987) *Journal of Biological Chemistry* **262,** 16212-16223

74. Huber, H. E., Tabor, S., and Richardson, C. C. (1987) *Journal of Biological Chemistry* **262,** 16224-16232

75. Lohmann, V., Roos, A., Korner, F., Koch, J. O., and Bartenschlager, R. (2000) *Journal of Viral Hepatitis* **7,** 167-174

76. Yanagi, M., St Claire, M., Shapiro, M., Emerson, S. U., Purcell, R. H., and Bukh, J. (1998) *Virology* **244,** 161-172

77. O'Farrell, D., Trowbridge, R., Rowlands, D., and Jager, J. (2003) *J.Mol.Biol.* **326,** 1025-1035

78. Sulewski, M., Marcheseragona, S. P., Johnson, K. A., and Benkovic, S. J. (1989) *Biochemistry* **28,** 5855-5864

79. Mohanty, B. K. and Kushner, S. R. (2000) *Proceedings of the National Academy of Sciences of the United States of America* **97,** 11966-11971

80. Reigadas, S., Ventura, M., Sarih-Cottin, L., Castroviejo, M., Litvak, S., and Astier-Gin, T. (2001) *Eur.J.Biochem.* **268,** 5857-5867

81. Sulewski, M., Marchese-Ragona, S. P., Johnson, K. A., and Benkovic, S. J. (1989) *Biochemistry* **28,** 5855-5864

82. De Lorimier, R. M., Smith, J. J., Dwyer, M. A., Looger, L. L., Sali, K. M., Paavola, C. D., Rizk, S. S., Sadigov, S., Conrad, D. W., Loew, L., and Hellinga, H. W. (2002) *Protein Science* **11,** 2655-2675

83. Polit, A., Blaszczyk, U., and Wasylewski, Z. (2003) *European Journal of Biochemistry* **270,** 1413-1423

84. Yengo, C. M., De la Cruz, E. M., Chrin, L. R., Gaffney, D. P., and Berger, C. L. (2002) *Journal of Biological Chemistry* **277,** 24114-24119

85. Brune, M., Hunter, J. L., Corrie, J. E. T., and Webb, M. R. (1994) *Biochemistry* **33,** 8262-8271

86. Hirshberg, M., Henrick, K., Haire, L. L., Vasisht, N., Brune, M., Corrie, J. E. T., and Webb, M. R. (1998) *Biochemistry* **37,** 10381-10385

87. Lunn, C. A., Kathju, S., Wallace, B. J., Kushner, S. R., and Pigiet, V. (1984) *Journal of Biological Chemistry* **259,** 469-474

88. Sherry, S. T., Ward, M. H., and Sirotkin, K. (1999) *Genome Research* **9,** 677-679

89. Twyman, R. M. (2004) *Current Topics in Medicinal Chemistry* **4,** 1423-1431

90. Hoyer, L. W. (1994) *New England Journal of Medicine* **330,** 38-47

91. Tuddenham, E. G. D., Schwaab, R., Seehafer, J., Millar, D. S., Gitschier, J., Higuchi, M., Bidichandani, S., Connor, J. M., Hoyer, L. W., Yoshioka, A., Peake, I. R., Olek, K., Kazazian, H. H., Lavergne, J. M., Giannelli, F., Antonarakis, S. E., and Cooper, D. N. (1994) *Nucleic Acids Research* **22,** 4850-4868

Vita

Yu-Chih Tsai was born in Taipei, Taiwan on October 17th, 1972, the son of Wen-Hsiung Tsai and Li-Hua Chen. After graduating from Cheng-Kuo High School in May 1992, Yu-Chih attended The National Taiwan University and received the degree of Bachelor of Science in May 1995. He served two years in Taiwan Army form July 1995 to June 1997 and worked as a research assistant in The National Taiwan University from July 1997 to July 1998. Yu-Chih then went to The University of Texas at Austin for graduate studies in August 1998.

Permanent address: 4F 25 LN 53 Chihfeng St., Taipei, Taiwan, 102

This dissertation was typed by the author.