2013

Nucleic Acid Binding Thermodynamics and Functions of DNA Polymerases

Hiromi Shimasaki Brown
Louisiana State University and Agricultural and Mechanical College

Follow this and additional works at: https://digitalcommons.lsu.edu/gradschool_dissertations

Recommended Citation
https://digitalcommons.lsu.edu/gradschool_dissertations/2650

This Dissertation is brought to you for free and open access by the Graduate School at LSU Digital Commons. It has been accepted for inclusion in LSU Doctoral Dissertations by an authorized graduate school editor of LSU Digital Commons. For more information, please contact gradetd@lsu.edu.
NUCLEIC ACID BINDING THERMODYNAMICS AND FUNCTIONS OF DNA POLYMERASES

A Dissertation

Submitted to the Graduate Faculty of the Louisiana State University and Agricultural and Medical College in partial fulfillment of the requirements for the degree of Doctor of Philosophy

in

The Department of Biological Sciences

by

Hiromi Shimasaki Brown
B.S., Louisiana State University, 2006
December 2013
# TABLE OF CONTENTS

ABSTRACT.......................................................................................................................... iv

CHAPTER 1. GENERAL INTRODUCTION .............................................................................. 1
  1.1 Discovery of DNA polymerase and reverse transcriptase ........................................ 1
  1.2 Studies of DNA polymerases .................................................................................... 2
  1.3 Cellular functions of DNA polymerase I ................................................................. 4
  1.4 Substrate preferences of human immunodeficiency virus-1 reverse transcriptase .... 13

CHAPTER 2. ENTHALPIC SWITCH-POINTS AND TEMPERATURE DEPENDENCIES
  OF DNA BINDING AND NUCLEOTIDE INCORPORATION BY POL I
  DNA POLYMERASES .................................................................................................... 22
  2.1 Abstract .................................................................................................................... 22
  2.2 Introduction .............................................................................................................. 23
  2.3 Materials and methods ........................................................................................... 24
  2.4 Results ..................................................................................................................... 27
  2.5 Discussion ............................................................................................................... 34
  2.6 Acknowledgements ................................................................................................ 39

CHAPTER 3. FACTORS INFLUENCING PARTITIONING BETWEEN
  POLYMERIZATION AND EDITING SITES OF KLENOW
  POLYMERASE ........................................................................................................... 40
  3.1 Abstract .................................................................................................................... 40
  3.2 Introduction .............................................................................................................. 41
  3.3 Materials and methods ........................................................................................... 44
  3.4 Results and discussions .......................................................................................... 46
  3.5 Concluding summary .............................................................................................. 62

CHAPTER 4. DIFFERENT MODES OF INTERACTION BETWEEN NUCLEIC ACIDS
  AND HIV-1 REVERSE TRANSCRIPTASE ARE DELINEATED BY THE
  SALT DEPENDENCE OF PRIMER/TEMPLATE BINDING ........................................... 64
  4.1 Abstract .................................................................................................................... 64
  4.2 Introduction .............................................................................................................. 64
  4.3 Materials and methods ........................................................................................... 66
  4.4 Results ..................................................................................................................... 71
  4.5 Discussion ............................................................................................................... 77
  4.6 Acknowledgements ................................................................................................ 84

CHAPTER 5. CONCLUSIONS ............................................................................................... 85
  5.1 Contribution of the negative enthalpy of DNA binding to the catalytic activity of
    Klenow and Klentaq polymerases .............................................................................. 85
  5.2 Networking of the polymerization and proofreading sites in Klenow for DNA
    replication fidelity .................................................................................................... 86
  5.3 Two different modes of nucleic acid binding of HIV-RT ......................................... 89
ABSTRACT

This study examines several different linkages between the nucleic acid binding thermodynamics and the functions of three different DNA polymerases. The focuses are correlation of the DNA binding thermodynamics and the functional behavior of Klenow and Klentaq polymerases (from *Escherichia coli* and *Thermus aquaticus*, respectively), identification of factors that influence the proofreading activity of Klenow, and examination of HIV reverse transcriptase (HIV-RT) binding to different primer-template nucleic acid constructs.

A comparison of the DNA binding thermodynamics and the incorporation activity of Klenow and Klentaq reveals that the enthalpic versus entropic balance upon binding may function as a modulator of the temperature dependence of the enzymatic activity. Both polymerases bind DNA with nanomolar affinity at significantly low temperatures, but have negligible enzymatic activity at these lower temperatures. For both polymerases it is found that the temperature of onset of significant enzymatic activity corresponds with the temperature where the enthalpy of binding crosses zero and becomes favorable (negative).

Proofreading activity improves the fidelity of DNA synthesis. Proofreading requires unwinding of the primer strand and shuttling of the 3’ terminus of the primer from the polymerization site to the proofreading site. The binding of Klenow to matched and mismatched primed-template DNA was examined by monitoring the steady state fluorescence intensity change of a 2-aminopurine base site-specifically substituted in DNA and reveals that both the equilibrium partitioning and the dynamic partitioning between sites are dependent on the absence, presence, and identity of specific divalent cations, as well as on the presence of mismatched bases at the primer/template junction.
HIV-RT performs both DNA and RNA template directed DNA synthesis. Direct binding equilibria have been characterized for the interaction of HIV-RT with several different primer/template nucleic acid constructs across a range of KCl concentrations. The thermodynamic affinities of the two homoduplexes (DNA/DNA and RNA/RNA) are shown to be nearly identical, while binding of the heteroduplexes is significantly tighter. At least two different modes of nucleic acid binding are revealed by the thermodynamic salt linkages of binding, and these different thermodynamic binding modes correlate with different recently structurally elucidated binding modes.
CHAPTER 1
GENERAL INTRODUCTION

1.1 Discovery of DNA polymerase and reverse transcriptase

The first accurate model of the DNA molecule was proposed by Watson and Click in 1953 (1). Their model describes the double helical nature of DNA, which is composed of two strands of DNA molecules with an invariant sugar-phosphate backbone on the outside, and paired nucleobases oriented toward the center of the molecule. The two DNA strands run in opposite directions to each other, and nucleobases from each strand base pair (A-T and C-G) via hydrogen bonds. A month after the publication of the model of the DNA structure, Watson and Click proposed semiconservative replication of DNA (2), predicting that the two strands would separate, and each strand would serve as a template for the formation of a new chain, resulting in two copies of the double stranded DNA molecule. Each new molecule will have one of the original chains and one new chain. In 1958, Meselson and Stahl demonstrated this semiconservative DNA replication mechanism in Escherichia coli (E. coli) by using high-speed cesium chloride (CsCl) density gradient ultracentrifugation to separate $^{15}$N isotope-labeled and non-labeled genomic DNA through the replication process (3). Meanwhile, Arthur Kornberg was seeking an enzyme that synthesized DNA. In 1956, he and his colleagues first demonstrated DNA synthesis in a cell free system (i.e., using a partially purified protein extract from E. coli) (4), and their subsequent studies lead them to the discovery of the first DNA polymerase: which is E. coli DNA polymerase I (4-6).

The “central dogma” of molecular biology that evolved in the 1950s and 1960s was that a cell's genetic information travels from DNA to RNA to proteins. In 1970, however, two individual groups of investigators, Baltimore and Temin and his colleagues, found an enzyme in RNA tumor viruses that synthesized DNA from an RNA template. (7, 8). It was first called
RNA-dependent DNA polymerase and later named “reverse transcriptase” in Rous sarcoma virus and in a mouse tumor virus (9). Their discovery contradicted the conventional wisdom at the time and resulted in an extension of the central dogma; from RNA to DNA to RNA to proteins.

1.2 Studies of DNA polymerases

Since the discovery of DNA polymerase I, the major approaches that have been used to study DNA polymerases are enzyme kinetics and crystallography. Early studies revealed that DNA polymerases can only add free nucleotides to a preexisting 3’-OH of a primer on a template DNA (4-6). Kinetic approaches revealed the mechanism/sequence of the nucleotide incorporation activity of the DNA polymerase (10-13). Crystallographic studies revealed the topology of the polymerase and its interaction with DNA (14, 15). Alternate series of biochemical and structural studies incorporating mutational analysis led to identification of the catalytic residues and the detailed mechanism of DNA polymerase activity (16-23).

A major focus in DNA polymerase research is the question of fidelity (24-27). DNA polymerases perform highly accurate replication utilizing two distinct mechanistic steps for error reduction. The first defense system for achieving high fidelity is discrimination against incorrect nucleotides, which takes place before the chemical step of nucleotide addition. Selection of the correct nucleotide from among a large pool of various nucleotides involves conformational changes of DNA polymerases (22, 23, 28). The dynamics of DNA polymerases has been mostly studied with fluorescence assays including Forster resonance energy transfer (FRET) and single-molecule FRET analyses (29-35). These studies have started revealing characteristics of the noncovalent steps that provide kinetic checkpoints for the overall specificity of the enzyme reactions. Proofreading activity of DNA polymerases takes place once an incorrect nucleotide is incorporated and enhances fidelity by as much as 300 fold, depending on the type of mismatches.
and the specificity of the DNA polymerase (25, 36). This is another extensively active area of study in DNA polymerase research. The mechanism of the proofreading activity has been studied with various approaches such as kinetic, thermodynamics, structural, and fluorescence analyses (22, 25, 26, 36-44). *E. coli* DNA polymerase I has been used as a model system for several decades.

DNA polymerases are subdivided into seven families presently based on sequence homology and structural similarities. These are: A, B, C, D, X, Y, and RT. Each family possesses different properties and plays various roles in DNA replication and/or DNA repair processes (45-51). Some DNA polymerases are classified as replicative polymerases, and others are classified as repair polymerases based on their primary function *in vivo*. In addition to the primary function of DNA polymerases, which is 5’-3’ polymerization activity, some also possesses 5’-3’ nuclease activity (family A), 3’-5’ exonuclease activity (families A, B, and D), and RNase H activity (family RT) (51). Retroviral reverse transcriptase (RT family) has three separate biochemical activities; RNA-dependent DNA polymerization, RNase H activity, and DNA-dependent DNA polymerization (52, 53). DNA polymerases that belong to the same family from different species may exhibit different roles in cells. For instance, family A polymerases are further divided into replicative and repair polymerases. T7 DNA polymerase from T7 bacteriophage and mitochondrial DNA polymerase γ are replicative, while *E. coli* DNA polymerase I and *Thermus aquaticus* (*T. aquaticus*) polymerase I from eubacteria are classified as repair polymerases.

Previous studies that took a closer look at DNA binding by *E. coli* DNA polymerase I and *T. aquaticus* polymerase I, which are the same type and belong to the same family, showed some similarities and differences in the binding reactions (54-59). The comparative studies of the thermodynamics of DNA binding by these polymerases in various solution conditions revealed
that the nature of the DNA-protein interactions is species specific but that some aspects of overall trends are similar. In this dissertation, I focus on these two polymerases to examine how the DNA binding of each polymerase is correlated with their catalytic functions. I also examine the exonuclease proofreading activity of *E. coli* DNA polymerase I in order to clarify some of the conflicting results raised in the previous studies (37-43). In addition, I examine the binding of various types of nucleic acids to human immunodeficiency-1 (HIV-1) reverse transcriptase at different salt concentrations, which yields detailed information about the physical properties of these DNA/polymerase interactions (54, 57, 58).

1.3 Cellular functions of DNA polymerase I

The DNA polymerase that was discovered by Arthur Kornberg is *E. coli* DNA polymerase I (4-6). *E. coli* DNA polymerase I is composed of a single polypeptide with a 103 kDa molecular weight (60). It possesses three distinct structure-function domains: the N-terminal 5’ nuclease domain, the C-terminal polymerase domain, and the intermediate 3’-5’ exonuclease/proofreading domain (17, 21). It is mostly involved in DNA repair and also in DNA replication (61-63). During prokaryotic DNA replication, DNA polymerase I removes RNA primers from the lagging strand with its 5’-3’ nuclease activity, and fills gaps with the 3’-5’ polymerase activity. DNA ligase then seals the nicks (63).

The major role of DNA polymerase I in the cell is involvement in DNA-repair pathways, such as base excision repair (64, 65) and nucleotide excision repair (66, 67). Base excision repair mainly repairs non-bulky lesions produced by alkylation, oxidation or deamination of bases. Nucleotide excision repair removes a variety of forms of DNA damage, including photoproducts induced by UV and other bulky lesions. In both types of excision repair, DNA polymerase I is involved in filling in short gaps.
E. coli Klenow polymerase and T. aquaticus Klentaq polymerase

Taq polymerase was first isolated in 1976 from a thermophilic eubacterium, T. aquaticus (68) which thrives at temperatures around 75 °C. In contrast, the optimal physiological growth temperature for E. coli is 37 °C (69). Taq polymerase is also a single polypeptide with a molecular weight of 94 kDa. Similar to E. coli polymerase I, Taq possesses three different structural domains: the N-terminal 5’ nuclease domain, the C-terminal polymerase domain, and the intermediate but nonfunctional proofreading domain. E. coli DNA polymerase and Taq polymerase share ~38 % overall sequence identity based on an amino acid sequence alignment (21). Removal of the N-terminal 5’ nuclease domains from these polymerases yield Klenow polymerase (68 kDa) and Klentaq polymerase (62 kDa) which retain the polymerase and proofreading domains (70). Although they have significant differences in thermostability (Tm = 40-60 °C for Klenow and Tm = ~100 °C for Klentaq) (71, 72), Klenow and Klentaq polymerases have highly similar architectures, and those are shown in Figure 1.1 (20, 73). Klenow and Klentaq share ~48 % sequence identity between their polymerase domains (21, 73) which adopt a shape that resembles a half-opened right hand with “fingers”, “thumb”, and “palm” subdomains.

DNA binding and polymerase activity

Both E. coli DNA polymerase I and Taq polymerase are DNA-dependent DNA polymerases, which attach a deoxyribonucleotide (dNTP) onto the 3’-OH of the DNA primer strand of a primer-template DNA (pt-DNA) guided by the DNA template strand (74). As shown in Figure 1.2, DNA polymerization is a multistep reaction starting with binding of the primer-template DNA (pt-DNA) to the polymerase domain of DNA polymerase. Once it binds to pt-DNA, the polymerase sequentially incorporates multiple nucleotides before dissociating from the
extended pt-DNA. Processivity is the average number of nucleotides added by a polymerase per association with the pt-DNA. During the processive incorporation process, the polymerase undergoes multiple nucleotide bindings, conformational changes, catalytic turnovers, pyrophosphate releases, and translocations. Then the extended product DNA is released (13, 28, 75, 76). *E. coli* DNA polymerase I has much lower processivity (< 30 bases) (11) than Taq polymerase does (50-80 bases) (77).

The catalytic step of nucleotide incorporation is mediated by two metal ions bound at the polymerase active site (78) and is shown schematically in Figure 1.3. Crystal structures of Klenow polymerase show the binding of Mg$^{2+}$ or Mn$^{2+}$ to Asp 705 and Asp 802 in the polymerase active site (17, 79). In addition to these carboxylate side chain residues, Glu 883, also with a carboxylate side chain, is also crucial for the catalytic activity (45, 80). An amino acid sequence alignment of Klenow and Taq polymerase showed that these residues are conserved as Asp 610, Asp785, and Glu 786 in Taq polymerase (21). In addition, these residues are well conserved among Type I DNA polymerases (45).

![Figure 1.1. Crystal structures of Klenow and Klentaq polymerases. The polymerase domain is shown in blue for Klenow and in red for Klentaq. The proofreading domain is shown in green for Klenow and in gray for Klentaq. The polymerase domain can be subdivided into three subdomains, thumb, fingers, and palm subdomains. The palm subdomain is indicated with a yellow arrow for each polymerase. The proofreading domain in Klentaq is inactive. The crystal structures of Klenow (1KRP) and Klentaq (1KTQ) are from references (20) and (73), respectively.](image-url)
Figure 1.2. An overview of DNA polymerases’ processive polymerization reaction. E is the enzyme, D_n is the DNA with an n nucleotide long primer. The first step of polymerization is DNA binding to the polymerase to form a binary complex (step 1). The following steps are nucleotide binding (step 2), a conformational change (step 3), catalysis (step 4), another conformational change (step 5), pyrophosphate release (step 6), and return to the binary complex (step 7a). This cycle repeats multiple times. The processivity (number of nucleotides addition per pt-DNA binding event) depends on the specific polymerase. After a round of processive polymerization, the product pt-DNA is released (step 7b) (13, 28, 75, 76).

Figure 1.3. Mechanism of the polymerase reaction. The roles of the two divalent metal ions are deprotonation of the 3’-OH of the primer strand and stabilization of the pentavalent transition state of the α-phosphate of the dNTP and for cleavage of the pyrophosphate. This figure is based on Figure 3 from reference (46) and was created using the program ChemBioDraw.
Each subdomain of the polymerase domain plays a distinct role in the polymerase reaction. Figure 1.4 shows a crystal structure of Klentaq polymerase binding to pt-DNA in the polymerase domain (19). The “thumb” subdomain binds the minor groove of the DNA duplex and places the 3’-OH of the primer near the conserved catalytic residues located in the “palm” subdomain, and the “fingers” subdomain interacts with the template stand and binds the incoming dNTP (19). The placement of dNTP in the nucleotide binding pocket, mediated by movement of the “fingers” subdomain, makes the tertiary complex into the activated tertiary complex as shown in step 3 of Figure 1.2.

Figure 1.4. Crystal structure of Klentaq bound to DNA showing the polymerization mode of DNA binding. The polymerase domain is divided into three subdomains: fingers (green), palm (magenta), and thumb (blue). The proofreading domain (inactive in Klentaq) is shown in gray. The bound DNA is shown in black. The polymerization catalytic residues, Asp 610, Asp 785, and Glu 786, are colored red and shown as spheres (19).

DNA binding thermodynamics and the overall enzymatic activity of Klenow and Klentaq polymerases

Primer-template DNA (pt-DNA) binding by DNA polymerases is the first step of the polymerization cycle. In Chapter 2 of this dissertation, I examine the relationship between the DNA binding thermodynamics and the overall enzymatic activity of Klenow and Klentaq.
polymerases. Like most sequence-specific DNA binding proteins, and many sequence-independent DNA binding proteins (81-84), both Klenow and Klentaq show significant heat capacity changes upon binding, yielding curved ΔG versus temperature dependences. This results in large changes in ΔH and TΔS with temperature, including a sign change from positive to negative for both enthalpy and entropy of DNA binding as temperature increases. For both polymerases, DNA binding is enthalpy-driven near their respective physiological temperatures (55, 56). Herein, I show a correlation between these common DNA binding thermodynamic patterns and the control of enzymatic activity in both polymerases. Both polymerases are enzymatically inactive until the temperature where the enthalpy of binding crosses zero (T_H) and becomes favorable. Because both polymerases bind DNA with nanomolar affinity at significantly lower temperatures (55, 56), these data indicate that a negative free energy of DNA binding alone is not sufficient to drive catalysis, and that a negative enthalpy of initial binding (ΔH) is required for nucleotide incorporation activity.

**Partitioning of DNA bindings between the pol and exo sites of Klenow**

During DNA synthesis, proofreading activity enhances replication fidelity by excising misincorporated nucleotides from 3’ end of the primer strand. The first step of the proofreading process is to send the misincorporated 3’-primer terminus to the 3’-5’ exonuclease (exo) site from the polymerase (pol) site, which requires separation of the primer terminus from the template strand (15, 17). A co-crystal X-ray structure of a binary complex composed of a duplex matched DNA and Klenow shows that the two active sites are located ~30 Å apart, and that during proofreading at least 3 bases of the 3’-primer terminus have been unwound from the template strand and bound to the 3’-5’ exonuclease site as a single strand (17). Figure 1.5 shows the editing mode of DNA binding with a matched DNA substrate. This is the only DNA bound
structural data that exists for Klenow. Comparison of the Klenow-DNA complex and the Klentaq-DNA complex (Figure 1.4) illustrates different fates of the 3’-terminus of the primer strand upon DNA binding in the polymerization versus exonuclease binding modes.

Figure 1.5. The editing mode of DNA binding. Klenow complexed with DNA shows the 3’ end of the primer strand bound at the 3’-5’ exonuclease active sites. DNA is shown in black, and the polymerase and proofreading domains are shown in blue and green, respectively. The polymerase catalytic residues, Asp 705, Asp 882, and Glu 883, are shown as red spheres, and the 3’-5’ exonuclease catalytic residues, Asp 355, Glu 357, Aps 424, and Asp 501, are shown as orange spheres (17).

The accuracy of passing parental genetic sequences to subsequent generations is highly dependent on faithful DNA replication by DNA polymerases. In *E. coli* this involves a balance between accurate nucleotide incorporation and exonucleolytic removal of incorrectly incorporated nucleotides. The 3’-5’ exonuclease activity also employs a two metal ion mechanism, and it is illustrated in Figure 1.6 (20). Metal ions A and B are bound to Asp 355, Glu 357, and Asp 501 at the catalytic site. In addition to these residues, Tyr 497 facilitates the attack of a hydroxide ion on the phosphorus, and Asp 424 helps stabilize metal ion B through two water molecules. Identities of these metal ions *in vivo* are unclear, but Mg$^{2+}$, Zn$^{2+}$, and Mn$^{2+}$ can all fill the two metal ion binding sites (16). Mutation D424A in Klenow abolishes metal ion B binding
but retains proper single-stranded DNA binding with a significantly reduced exonuclease activity (14), and D424A, also known as Klenow exo- (Klenow exo minus), is the most common Klenow molecule used for biochemical and structural studies.

Figure 1.6. Mechanism of the 3’-5’ exonuclease reaction. The two divalent metal ions facilitate formation of an attacking hydroxide, the departure of the 3’-O group, and the stabilization of the transition state. The water molecules are shown as black balls. This figure is adapted from reference (20) and was created using the program ChemBioDraw.

DNA polymerase I from *E. coli* has served as a model for the elucidation of the polymerization and the proofreading activity since its discovery. It still remains one of the most intensively studied polymerases. Equilibrium DNA binding at the pol and exo sites of D424A Klenow has been studied with matched and mismatched DNA for some time; however, establishing the initial partitioning of equilibrium DNA binding between the two sites remains controversial. Several studies have reported quite different levels of the partitioning of the primer strand of correctly matched pt-DNA at the exo site of D424A Klenow, ranging from nearly 0 %
up to 43 % (37-43). Most of the binding studies were done in the presence of Mg$^{2+}$ (37-40, 42, 43).

Determination of equilibrium binding partitioning at the pol and exo sites of D424A Klenow from measurements using the fluorescence depolarization properties of a dansyl fluorophore attached to pt-DNA by Millar and co-workers showed a 7-43 % occupancy at the exo site in the presence of Mg$^{2+}$ (37-39, 42). smFRET studies in the presence of Mg$^{2+}$ measuring real time polymerization dynamics with Cy3 fluorophore labeled pt-DNA and Cy5 labeled D424A Klenow showed 24 % occupancy at the exo site (40). Inexplicably, another smFRET study conducted by the same group using the same methodology showed that the primer strand of matched DNA has almost 0 % occupancy at the exo site in the presence of Mg$^{2+}$ (43). In all the above studies, the interpretation of the data is dependent on the assumption that D424A Klenow does not excise nucleotides from the 3’ end of the given pt-DNA in the presence of Mg$^{2+}$. However, the original characterization of Klenow exo- clearly notes that the exonuclease activity has significantly reduced but not totally eliminated. In addition, Datta et al. noted a slow residual exonuclease activity of D424A (over several hours) was present in buffers containing Mg$^{2+}$ during circular dichroism (CD) measurements of the binding of primers containing 2-aminopurine (2AP) dimers at the 3’ terminus of pt-DNA. 2AP is a fluorescent analog of adenine, which has been used in a variety of nucleic acid binding studies. In order to maintain the integrity of their DNA substrates during the CD measurements, Datta et al. replaced Mg$^{2+}$ with Ca$^{2+}$ since Ca$^{2+}$ is not an effective cofactor for the exonuclease or polymerase activity of D424A Klenow. Their results monitoring the local conformation of a 2AP dimer at the 3’ end of the primer showed about 43 % occupancy at the exo site for matched pt-DNA in buffers with Ca$^{2+}$ (41).
The major differences in the assay conditions (besides experimental approaches) among these studies are presence/absence of divalent ions, the type of divalent ions (Mg\textsuperscript{2+} vs. Ca\textsuperscript{2+}), the type and position of labeling on pt-DNA, the time periods of data acquisition, and the DNA sequences used.

In chapter 3 of this dissertation, I examine the binding of D424A Klenow to matched and mismatched pt-DNA by monitoring the steady state fluorescence intensity change of a single 2-aminopurine base site-specifically substituted in the template strand within the duplex part of pt-DNA, and also measure the residual exonuclease activity. The changes in fluorescence intensity allow detection of shuttling of the primer terminus between the pol and exo sites. The results showed that the rate of partitioning of the primer between the two active sites strongly depends on: 1) the number of mismatched bases at the primer-template junction, 2) the presence or absence of divalent ions, and 3) the type of divalent ions. I show that magnesium and calcium ions have opposite effects on the direction of the shift between the pol and exo sites. Substitution of the normal phosphodiester linkage between the last two bases of the primer strand with a non-hydrolysable phosphorothioate linkage also has significant effects on the partitioning between sites. In addition, the residual exonuclease activity of D424A Klenow also helps explain previous conflicting reports of the equilibrium partitioning between sites by different DNA substrates.

1.4 Substrate preferences of human immunodeficiency virus-1 reverse transcriptase

Human immunodeficiency virus-1 (HIV-1) is a lentivirus (slow replicating retrovirus) (85) that contains two copies of a single-stranded, positive sense single-stranded RNA genome (~10,000 base long) in each virion (52, 86). Fusion of HIV-1 to immune cells such as T cells and macrophages is mediated by primary interaction of the surface glycoprotein, gp120 of HIV-1 with CD 4 receptors on the surface of the host cells (87). As a result of the successful fusion to the host cell, HIV releases genomic RNA, reverse transcriptase, integrase, and other viral
proteins into the cytoplasm of the host cell where the process of reverse transcription occurs. The reverse transcriptase (RT) encoded by HIV-1 carries out a complex, multi-step reaction to convert the retroviral single-stranded RNA genome into double-stranded DNA, which is subsequently transported to the host cell nucleus and integrated into the host genome by viral integrase (52).

RT performs RNA- and DNA-dependent DNA polymerization and removes the RNA template of DNA/RNA hybrids through RNase H activity (52, 53). The full reverse transcription cycle of HIV-1 RT is shown in Figure 1.7. Reverse transcription initiates with minus strand DNA (cDNA) synthesis. A cellular tRNA$_{Lys}^3$ whose 3’ end partially unfolds, forms 18 basepairs of duplex with the viral RNA primer binding site (PBS) near the 5’ end of the viral RNA (88, 89). RT uses this RNA/RNA-primer/template to start incorporating dNTPs; then the substrate becomes a DNA/RNA-primer/template for DNA polymerization; and it continues to the 5’ terminus (step 1). The U5 and R (repeat) regions are copied into the minus strand DNA. This short (181 nucleotides long) cDNA is known as minus strong stop DNA (90). In order to continue minus strand DNA synthesis the virus employs a complicated mechanism, which involves transferring of the growing minus strand DNA to a remote position at the 3’ end of the viral RNA genome. This is called minus strand DNA transfer. The RNase H of RT cleaves RNA from the DNA/RNA hybrid produced by the RNA-dependent DNA polymerization activity of RT near the 5’ end of the genomic RNA (step 2), which helps induce the short segment of cDNA to translocate to the 3’ end of the genomic RNA (step 3). There are two repeating sequences, R, near the 5’ and 3’ ends of the viral RNA. Since both R (repeat) elements in the viral genome are identical, the minus single stranded DNA will interact with the 3’ end of the RNA genome through complementarity of their sequences (91-93). RNA-dependent DNA polymerization then
resumes completing minus strand DNA synthesis (step 4), accompanied by RNase H degradation of the viral RNA genome from the DNA/RNA hybrid product (step 5). Plus strand DNA synthesis initiates at an RNase H-resistant RNA sequence known as the polypurine tract (PPT) located near the U3 region at the 5’ end of the minus strand DNA. The PPT RNA serves as a primer, and the minus strand DNA serves as a template creating the RNA/DNA-primer/template for the initiation of DNA-dependent polymerization (step 6) (94, 95). The substrate then becomes a DNA/DNA-primer/template as polymerization proceeds. This polymerization process stops after copying the annealed portion of the tRNA\textsuperscript{Lys3} which is covalently attached to the minus strand DNA to generate the primer binding site (PBS) on the plus strand DNA. The tRNA\textsuperscript{Lys3} is then cleaved by RNase H activity of HIV-1 RT (step 7). This facilitates annealing of the PBS on the plus strand DNA to the primer binding site (PBS) complement on the minus strand DNA, which provides the complementarity for the plus strand DNA transfer (step 8) (91). DNA-dependent DNA synthesis continues to produce a linear duplex DNA with long terminal repeats (step 8). This process involves strand displacement synthesis as well (96).

**Structure of HIV-1 RT**

HIV-1 RT is a heterodimer consisting of two subunits: p66 (66-kDa) and p51 (51-kDa). The p51 subunit is derived from the p66 subunit by proteolytic cleavage by the viral protease (97, 98). The p66 subunit is composed of 560 amino acids, and the p51 subunit is composed of the first 450 amino acids of the p66 subunit (99). The two subunits have in common “fingers” (residues 1-84 and 120-150), “thumb” (244-322),” palm” (85-119 and 151-243), and “connection” (323-437) domains; however, p51 subunit lacks the C-terminal RNase H (438-556) domain (100). The arrangement of the p66 and p51 subdomains differs significantly in the heterodimer (Figure 1.8) (79, 100, 101).
Figure 1.7. Schematic representation of the reverse transcription process in HIV-1. The genomic RNA and tRNA are shown as a red bar and a pink toggle, respectively. Minus strand DNA and plus strand DNA are shown as a blue bar and a light blue bar, respectively. Step 1: Minus strand synthesis is initiated using a cellular tRNA annealed to the PBS and proceeds to the 5’ end of the RNA genome. Step 2: The RNA portion of the RNA/DNA hybrid is digested by the RNase H activity of RT. Step 3: First strand transfer event in which the newly synthesized DNA hybridizes with the R (repeat sequence) region at the 3’ end of the RNA genome. Step 4: Minus strand synthesis resumes. Step 5: A polypurine tract (PPT) withstands RNase H digestion. Step 6: Plus strand synthesis is initiated using the PPT as a primer. Step 7: The RNase H activity removes the PPT and tRNA. Step 8: Second strand transfer occurs by annealing of the PBS on the plus strand DNA to the PBS complement on the minus strand DNA. Step 9: DNA-directed DNA polymerization to complete the cDNA synthesis.
Figure 1.8. Domain structure of HIV-1 RT. The “fingers”, “palm”, “thumb”, “connection”, and RNase H domains of the p66 subunit are shown in green, magenta, blue, black, and red, respectively. The “fingers”, “palm”, “thumb”, and “connection” domains of the p51 subunit are shown in forest green, light-magenta, light-blue, and gray, respectively (101).

**Polymerase and RNase H activities of HIV-1 RT**

Both the DNA polymerase and RNase H domains reside in the p66 subunit, and a nucleic acid binding cleft runs between the two active sites (79, 100, 102-104). The polymerase active site, located in the “palm” subdomain of p66, catalyzes a nucleophilic attack by the 3’ hydroxyl of the primer terminus on the α-phosphate of dNTP, with release of pyrophosphate. The sterochemical outcome of the reaction is consistent with several other DNA polymerases, including Klenow, although the synthesis is either DNA or RNA templated for RT. RT also employs a two metal ion mechanism. The three catalytic aspartate residues (Asp 110, Asp 185, and Asp 186) chelate the two metal ions, and a third metal ion chelates the β and γ phosphates of dNTP and stabilizes the leaving of the pyrophosphate (45, 105). A possible mechanism for the polymerase reaction of HIV-1 RT is shown in Figure 1.9.
The RNase H activity (endonuclease) of RT hydrolyzes the RNA strand of RNA/DNA heteroduplexes to generate 5’ phosphate and 3’ hydroxyl ends, which requires two divalent ions in the RNase H active site. The four acidic residues, Glu 478, Asp 443, Asp 498, and Asp 549, are presumed to coordinate the binding of the two metal ions. The coordination of the metal ions stabilizes the substrate binding and promotes a nucleophilic attack by a hydroxyl ion on the scissile phosphate (106). A possible mechanism for the RNase H catalytic activity is shown in Figure 1.10.

Figure 1.9. Mechanism for the polymerase reaction of HIV-1 RT. The roles of the two divalent metal ions are deprotonation of the 3’-OH of the primer strand and stabilization of the pentavalent transition state of the α-phosphate of the dNTP. A third Mg$^{2+}$ is shown chelated by the β and γ phosphates of dNTP and stabilizes the leaving of the pyrophosphate. This figure is based on Figure 11 from reference (45) and was created using the program ChemBioDraw.
Figure 1.10. Mechanism for the RNase H activity of HIV-1 RT. The two divalent metal ions facilitate formation of an attacking hydroxide and the stabilization of the transition state. This figure is based on Figure 2 from reference (106) and was created using the program ChemBioDraw.

Polymerization and RNase H modes of nucleic acid binding of HIV-1 RT

Co-crystal structures of HIV-1 RT complexed with DNA/DNA (102, 107-109), PPT-sequence containing DNA/RNA (110), and non-PPT sequence containing DNA/RNA are now available (111). No structural data with RNA/RNA yet exists. The structures of HIV-1 RT complexed with a DNA/DNA substrate show that the 3' end of DNA is bound near the catalytic residues in the polymerase active site. The duplex near the active site adopts the A-form, a kink is adjacent to the p66 thumb, and B-form geometry persists near the RNase H active site (Figure 1.11) (102, 107-109). These DNA/DNA structures show nucleic acid binding in the polymerase active site. The 2013 structure of HIV-1 RT containing a non-nucleoside reverse transcriptase
inhibitor (NNRTI) and a non-PPT DNA/RNA differs from all previously reported HIV-1 RT-nucleic acid structures and is compatible with RNA cleavage. The bound hybrid DNA/RNA is mostly in the A-form and has an additional kink before entering the RNase H domain. The 3’ end of the DNA reaches into the palm but is ~5Å away from the polymerase active site (111). In addition, in this recent structure both p66 and p51 also showed conformational changes resulting in novel interactions with the hybrid and a more expanded architecture of p66/p51 (111). The regions that showed the changes are the p66/p51 connection domain, the p66 RNase H domain, and the p51 C terminus, which all contribute to positioning of the RNA strand of the hybrid in the RNase H domain (Figure 1.11) (111).

![Figure 1.11](image)

**Figure 1.11.** Polymerization and RNase H compatible nucleic acid binding modes of HIV-1 RT. The p66 and p51 subunits are shown in green and cyan, respectively. The DNA primer, DNA, template, and RNA template are colored red, pink, and blue, respectively. The polymerase catalytic residues (Asp 110, Asp 185, and Asp 186) are colored black, and the RNase H catalytic residues (Glu 478, Asp 443, Asp 498, and Asp 549) are colored brown. Catalytic residues in both structures are shown as spheres. PDB: 2HMI is from (107), and PDB: 4B3O is from (111).

The non-covalent driving forces that lead to stable protein-substrate complexes are strongly influenced by the solution environment (salt concentration and type, temperature, pH,
etc.). As a result of this dependence on the solution conditions, the functional properties (thermodynamics and kinetics) of these interactions must be investigated as a function of solution conditions to understand the origins of the stability of the complexes. Understanding which aspect of the solution composition leads to the stable complex formation helps understand what drives interactions of proteins and their substrates inside cells.

Although HIV-1 RT is one of the major drug-targeted viral proteins for treatment of HIV-1 infection (112, 113), direct comparative substrate binding studies among different pairs of nucleic acids are relatively scarce. In chapter 4 of this dissertation, I examine the affinities and the thermodynamic salt linkages for HIV-RT binding to four different 20/27mer (primer/template) nucleic acids; DNA/DNA, DNA/RNA, RNA/DNA, and RNA/RNA with equivalent sequences by using a fluorescence anisotropy based direct binding assay. The results showed that heteroduplexes bind to RT with approximately 25X higher affinity than do homoduplexes, and that RT-heteroduplex binding is linked to the release of approximately 2.5 more ions than is RT-homoduplex binding. Because the salt dependence of DNA binding is extremely sensitive to the DNA binding footprint, the thermodynamic linked ion releases upon binding indicate two distinct molecular binding modes for the HIV-1 RT to the different nucleic acid constructs. These thermodynamic results concur with the recent crystallographic finding of a second nucleic acid binding topology (RNase H compatible mode) by Yang and associates (111), although exact correspondence between the structurally and thermodynamically revealed binding modes cannot be established as yet. Interestingly, the affinity and salt linkage for HIV-1 RT binding to RNA/RNA and DNA/DNA constructs were nearly identical. Since little structural or binding information on HIV-1 RT interactions with RNA/RNA constructs exists, this data provides the first evidence that RNA/RNA and DNA/DNA likely bind to HIV-RT identically.
CHAPTER 2
ENTHALPIC SWITCH-POINTS AND TEMPERATURE DEPENDENCIES OF DNA BINDING AND NUCLEOTIDE INCORPORATION BY POL I DNA POLYMERASES

2.1 Abstract

This study examines the relationship between the DNA binding thermodynamics and the enzymatic activity of the Klenow and Klentaq Pol I DNA polymerases from Escherichia coli and Thermus aquaticus. Both polymerases bind DNA with nanomolar affinity at temperatures down to at least 5 °C, but have lower than 1% enzymatic activity at these lower temperatures. For both polymerases it is found that the temperature of onset of significant enzymatic activity corresponds with the temperature where the enthalpy of binding ($\Delta H_{\text{binding}}$) crosses zero ($T_H$) and becomes favorable (negative). This $T_H$/activity upshift temperature is 15 °C for Klenow and 30 °C for Klentaq. The results indicate that a negative free energy of DNA binding alone is not sufficient to proceed to catalysis, but that the enthalpic versus entropic balance of binding may be a modulator of the temperature dependence of enzymatic function. Analysis of the temperature dependence of the catalytic activity of Klentaq polymerase using expanded Eyring theory yields thermodynamic patterns for $\Delta G^\ddagger$, $\Delta H^\ddagger$, and $T\Delta S^\ddagger$ that are highly analogous to those commonly observed for direct DNA binding. Eyring analysis also finds a significant $\Delta C_p^\ddagger$ of formation of the activated complex, which in turn indicates that the temperature of maximal activity, after which incorporation rate slows with increasing temperature, will correspond with the temperature where the activation enthalpy ($\Delta H^\ddagger$) switches from positive to negative.

This chapter previously appeared as: Brown, H. S., and V. J. Licata. 2013. Enthalpic switch-points and temperature dependencies of DNA binding and nucleotide incorporation by Pol I DNA polymerases. Biochim Biophys Acta 1834:2133-2138. It is reprinted by permission of Elsevier.
2.2 Introduction

Primer-template DNA (pt-DNA) binding by a DNA polymerase is the first step of the DNA polymerization cycle. We have previously characterized the thermodynamics of pt-DNA binding with respect to salt and temperature for the Klenow and Klentaq “large fragments” of DNA polymerase I from *Escherichia coli* and *Thermus aquaticus* (55, 56). The present study asks the question of how these initial binding thermodynamics are related to or potentially influence the characteristics of the overall enzymatic cycle of the polymerases.

The Pol I DNA polymerases from *E. coli* and *T. aquaticus* are both single polypeptide chains comprised of three structure/function domains: a polymerization domain, a proofreading domain (inactive in Taq), and a 5′ nuclease domain. Removal of the 5′ nuclease domains yields the Klenow and Klentaq “large fragments” of each polymerase, both of which are fully functional polymerases on their own.

Previous studies of the temperature dependence of DNA binding of the two different polymerases established a variety of different thermodynamic characteristics for the two binding reactions (55, 56). Both binding reactions have been shown to display a classic curved Gibbs–Helmholtz plot (ΔG versus temperature), indicative of a significant heat capacity of binding and a strong temperature dependence of both the enthalpy and entropy of binding. Enthalpy–entropy compensation is observed for the binding of both polymerases, such that both the ΔH and the ΔS progress from positive values at low temperature to negative values at high temperature roughly in parallel. Enthalpy–entropy compensation for protein–DNA interactions and other biomolecular interactions has been recognized, and its meaning debated, for decades (e.g. (114-117)).
In the present study, nucleotide incorporation activity was measured with respect to temperature for both polymerases to characterize the kinetic behavior of the polymerases under a variety of solution conditions. What is somewhat more unusual, however, is that we examine possible connections between the thermodynamics of initial binding of the polymerases and their kinetic behavior. For both polymerases it is observed that the temperature where the ΔH of DNA binding switches from positive to negative (T_H) correlates with the temperature where nucleotide incorporation activity becomes significant (> 1%). The data suggest that the enthalpy–entropy compensation pattern displayed by polymerase binding may be related to the modulation of enzymatic activity. Eyring analysis of the temperature dependence of polymerization kinetics for Klentaq further suggests a heat capacity change upon formation of the activated complex resulting in a correspondence between the temperature where the activation enthalpy crosses zero and the temperature where the enzymatic activity reaches a maximum and begins to decrease. The findings suggest that temperatures where enthalpy (of binding or of activation) changes sign are possible regulatory points in the control of polymerase activity.

2.3 Materials and methods

Proteins and DNA

Purification of the proteins has been described previously (54). M13mp18(+) DNA was purchased from Amersham Pharmacia Biotech/GE Healthcare Life Sciences. E. coli competent strain JM109 from Promega was transformed with M13mp18(+) as described in the company's technical bulletin. The isolated single stranded M13 DNA was annealed with a 63-deoxyribonucleotide oligomer (P63):

5’CCATCCTAATTTACGAGCATGTAGAAACCAATCAATAATCGGCTGTCTTTCCTTACATTCCA3’ (from Integrated DNA Technologies, Inc.) which complements positions 3489–
3551 of the single stranded M13 DNA. Equimolar amounts of P63 and M13 DNA were annealed in TE buffer by heating at 95 °C for 5 min and gradually returning to room temperature. Use of this specifically primed-M13 nucleotide incorporation system, rather than using “activated” or randomly sheared calf-thymus or salmon sperm DNA as used in most polymerase nucleotide incorporation assays, means that the primer–template junction that the polymerases encounter in the nucleotide incorporation assay is the same as in the previous direct binding assays (55, 56).

Other materials

dATP, dTTP, dCTP, and dGTP were purchased from Promega. [α-32P]dATP (3000 Ci/mmol) was from PerkinElmer. DE81 filters were from Whatman, ScintiVerse BD Cocktail was from Fisher Scientific, and all other reagents were from Sigma Chemical or Fisher Scientific.

Nucleotide incorporation assay

The nucleotide incorporation reaction mixture was composed of 10 nM of annealed P63/M13mer, 200 μM each of dTTP, dCTP, and dGTP, 150 μM dATP, 100 μCi/mL [α-32P]dATP, and 1 μM DNA polymerase in a total volume of 20 μL. All reactions were initiated by addition of enzyme. The polymerases and the rest of the reaction mixture were separately equilibrated at reaction temperatures for 5 min before addition of enzyme. After timed incubations, reactions were quenched by adding 20 μL of 300 mM EDTA. Background measurements, and the zero timepoint, were obtained by replacing the enzyme with the corresponding buffer. Background measurements were insensitive to time of incubation. Each measurement was repeated three times. The incorporation of [α-32P]dATP was measured on DE81 filters as originally described by Bryant et al. (12) with small modifications in the washing steps. The procedure used herein was as follows: 10 μL of the quenched reaction mixture was
spotted on a DE81 filter. The filters were air-dried, then washed in a buffer containing 300 mM sodium phosphate, pH 7 for 5 min, with gentle swirling every 30 s. The filters were washed 3 times in the washing-buffer followed by a wash with 70% ethanol. 5 mL of washing-buffer and 2 mL of 70% ethanol per individual DE81 filter disk were used for each wash. The filters were then air-dried again, then immersed in 4 mL of scintillation cocktail. Radioactivity present in DNA was counted with a Packard BioScience Tri-Carb 2900TR Liquid Scintillation Analyzer. In order to determine the amount of dNMP incorporated by the polymerases, background and total radioactivity measurements were carried out. In the total radioactivity measurements, the washing steps were omitted. The total number of nucleotides (nts) incorporated per 10 nM P63/M13mer was calculated with the following equation: \[ \text{nts} = \frac{\text{sample count (cpm) − background count (cpm)}}{\text{total count (cpm)}} \times \frac{150 \mu M \text{ dATP} \times 3.07/10 \text{ nM (concentration of P63/M13mer)}}{150 \mu M \text{ dATP} \times 3.07/10 \text{ nM (concentration of P63/M13mer)}} \]. The value of 3.07 was calculated using a thymine content for single stranded M13mp18(+) of 32.6%.

**Temperature dependence of polymerase activity**

The buffers used for temperature dependent nucleotide incorporation assays were the same as those previously used for the temperature dependent primer–template DNA binding assays (55, 56), and were composed of 10 mM Tris and 5 mM MgCl₂ with 300 mM KCl for Klenow and 75 mM KCl for Klentaq. The pH was adjusted to pH 7.9 at the temperature for each experiment. The assayed temperatures ranged from 5 °C to 70 °C for Klentaq, and from 5 °C to 45 °C for Klenow. The assayed incubation times were from 0 to 10 min for both polymerases. Timepoint data were collected in triplicate, and steady state rates (k_{ss}) of incorporation were obtained by linear regression analysis using the program Kaleidagraph (Synergy Software).
Other buffer systems

Temperature dependence of Klentaq polymerase activity was also measured in a variety of different 10 mM buffer solutions using a 2 minute incubation.

Eyring analysis

The steady state rates of nucleotide incorporation for Klentaq from 25 °C to 70 °C were analyzed using an Eyring plot (ln k/T vs. 1/T), and fit to a non-linear form of the Eyring equation (118):

\[ \ln \frac{k_{ss}}{T} = -\frac{\Delta H^\ddagger}{R} \times \frac{1}{T} + \left( \ln \frac{k_B}{h} + \frac{\Delta S^\ddagger}{R} \right) \]

where:

\[ \Delta H^\ddagger(\text{at } T) = \Delta H_r^\ddagger + \Delta C_p^\ddagger \times (T - T_r), \]

\[ \Delta S^\ddagger(\text{at } T) = \Delta S_r^\ddagger + \Delta C_p^\ddagger \times \ln(T/T_r), \]

\( k_B/h \) is Boltzman's constant/Planck's constant.

The subscript ‘r’ denotes a single selected reference temperature in Kelvin while T is any other measured temperature. The activation enthalpy (\( \Delta H^\ddagger \)) and activation entropy (\( \Delta S^\ddagger \)) at any assayed temperature can then be extracted from the curve fit, and the free energy of activation (\( \Delta G^\ddagger \)) at each assayed temperature can be calculated (\( \Delta G^\ddagger = \Delta H^\ddagger - T \Delta S^\ddagger \)) or obtained directly from \( k_{ss} = k_B T/h \times \exp(-\Delta G^\ddagger / RT) \). Non-linear analysis was done with the program Kaleidagraph (Synergy Software).

2.4 Results

We have examined the temperature dependence of nucleotide incorporation for Klenow and Klentaq polymerases, with the purpose of comparing these results with their previously reported temperature dependences of initial primer–template (pt-DNA) binding (55, 56). The substrate for nucleotide incorporation was single-stranded M13 DNA primed with the same
single-stranded 63mer used in binding studies, such that the sequence at the primer–template junction is the same in both the binding and incorporation assays.

Nucleotide incorporation activity was followed using [α-32P]dATP, and steady state kinetic data are shown in Figure 2.1. Polymerase nucleotide incorporation rates are often determined from a single timepoint: a strategy that will yield equivalent results if the reactions all remain in the initial linear kinetic phase at that timepoint under all conditions. Polymerase incorporation kinetics often switch to a second, slower rate at times longer than approximately 5–10 min, depending on temperature ((11, 119-121), and data not shown). Due to the wide range of temperatures examined in this study, we collected data as a function of time to ensure that data compared across temperatures are all in the initial steady state kinetic phase. The kinetics are linear rather than exponential in these experiments because a high enzyme:DNA ratio is used, making the reactions pseudo-zero order. A linear plot of [product] or [reactant] versus time is a classic diagnostic for zero order kinetics (122).

![Figure 2.1](image.png)

Figure 2.1. Nucleotide incorporation kinetics of Klenow and Klentaq as a function of temperature. Panel A shows nucleotide incorporation for Klenow from 5 to 45 °C. Panel B shows incorporation activity of Klentaq from 5 to 60 °C. The data are fit to a linear equation. Each data point is an average of three measurements, and the error bars show the standard deviation of the triplicates. For both polymerases, incorporation rates increase with the rise in temperature up to 45 °C for Klenow and 60 °C for Klentaq. After 60 °C, the incorporation rate for Klentaq decreases with increasing temperature. Data were not collected above 45 °C for Klenow due to known onset of protein denaturation near this temperature.
The nucleotide incorporation rates of Klenow and Klentaq were determined at five-degree intervals as shown in Figure 2.1. For both polymerases, the rate of nucleotide incorporation is negligible at lower temperatures, then appears to switch “on” and increase with temperature. Data for Klentaq additionally show a maximal incorporation rate at 60 °C under these conditions, after which the incorporation rate declines. Klentaq does not begin to denature until above 95 °C (71), and the 63mer used to prime the DNA does not begin melting until above 70 °C under these buffer conditions (data not shown), thus the decreased nucleotide incorporation rate above 60 °C is not due to either protein unfolding or DNA denaturation. Klenow data were not collected past 45 °C, as the protein is known to begin to denature above this temperature (71, 72).

Figure 2.2. Incorporation rates vs. temperature for Klenow and Klentaq polymerases. Rates are shown in nucleotides (nts) incorporated per second. Rates were obtained from the linear regression analyses of the data in Figure 1, and error bars show the errors of the fits from Figure 2.1. Incorporation activity rises above 1% of maximum at 15 °C for Klenow (open squares) and at 30 °C for Klentaq (open circles). For Klentaq, the activity reaches a maximum at 60 °C and then decreases with further increase in temperature.

Previous binding studies have provided Gibbs–Helmholtz plots (ΔG versus T) for pt-DNA binding of the two polymerases (55, 56), and in Figure 2.3 the ΔG, ΔH, and TΔS data from those studies are shown superimposed with the temperature dependent nucleotide incorporation
rate data from Figure 2.2. These earlier binding studies showed that the $\Delta G$ of binding versus temperature for both polymerases is non-linear, with a minimum (tightest binding) near 47 °C for Klentaq and near 33 °C for Klenow (dashed lines in each panel of Figure 2.3). A non-linear $\Delta G$ versus temperature indicates that the binding reaction proceeds with a significant change in heat capacity, and this is reflected in the significant temperature dependencies of both $\Delta H$ (solid diagonal lines in both panels) and $T\Delta S$ (dotted diagonal lines). For both polymerases, enthalpy and entropy both proceed from positive values at lower temperatures (where binding will be entropy driven) to negative values at higher temperatures (where binding will be enthalpy driven). The temperature points where enthalpy and entropy cross zero and change signs are denoted $T_H$ and $T_S$, respectively.

Figure 2.3. Combined plots of polymerase–DNA binding thermodynamics and normalized polymerase activity of Klenow and Klentaq from Figure 2.2. The thermodynamic binding data are from Refs. (55) and (56). Nucleotide incorporation rate data are shown as closed squares for Klenow (panel A) and closed circles for Klentaq (panel B). The previous binding studies provided $\Delta G$, $\Delta H$, and $T\Delta S$, of binding, which are shown as dashed ($\Delta G$), solid diagonal ($\Delta H$), and dotted diagonal ($T\Delta S$) lines in both panels.
It can be seen in Figure 2.3 that the apparent onset of the enzymatic activity corresponds to where the enthalpy of DNA binding switches signs (and favorability). In other words, nucleotide incorporation is almost negligible (< 1% of maximum) until the temperature where \( \Delta H \) switches from positive to negative (\( T_H \)). This temperature is 15 °C for Klenow and 30 °C for Klentaq. Note that the enzymes are never really predicted to completely “switch off” — they should continue to decrease exponentially in activity as the temperature decreases, but from a physiological point of view, in most situations, < 1% activity is effectively “off”.

Nucleotide incorporation activity is sensitive to a variety of solution conditions, including but not limited to pH, KCl concentration, MgCl\(_2\) concentration, and the nature of the DNA substrate (123, 124). Figure 2.4 shows the temperature dependence of nucleotide incorporation under a variety of different solution conditions. It can be seen that different solution conditions can shift the temperature for maximal activity of the polymerase by up to about 10 °C. The incorporation rates are normalized in Figure 2.4 to emphasize transitions in the activity (onset of significant activity and temperature of maximum activity) — the absolute rates at the temperature maxima vary significantly among the different conditions. The normalized data show that changing solution conditions alters the shape and steepness of the temperature dependence for activity. Interestingly, however, the temperature where enzymatic activity rises above 1% of maximum is nearly identical under all of the conditions examined, meaning that the correspondence of \( T_H \) and rise in activity shown for the carefully matched buffer conditions in Figure 2.3, would still hold for the wide variety of buffer conditions examined in Figure 2.4. It is also notable that for all conditions the rise in activity tapers off or even reverses at higher temperatures. This non-Arrhenius behavior has several possible origins (see Discussion).
Figure 2.4. Temperature dependence of the nucleotide incorporation activity of Klentaq in various buffer systems. Incorporation activities in this plot were measured at 2-min of incubation and then normalized. The different buffer systems examined are as follows: 1) X symbols: 10 mM succinate, 75 mM NaCl, 5 mM MgCl₂, pH 6; 2) open squares: 10 mM PIPES, 75 mM KCl, 5 mM MgCl₂, pH 7; 3) open triangles: 10 mM phosphate, 75 mM KCl, 5 mM MgCl₂, pH 7.9; 4) open circles: 10 mM Tris, 75 mM KGlutamate, 5 mM MgCl₂, pH 7.9; 5) open diamonds: 10 mM Tris, 75 mM KCl, 5 mM MgCl₂, pH 9; 6) plus symbols: 10 mM CAPS, 75 mM KCl, 5 mM MgCl₂, pH10; and 7) closed triangles: 25 mM Tris, 25 mM KCl, 4.5 mM MgCl₂, pH 8.3 (buffer condition # 7 is from Ref. (11), others are designed to explore a range of pH’s and to examine conditions analogous to several commercial PCR buffer conditions). The normalized incorporation rate data from Figure 2.2 is shown as the closed diamonds.

Eyring analysis of the Klentaq kinetic data from Figure 2.2 is shown in Figure 2.5. Because the activity rate goes through a maximum after which it begins to slow with increasing temperature, the data are non-linear in both Eyring and Arrhenius plots. Eyring analysis can be expanded (by analogy to non-linear Gibbs–Helmholtz plots) to account for such non-linear temperature dependence (118). Because the kinetic data (in Figure 2.1) are pseudo-zero order, the rate = k_{ss} (the rate constant) (122), for use in the Eyring analysis. Fersht and associates have previously used such expanded Eyring analyses to explain how a protein folding reaction can show a rate maximum, when Arrhenius theory predicts that the folding rate should increase continuously with temperature without going through a maximum (118). Non-linear Eyring
analysis provides quasi-thermodynamic information about the activated complex of a kinetic reaction: including the $\Delta G^\ddagger$, $\Delta H^\ddagger$, $\Delta S^\ddagger$, and the $\Delta C_p^\ddagger$ of activation (118, 125). Figure 2.5 shows that an analogous analysis can be applied to examine the enzymatic activity maximum that is observed for Klentaq at 60 °C. The fitted $\Delta C_p^\ddagger$ of activation for formation of the Klentaq-DNA-dNTP transition state complex is $-1.71 \pm 0.15$ kcal/molK, which is approximately 2 × larger than the $\Delta C_p$ of DNA binding itself (55). Figure 2.5 shows the Eyring plot for the data of Figure 2.2, used to obtain the $\Delta C_p^\ddagger$, along with the temperature dependencies of $\Delta G^\ddagger$, $\Delta H^\ddagger$ and $T \Delta S^\ddagger$ calculated as described in Materials and methods. Klenow does not exhibit such an activity maximum, instead its activity increases with temperature until the protein denatures. The 60 °C maximal activity temperature for Klentaq in Figure 2.2 occurs long before either the protein or the DNA begins denaturing under those solution conditions (71).
While most chemical reactions normally exhibit a positive enthalpy of activation ($\Delta H^\ddagger$), when the reaction also has a $\Delta C_p^\ddagger$, the $\Delta H^\ddagger$ will change with temperature, and at higher temperatures will cross zero and become negative, as shown in Figure 2.5. It is at this $T_H^\ddagger$ temperature that the enzyme activity begins to decline with increasing temperature, as dictated by the Gibbs–Helmholtz based extension of the Eyring equation (see Materials and methods). Thus, not only is the $T_H$ for the enthalpy of binding serving as an onset point for enzymatic activity, but also the quasi-thermodynamic $\Delta C_p^\ddagger$ for the formation of the activated complex means that there will be a $T_H^\ddagger$ for the enthalpy of activation corresponding to the reversal of the temperature dependence of nucleotide incorporation.

2.5 Discussion

The purpose of this study was both to characterize the temperature dependence of the enzymatic activities of Klenow and Klentaq DNA polymerases, and to begin to ask the question of how the thermodynamics of DNA binding might potentially influence the overall activity of Klenow and Klentaq DNA polymerases. Typically, binding thermodynamics and enzymatic activity are examined independently for any enzyme, usually by different researchers, rather than being studied in parallel to ask how they might be energetically related. Herein we find that the temperature dependent shift in the enthalpy–entropy balance of initial DNA binding corresponds with specific changes in the overall functional behavior of the polymerases.

Both polymerases are tightly bound to DNA ($\Delta G$ tighter than $-9 \text{ kcal/mole}$) at temperatures well below where their enzymatic activity initiates. Simple DNA binding with a negative $\Delta G$ is not sufficient for polymerase activity. Polymerase activity does not correlate with any trend in $\Delta G$ of binding, instead, the specific combination of enthalpy and entropy that produce the $\Delta G$ of binding appears to correlate with the catalytic activity of the polymerases.
Nucleotide incorporation activity does not rise above 1% until the binding enthalpy begins to turn negative, and physiologically relevant enzymatic activity increases as the favorable enthalpy gets larger. In the earlier thermodynamic binding studies of these polymerases we questioned whether there might be such a correspondence (56), and herein have been able to obtain appropriate enzymatic data to establish that there is.

It is possible that these matched $T_H$/activity onset temperatures are just fortuitous, but the fact that the corresponding temperatures for Klentaq and Klenow differ by 15 °C argues against a simple coincidence, and also indicates that polymerase catalysis does not simply initiate at some universal start temperature, but actually does track individually for the two different enzymes.

The characteristic enthalpy–entropy compensation pattern for protein–DNA interactions (as illustrated by the parallel $\Delta H$ and $T\Delta S$ dependencies in Figure 2.3), which is also seen for other biomolecular interactions with associated heat capacity changes, has been observed and its meaning debated for many years (e.g. (114-117)). One widely recognized consequence of enthalpy–entropy compensation is that the $\Delta G$ of binding remains relatively constant versus temperature (as also illustrated in Figure 2.3). This study seems to be the first to suggest a new possible physiological utility of such a thermodynamic pattern: if an enzyme is effectively inactive (from a physiological point of view) when the enthalpy of substrate binding is positive, and active when the enthalpy of substrate binding is negative, then the thermodynamics of binding will serve as a sensitive means for regulating the activity of that enzyme.

This is not the first time that the specific $\Delta H$ versus $T\Delta S$ makeup of a $\Delta G$ has been proposed to be critical for manifesting a specific molecular/physiological function. Recent examples include the specific use of enthalpy, rather than simply $\Delta G$, to drive heterotropic regulation in the biotin repressor (126), and the finding that the specific enthalpy–entropy ratio
for similar or identical binding free energies can dramatically alter the efficacy of HIV drugs (e.g. (127-129)).

The data in this study also question the longstanding concept of “corresponding states” of thermophilic and mesophilic proteins, which posits that the function of thermophilic proteins is inhibited at lower temperatures because they are too rigid (130, 131). The Klenow and Klentaq activity data in this study suggest that the situation is more complicated, as both proteins bind DNA quite tightly at low temperatures (i.e., there is no temperature dependent inhibition of the binding function of the thermophile at lower temperatures). Instead, the binding profiles are very similar for the two proteins versus temperature (55, 56), while the up-shift temperature for enzymatic activity in the thermophile appears more closely correlated with the binding enthalpy rather than any characteristic of the protein's thermal stability (71, 72). It is, of course, possible that the DNA binding enthalpy (rather than binding free energy) is somehow correlated with protein rigidity, but the data of this study do not address this possibility.

Regarding the temperature maximum in the nucleotide incorporation rate, standard Arrhenius/Eyring theory holds that a reaction rate will always increase with increasing temperature. The rate maximum (and subsequent decrease in rate with further increase in temperature) exhibited by Taq polymerase is thus somewhat unusual, but is one of a slowly growing number of examples of such behavior in biochemistry. There have been a few different approaches to the analysis and interpretation of such non-linear Arrhenius/Eyring behavior. Winzor and Jackson caution that this can often be a result of a change in reaction mechanism or involvement of a new thermodynamic linkage at high temperatures (such as temperature dependent linkages to ion or proton binding) (132). Such potential unforeseen linkages are almost impossible to completely rule out. In another approach, Daniel and Danson (133, 134)
have postulated that such effects can occur if there is a pre-existing equilibrium between two different protein conformers. If this pre-existing equilibrium shifts with temperature, it can produce an effective reversal of enzyme behavior under conditions where one would normally expect continued acceleration of the reaction (e.g. increased temperature, increased [substrate], and such). In yet another approach to interpreting such a temperature maximum, Alan Fersht analyzed such Eyring plot curvature in a protein folding reaction by introducing a heat capacity of activation ($\Delta C_p^\dagger$) (118), by analogy to analysis of non-linear van't Hoff or Gibbs–Helmholtz plots. Application of such an approach to the analysis of the temperature maximum for Klentaq activity suggests that there exists a significant $\Delta C_p^\dagger$ of activated complex formation ($2 \times$ that of direct binding (55)), suggesting large changes in structure or hydration in the activated complex relative to the DNA–protein binary complex, and resulting in another enthalpic correlation: as shown in Figure 2.5, the temperature where the activation enthalpy crosses zero ($T_H^\dagger$) correlates with the maximal activity temperature. Because there are several other potential explanations for non-linear Eyring behavior (e.g. (132-134)), this $\Delta C_p^\dagger$ induced correlation between $T_H^\dagger$ and the activity maximum correlation might normally be discounted, but when considered alongside the binding enthalpy correlation ($T_H$ and activity onset) it suggests another intriguing enthalpy based control point for enzymatic activity.

Biomacromolecules display a number of strategies for creating effective “on/off switches”. Examples include homotropic and heterotropic allostery, two-state protein unfolding, and covalent modifications such as phosphorylation. These data give a glimpse into another potential type of molecular on/off switch that takes advantage of the fact that many biomolecular interactions (i.e. all those with heat capacity changes) display enthalpy–entropy compensation with large excursions of $\Delta H$ and $T\Delta S$. The physiological utility of such a thermodynamic based
“on/off” switch in the case of the DNA polymerases examined here is that both polymerases will stay bound to DNA (with a strongly negative ΔG) even at low, suboptimal growth temperatures, where it would not be advantageous to replicate the DNA, but then will increase to physiologically useful nucleotide incorporation rates at temperatures more favorable for growth.

Immediate questions begged by these data are: what molecular events correlate with the TH transition (and the TH⁺ transition), and how might they translate into catalytic rate changes? Probably the biggest obstacle in attempting to answer such questions is the fact that the underlying molecular origins of the heat capacity of binding and the temperature dependencies of the enthalpy and entropy themselves are not yet understood. In other words, as mentioned above, the molecular origins for the existence of TH and TS, and the parallel ΔH and TΔS plots for binding (denoted enthalpy–entropy compensation), continue to be debated (e.g. (114-117)). Thus, understanding the molecular origins of these underlying thermodynamics (i.e. what bonds are made/broken, strengthened/weakened) is required before we can begin to truly understand what might be happening at the molecular level at TH to “switch on” catalysis. Perhaps upshift of enzymatic activity is not even specifically dependent on specific changes in non-covalent bonding that cause enthalpy–entropy compensation and manifest TH, but simply requires a little internal heat, a secondary effect generated by a negative enthalpy. In addition, a crucial question is simply asking whether TH/activity correlations, i.e. the patterns depicted in Figure 2.3, are general for other enzymes, or whether they are simply correlations for these two polymerases. Currently, the thermodynamic plus enzymatic activity data presented herein are the only existing data sets for any enzyme system where such a comparison can be made.
2.6 Acknowledgements

The authors thank the Michael Doughty lab at Southeastern Louisiana University for initial advice on the nucleotide incorporation assay. This work was supported by the National Science Foundation.
CHAPTER 3
FACTORS INFLUENCING PARTITIONING BETWEEN POLYMERIZATION AND EDITING SITES OF KLENOW POLYMERASE

3.1 Abstract

Klenow polymerase is the large fragment of DNA polymerase I from *E. coli*. Klenow possesses 5’-3’ polymerase and intrinsic 3’-5’ exonuclease activities with two distinct active sites that are located ~30 Å apart. During DNA replication, proofreading activity enhances replication fidelity by excising misincorporated nucleotides from the 3’ end of primer strand. The first step of the proofreading process is sending the 3’-primer terminus to the proofreading site, which requires separation of the primer terminus from the template strand. Here we examined the binding of Klenow to matched and mismatched primed-template DNA (pt-DNA) by monitoring the steady state fluorescence intensity change of a single 2-aminopurine base site-specified in the template strand within the duplex part of pt-DNA. The changes in fluorescence intensity allow us to follow shuttling of the primer terminus between the polymerization and proofreading sites.

We have found that the rate of partitioning of the primer between the two active sites depends on: 1) the number of mismatched bases at the primer-template junction, 2) the presence or absence of divalent ions, and 3) the type of divalent ions. Magnesium and calcium ions have opposite effects on the direction of the shift between the pol and exo sites. Substitution of the normal phosphodiester linkage between the last two bases of the primer strand with a non-hydrolysable phosphorothioate linkage and 2AP-T basepair at the primer-junction also have significant effects on the partitioning between sites.
3.2 Introduction

DNA polymerase I from *E. coli* is involved in both replication and repair of DNA. The Klenow “large fragment” of *E. coli* DNA polymerase I retains the 5’-3’ polymerase, 3’-5’ exonuclease (proofreading) domains. During DNA synthesis, proofreading activity enhances replication fidelity by excising misincorporated nucleotides from the 3’ end of the primer strand. The first step of the proofreading process is to send the misincorporated 3’-primer terminus to the proofreading site, which requires separation of the primer terminus from the template strand (15, 17). A cocrystal X-ray structure of a binary complex composed of a duplex matched DNA and Klenow shows that the two active sites are located ~30 Å apart, and that at least 3 bases of the 3’-primer terminus have been unwound from the template strand and bound to the 3’-5’ exonuclease (exo) site as a single strand (17). This structure shows the editing mode of DNA binding is populated even with matched DNA, and it is the only mode for which structural data exist for Klenow. On the other hand, a cocrystal structure of DNA bound to Klentaq polymerase (homologous to Klenow) from *T. aquaticus* shows the polymerization (pol) mode of DNA binding in which the 3’-terminus of the primer is base-paired with the template strand (19). These two structures represent different fates of the 3’-terminus of the primer strand upon DNA binding in the two different binding modes.

Accuracy of passing parental genetic codes to next generations is highly dependent on faithful DNA replication by DNA polymerases with a network of accurate nucleotide incorporation and exonucleolytic removal of incorrectly incorporated nucleotides once the misincorporation occurs. Several studies have attempted to elucidate the mechanism of sensing mismatched incorporation in the polymerase domain and switching of the primer strand from the pol site to the exo site (35, 44, 135), and the partitioning of the primer stand between the two
active sites of Klenow has been also studied with fluorescence based assays for some time (37-43, 136). Through these studies, now it is known that the number of mismatched bases at the primer-template junction influences the apparent equilibrium partitioning between the pol and exo sites; however, the relative partitioning of correctly matched pt-DNA between the two sites has still been controversial for D424A Klenow exo minus (KF-), with reports ranging from nearly 0% up to 43% occupancy at the exo site (37-43, 136). D424A is the most commonly used mutant Klenow for binding studies. A mutation from aspartic acid to alanine at position 424 amino-acid abolishes one of two divalent metal bindings at the exo site but retains proper single-stranded DNA binding with a reduced exonuclease activity (14).

DNA polymerase I from *E. coli* has served as a model for the elucidation of the polymerization and the proofreading activity since its discovery, and it still remains one of the most intensively studied polymerases. In order to enhance our understanding of the proofreading mechanism and to pinpoint the causes of the diverse reports of partitioning of KF- binding to correctly matched pt-DNA, we examined effects of divalent metal ions, residual exonuclease activity, length of incubation time, and DNA sequence on equilibrium partitioning between the pol and exo sites of KF-. The binding of KF- to matched and mismatched pt-DNA was monitored with a steady state fluorescence intensity change of a single 2-aminopurine (2AP) base site-specifically substituted in the template strand within the duplex part of pt-DNA in the absence/presence of divalent cations (Mg$^{2+}$ or Ca$^{2+}$). 2AP is a fluorescent analogue of adenine and capable of forming a base pair with thymine with minimal disruption of normal helical DNA geometry (137). The fluorescence of 2AP is very sensitive to its environment and quenched by stacking with its immediate 5’ and 3’ neighbors in the same DNA strand and/or by forming a basepair (138, 139). The fluorescence of 2AP can be recovered upon DNA melting where base
stacking interactions with the 2AP are disrupted. These changes in the fluorescence intensity of 2AP allow us to differentiate between the two modes of DNA binding and to follow shuttling of the primer terminus between the pol and exo sites.

In addition, the result from an exonuclease assay showed that KF- is still capable of exonuclease digestion in the presence of Mg$^{2+}$ and incapable in the absence of divalent ions or in the presence of Ca$^{2+}$. Consequently, we observed a significant increase in the 2AP fluorescence intensity change as a function of time from KF- binding to the 2AP incorporated DNA in the presence of Mg$^{2+}$. As KF- shuttles the primer from the pol site to the exo site and removes nucleotides from the 3’-end of the DNA construct, the buried 2AP on the template strand will be exposed to the solution by the loss of its base-pairing. In order to prevent the fluorescence intensity change caused by residual removal of nucleotides from the 3’-end, we replaced the 3’ terminal phosphodiester bond with a phosphorothioate linkage which is resistant to the 3’-5’ exonuclease activity (20, 140, 141).

Using various pt-DNA constructs we have found that partitioning of the primer between the two active sites depends on: 1) the number of mismatched bases at the primer-template junction, 2) presence or absence of divalent ions, 3) the type of divalent ions, 4) time in the presence of the divalent ions, and 5) DNA sequence. A significant time dependence of the partitioning between the two active sites was observed with the matched and single mismatched pt-DNA in the presence of divalent ions. Magnesium and calcium ions have opposite effects on the direction of the shift between the pol and the exo sites. In addition to the number of mismatched bases at the primer/template junction, DNA sequence and a non-hydrolysable phosphorothioate linkage also have significant effects on the partitioning between the sites.
3.3 Materials and methods

Proteins and DNA

Purification of the proteins has been described previously (54). All DNA oligonucleotides used in this study were purchased from Integrated DNA Technologies (Coralville, IA), and the nomenclature, sequence, structure, position of 2AP, and presence or absence of the phosphorothioate linkage of DNA constructs are summarized in Table 3.1. The pt-DNA constructs with a hairpin end were prepared in STE buffer (50 mM NaCl, 10 mM Tris, and 1 mM EDTA, pH 8) by heating 100 μM of single stranded DNAs at 94 °C for 4 minutes and then were gradually cooled down to room temperature. For the pt-DNA with a blunt end, annealing was done in the presence of equimolar amounts of primer and template strands (100 μM each).

Table 3.1. Sequences of DNA constructs. P represents the 2AP base. The underlined bases indicate mismatched bases introduced into the template region of primer-template DNA. Phosphorothioate bonds are shown as star symbols (*).

<table>
<thead>
<tr>
<th>DNA constructs</th>
<th>DNA sequences</th>
</tr>
</thead>
<tbody>
<tr>
<td>0mmDNA</td>
<td>AAGTCGCA GCGTCTA-3’&lt;br&gt;AGCAGCGT CGGCGAGPTTCCC A A A-5’</td>
</tr>
<tr>
<td>S-0mmDNA</td>
<td>AAGTCGCA GCGTCT*A-3’&lt;br&gt;AGCAGCGT CGGCGAGPTTCCC A A A-5’</td>
</tr>
<tr>
<td>S-1mmDNA</td>
<td>AAGTCGCA GCGTCT*C-3’&lt;br&gt;AGCAGCGT CGGCGAGPTTCCC A A A-5’</td>
</tr>
<tr>
<td>S-3mmDNA</td>
<td>AAGTCGCA GCGTGCTAC*C-3’&lt;br&gt;AGCAGCGT CGGCGAGPTTCCC A A A-5’</td>
</tr>
<tr>
<td>†0mmDNA*</td>
<td>5’-GCAA GAAACC GAAACCAP-3’&lt;br&gt;3’-CGTTCTTGGCTTGTTAAACTAAC-5’</td>
</tr>
<tr>
<td>†3mmDNA*</td>
<td>5’-GCAA GAAACC GAAACCAP-3’&lt;br&gt;3’-CGTTCTTGGCTTGTTAAACTAAC-5’</td>
</tr>
</tbody>
</table>

† The sequences of 0mmDNA* and 3mmDNA* are nearly same as the ones from reference (41). The only difference is the 2nd base from the 3’ end of the primer strand is adenine in this study and 2AP in reference (41).
Other materials

All reagents were from Amresco, Sigma Chemical, or Fisher Scientific. The buffers used for the steady state fluorescence measurements were composed of 10 mM Tris, 50 mM KCl, and either 5 mM MgCl$_2$ (Mg$^{2+}$ buffer), 5 mM CaCl$_2$ (Ca$^{2+}$ buffer), or 3 mM EDTA (EDTA buffer) at pH 7.9. *Ca$^{2+}$ buffer was composed of 20 mM Hepes, 50 mM sodium acetate, 5 mM Ca(OAc)$_2$, and 1 mM dithiothreitol at pH 7.9 and from reference (41).

Steady state fluorescence intensity measurement of 2AP labeled DNA constructs in the presence or absence of KF- or Klentaq polymerase

Steady-state fluorescence measurements were performed at 25 °C on a FluoroMax-4 spectrofluorometer (HORIBA Scientific). Fluorescence emission spectra of 500 nM of 2AP labeled DNA constructs were collected from 330 to 460 nm (6 nm bandwidth) with excitation at 315 nm (2 nm bandwidth) in the absence or presence of 500 nM proteins. The 2AP signal was obtained by subtracting signal contributions from buffer and/or protein. For time course of intensity measurements, the 2AP signal at 370 nm was followed for various time periods (see the result section). The average of 3 measurements of the buffer signal or the buffer-protein signal at 370 nm was subtracted from the raw signal. Time dependence of fluorescence intensity changes, data were fit with a half-life equation:

$$I(t) = I_P + I_0 \times \left(\frac{1}{2}\right)^{t/t_{1/2}}$$

where: $I(t)$ is the intensity after a time $t$, $I_P$ is the intensity at a plateau, $I_0$ is the initial intensity at time zero, and $t_{1/2}$ is the half-life of the intensity change. Non-linear analysis was done with the program Kaleidagraph (Synergy Software).
Measurements of the residual 3’-5’ exonuclease activity of KF-

A total volume of 50 μL reaction mixture containing 500 nM each of KF- and 0mmDNA/S-0mmDNA was incubated for 0, 1, 2, 3, 4, 6, and 8 hours in Mg\(^{2+}\) buffer at room temperature. At each time point, 5 μL was removed from the reaction mixture and added into the same volume of gel loading dye (formamide with 10 mM EDTA, 5 M urea, 13 % w/v sucrose, bromophenol blue, and xylene cyanol). The mixture was heated at 95 °C for 5 minutes and placed on ice until loaded on to a denaturing gel (1×TBE, 20 % polyacrylamide, 8 M urea). The gel was stained with SYBR Green I (Life Technologies), and quantified with ImageQuant Version 5.1 (Molecular Dynamics). For the quantification of the residual exonuclease activity of D424A Klenow, degradation of the original pt-DNA bands (intact pt-DNA) were quantified, and over 90 % of the original band was digested after 8 hours of incubation. The lowest and highest values are normalized to 0 and 100 %, respectively. The residual exonuclease activity was shown as percent change in digestion in Figure 3.1; B, and the half-life of the activity was determined.

3.4 Results and discussions

Effects of residual exonuclease activity of KF- on steady state fluorescence intensity of 2AP labeled DNA

We examined KF- binding to 2AP-labeled matched 0mmDNA in Mg\(^{2+}\), Ca\(^{2+}\), and EDTA buffer. 0mmDNA contains 2AP on the template strand, which is base paired with thymine (T) at the 2\(^{nd}\) base-pair from the primer/template junction (Table 3.1). We would not expect an increase in fluorescence intensity when KF- binds to this 2AP-labeled DNA at the pol site since 2AP should remain base paired with T on the primer strand. On the other hand, if KF- binds to this DNA substrate in the exo mode, our expectation would be an increase in fluorescence intensity since the exo mode of binding requires 3-4 bases of the primer strand to unwind from the template strand. As a result, the 2AP in 0mmDNA would lose its base pairing partner and
experience less base stacking, which would cause an increase in fluorescence intensity.

Figure 3.1; A shows the fluorescence intensity change of 0mmDNA upon complex formation with KF- in the presence of Mg$^{2+}$. The fluorescence intensity of 2AP in the 0mmDNA/KF- complex starts at nearly same level of intensity as 0mmDNA itself (data not shown), then increases, and reaches a plateau after 4 hours of incubation. The intensity change was gradual and significant (~10X higher intensity at the plateau compared with the initial complex intensity). It thus appears that KF- binding to 0mmDNA was initially at the pol site and that the primer strand gradually shuttled to the exo site. Figure 3.1; B shows the residual exonuclease activity of KF- on 0mmDNA in the presence of Mg$^{2+}$. In a time course of 8 hours, over 90% of the original DNA was digested. In Figure 3.1; A, the kinetic data for the exonuclease activity were superimposed with the fluorescence intensity change of 0mmDNA/KF-. Half-lifes of the exonuclease activity and the intensity change were obtained by fitting data into the half-life equation and determined as 1.43 and 1.03 hours, respectively. The correspondence between the residual exonuclease activity and the 2AP intensity change suggests that the residual exonuclease activity causes to the change in intensity of 2AP labeled 0mmDNA/KF- complex. As KF- shuttled the primer from the pol site to the exo site and removed nucleotides from the 3’ end of the DNA construct through the residual exonuclease activity, the 2AP on the template strand becomes exposed, and fluorescence intensity increases. These changes were observed only in the presence of Mg$^{2+}$ which is a catalytic co-factor for the exonuclease activity (20). Surprisingly, no previous studies of the pol to exo site partitioning of matched DNA appear to take this effect into account. Clearly the time at which one measures the pol:exo ratio will yield different answers unless extrapolated back to time zero. Such extrapolation to time zero does not appear to have been performed in any previous study, most of which have collected data in the 0-
2 hour time window, and this may account for some of the reported variation in the initial pol:exo mode partitioning. We show below, however, that even if this time dependent effect is accounted for, other reaction factors also strongly influence the pol:exo partitioning.

Figure 3.1. Steady state fluorescence intensity change of 2AP labeled matched 0mmDNA/KF-complex and the effect of the residual exo nuclease activity of KF-. In panel A, 0mmDNA fluorescence intensity change upon complex formation with KF- in the presence of Mg$^{2+}$ ions (open circles). Panel B shows the residual 3’-5’ exonuclease activity of KF- on 0mmDNA in the presence of Mg$^{2+}$ ions. Within 8 hours of incubation, over 90 % of the original pt-DNA was digested. The lowest and highest values of the quantified intact DNA (top bands in Panel B) were normalized to 0 and 100 %, respectively, and they are shown as closed squares in Panel A. Half-lifes of the exonuclease activity and the intensity change were determined as 1.43 and 1.03 hours, respectively.

Effects of divalent ions on the partitioning of the primer strand between the pol and exo sites

KF- is still capable of exonuclease digestion (Figure 3.1). In order to eliminate fluorescence intensity changes caused by digestion of the 2AP labeled DNA substrates, a
phosphodiester bond at the 3’ end of DNA was replaced with a phosphorothioate linkage which is resistant to the 3’-5’ exonuclease activity (140, 141). Exonuclease activity of KF- on the phosphorothioate linked DNA substrates was absent in the presence of Mg$^{2+}$ (Figure 3.2).

![Incubation Time (hr)](image)

Figure 3.2. Inhibition of the residual 3’-5’ exonuclease activity of KF- by the phosphorothioate linked S-0mmDNA in the presence of Mg$^{2+}$ ions. S-0mmDNA and KF- were incubated up to 8 hours in the presence of Mg$^{2+}$ ions.

Steady state fluorescence intensity of 2AP upon DNA/polymerase complex formation was measured in the presence of divalent ions (Mg$^{2+}$/Ca$^{2+}$) or in the absence of divalent ions (EDTA) with phosphorothioate linked S-0mmDNA, S-1mmDNA, and S-3mmDNA. The sequence of the template strand region is kept the same among these pt-DNAs since fluorescence of 2AP is greatly influenced by stacking interactions with nearest neighbor nucleotide bases (138). In both S-0mmDNA and S-1mmDNA, 2AP is base paired with a thymine on the primer strand and 2AP is mispaired and exposed in S-3mmDNA.

Figure 3.3; A, B, and C show the fluorescence intensity changes of 2AP within the matched and single mismatched pt-DNAs upon complex formation with KF- and Klentaq. Klentaq is used as a control in these experiments since it will only bind DNA in the pol mode. In the presence of divalent ions, time dependence of fluorescence intensity changes was observed with S-0mmDNA and S-1mmDNA complexed with KF- (Figure 3.3; A and B). Magnesium and calcium ions have opposite effects on the direction of the intensity change. The shift of the
intensity is from low to higher with Mg\textsuperscript{2+} (Figure 3.3; A) and vice versa with Ca\textsuperscript{2+} (Figure 3.3; B). Half-lifes of the intensity shifts were determined by fitting the data with a half-life equation and reported in Table 3.2. Longer half-lifes were observed in the presence of Mg\textsuperscript{2+} compared to Ca\textsuperscript{2+} and with S-1mmDNA/KF- complex compared to S-0mmDNA/KF- complex. The time dependence of the intensity shift was absent in the absence of divalent ions (Figure 3.3; C). In addition, Klentaq binding to S-0mmDNA (Figure 3.3; A, B, and C) and KF- binding to S-3mmDNA (data not shown) did not show any time dependence of fluorescence intensity under any solution conditions. It bears emphasis that because these phophorothioated DNA substrates are nuclease resistant, these time dependent shifts in the pol:exo partitioning for Klenow are a second exonuclease activity independent time dependent shift in the binding mode partitioning. For Mg\textsuperscript{2+}, given the relative t\textsubscript{1/2} values, it is probable that this initial pol → exo time dependent shift feeds that ongoing exonuclease activity in normal DNA by mass action.

Figure 3.3. Time dependence of the fluorescence intensity changes of 2AP labeled matched and single mismatched DNA with a phosphorothioate bond upon complex formation with KF- and Klentaq (KTQ). Time dependence of fluorescence intensity changes upon DNA binding were measured in the presence of 5 mM Mg\textsuperscript{2+} (panel A), 5 mM Ca\textsuperscript{2+} ions (panel B), and 3 mM EDTA (panel C). The data were fitted either into the half-life equation (KF- binding to S-0mmDNA and S-1mmDNA with Mg\textsuperscript{2+} and Ca\textsuperscript{2+} ions) or the linear equation (Klentaq binding to S-0mmDNA and KF- binding to S-0mmDNA and S-1mmDNA with EDTA).
**Effects of DNA context on 2AP spectral changes**

Figure 3.4; A, B, and C show the emission spectra of 2AP embedded in S-0mmDNA, S-1mmDNA, and S-3mmDNA in Mg$^{2+}$, Ca$^{2+}$, and EDTA buffers, respectively. Steady state fluorescence intensity of 2AP was sensitive to base stacking and the solution conditions. The 2AP intensity was greater with increase in mismatched bases on the primer strand. 2AP was least stacked with the adjacent bases within the single-strand region of S-3mmDNA compared to the 2AP base-paired with a thymine on the primer strand in S-0mmDNA and S-1mmDNA. The presence of divalent cations, Mg$^{2+}$ and Ca$^{2+}$, enhanced the fluorescence intensity of 2AP (i.e. destabilized the base-stacking of 2AP with the adjacent bases). The effect of the intensity enhancement was the greatest on S-3mmDNA and the least on S-0mmDNA. Such enhancement by Mg$^{2+}$ and Ca$^{2+}$ ions was previously observed (142).

Figure 3.4; D, E, and F show the emission spectra of 2AP of the pt-DNA/polymerase complexes, which were taken after the fluorescence intensity reached the plateaus in Mg$^{2+}$, Ca$^{2+}$, and EDTA buffers, respectively. The fluorescence intensity was the lowest when Klentaq bound to S-0mmDNA in the presence of divalent ions (Figure 3.4; D and E). In the absence of divalent ions, KF- binding to S-0mmDNA showed the least fluorescence intensity (Figure 3.4; F). Klentaq binding to S-0mmDNA did not change the fluorescence intensity of 2AP compared to that of S-0mmDNA itself in any solutions. A co-crystal structure of Klentaq bound to DNA at the pol site shows that the DNA is mostly in the B-form, except for the three base pairs at the end of the duplex DNA adjacent to the active site, which are A-form (19). However, the binding results for Klentaq did not indicate any change in conformation of S-0mmDNA from B-form upon formation of the binary complex. On the other hand, KF- binding to S-0mmDNA in the absence of divalent ions showed ~10 % reduction in the fluorescence intensity from the intensity of S-
0mmDNA itself (Figure 3.4; C and F), which indicates that 2AP has higher base-stacking with the adjacent bases in this complex. This reduction in intensity may reflect the conformational change of S-0mmDNA from B-form to A-form at the pol site of KF-. Such observations were previously made with fluorescence and CD measurements to monitor local conformation of 2AP dimer at the 3’ end of the primer when pt-DNA bound to KF- or Klentaq by Datta et al., as well (41).

The fluorescence intensity of 2AP was the highest when KF- binds to S-3mmDNA in any solutions (Figure 3.4; D, E, and F). S-3mmDNA binding to KF- produced a ~2.3X higher intensity in the presence of divalent ions and a ~2.7X higher intensity in the absence of divalent ions compared to the intensity of S-3mmDNA alone (A and D for Mg$^{2+}$, B and E for Ca$^{2+}$, and C and F for EDTA in Figure 3.4). A co-crystal structure of DNA/KF- complex showing the editing mode of DNA binding (the primer strand bound at the exo site) misses the template bases which the 3-4 bases of the unwound primer strand could base pair with (17). Hence it is unclear how the 2AP and nearest neighboring bases interact with KF- in the editing mode of DNA binding since the positions of these bases fall into the positions of the three missing template bases in the crystal structure. Our data of S-3mmDNA binding by KF- indicate that at least one of the three bases interacts with KF- resulting in the unstacking of 2AP with the adjacent bases.

**Pol ↔ Exo partitioning**

The percent partitioning of the primer strand at the exo site was calculated from the fluorescence intensities of DNA/Polymerase complexes from Figure 3.3; A, B, and C. Here we assumed that the intensity of triple-mismatched S-3mmDNA/KF-complex represents the 100 % exo mode binding, and that the intensity of matched S-0mmDNA/Klentaq complex represents the 0 % exo (100 % pol) mode of binding. Klentaq lacks the 3’-5’ exonuclease activity (143) and
does not bind to single-stranded DNA (58). KF- binding to three terminal mismatched pt-DNA has been shown to be the optimal substrate binding at the exo site (41). By setting the two end states, we estimated percent occupancies of the 3’ end of S-0mmDNA and S-1mmDNA at the exo site of KF- from the fluorescence intensities produced when the binary complex formed. We used the following equation to calculate the exo site binding of KF-: 

\[ \% \text{exo site binding of KF-complex} = \left( \frac{F_{I_{\text{KF-complex}}} - F_{I_{S-0mmDNA/Klentaq}}}{F_{I_{S-3mmDNA/KF-}} - F_{I_{S-0mmDNA/Klentaq}}} \right) \times 100 \]

where \( F_{I} \) is fluorescence intensity and KF-complex is either S-0mmDNA/KF- or S-1mmDNA/KF-complex. The average fluorescence intensities of triplicates of S-0mmDNA/Klentaq and S-3mmDNA/KF- measurements were used for \( F_{I_{S-0mmDNA/Klentaq}} \) and \( F_{I_{S-3mmDNA/KF-}} \) in the above equation. For the estimation of S-0mmDNA and S-1mmDNA binding to KF- in the presence of divalent ions where we observed the time dependence of the fluorescence intensity change, first we fit the data to the half-life equation to obtain the initial and plateau intensity values. Then those values were individually used to calculate the initial and plateaued percent partitioning of the primer at the exo site.

We summarize the percent partitioning of the primer strand at the exo site along with the half-lifes of the transitions in Table 3.2. The reported values were the average of either duplicates or triplicates. We obtained negative values when KF- was bound to S-0mmDNA in Mg\(^{2+}\) and EDTA buffers. The negative values originate from the fluorescence intensity of DNA/KF-complex being lower than that of DNA/Klentaq complex, which indicates that the conformation of S-0mmDNA at the KF- pol site has a higher degree of base stacking (potentially indicating A-form DNA) compared with the bound S-0mmDNA on Klentaq (potentially indicating B-form DNA). Datta et al. previously showed that the conformation of DNA at the primer/template junction is A-form when KF- bound to a matched DNA in the absence of metal ions and B-form
when Klentaq bound to a matched DNA in the presence or absence of divalent ions (41).

However, since it is not known whether there exists A- to B- or B- to A-form transitions of DNA at the pol site of Klenow, we simply kept the level of intensity produced from the S-0mmDNA/Klentaq complex as our control for the 0 % level of exo site binding in each buffer. Hence our results could be a slightly lower estimation of the primer binding at the exo site.

Figure 3.4. Fluorescence emission spectra of the 2AP labeled matched and mismatched DNA substrates before and after complex formation with Klenow (KF-) and Klentaq (KTQ) in Mg$^{2+}$, Ca$^{2+}$, and EDTA buffer. Fluorescence emission spectra of the DNA substrates in Mg$^{2+}$, Ca$^{2+}$, and EDTA buffer are shown in panel A, B, and C, respectively. Fluorescence emission spectra of DNA/Polymerase complexes in Mg$^{2+}$, Ca$^{2+}$, and EDTA buffer were taken after the fluorescence intensity reached the plateaus and shown in panel D, E, and F, respectively. Fluorescence emission spectra were collected from 330 to 460 nm with excitation at 315 nm. The 2AP signal was obtained by subtracting signal contributions from buffer and/or protein. All spectra show the peak around 370 nm.
Table 3.2. Summary of the primer strand partitioning of the phosphorothioate linked matched and single mis-matched ptDNA at the exo site of KF-.

<table>
<thead>
<tr>
<th></th>
<th>Mg</th>
<th>Ca</th>
<th>EDTA</th>
</tr>
</thead>
<tbody>
<tr>
<td>S-0mmDNA</td>
<td>Time Dependent</td>
<td>Time Dependent</td>
<td>No Time Dependent</td>
</tr>
<tr>
<td></td>
<td>-0.7 % → 2.5 %</td>
<td>1.19 % → 0.7 %</td>
<td></td>
</tr>
<tr>
<td>t&lt;sub&gt;1/2&lt;/sub&gt;</td>
<td>= 21.7 min</td>
<td>= 3.5 min</td>
<td></td>
</tr>
<tr>
<td>S-1mmDNA</td>
<td>Time Dependent</td>
<td>Time Dependent</td>
<td>No Time Dependent</td>
</tr>
<tr>
<td></td>
<td>18.4 % → 21.9 %</td>
<td>35.9 % → 31.3 %</td>
<td>9.5 %</td>
</tr>
<tr>
<td>t&lt;sub&gt;1/2&lt;/sub&gt;</td>
<td>= 30.3 min</td>
<td>= 5.8 min</td>
<td></td>
</tr>
</tbody>
</table>

The results indicate that the primer binding at the exo site depends on the presence of a mismatched base at the primer/template junction, the presence of divalent ions, and the type of divalent ions. For instance, KF- binding to the single mismatched pt-DNA (S-1mmDNA) promotes the primer binding more at the exo site under any solution conditions (S-0mmDNA versus S-1mmDNA). This mismatch effect is well known and has been characterized in detail by Millar and associates (37-39, 42, 136). Our results for S-1mmDNA binding to KF- highlight the effect of divalent ions on the primer strand binding at the exo site. The presence of divalent ions promotes the primer binding at the exo site. Here we observed around twice and three times higher partitioning of the primer strand at the exo site in the presence of Mg<sup>2+</sup> and Ca<sup>2+</sup>, respectively compared with the absence of metal ions (EDTA buffer). The presence of Ca<sup>2+</sup> promoted the primer binding at the exo site with higher partitioning than Mg<sup>2+</sup>, indicating that Ca<sup>2+</sup> likely binds at the exo site even though Ca<sup>2+</sup> ion is not a cofactor for the exonuclease activity. In this study we also captured another unexpected behavior of the primer strand binding between the pol and exo sites in the presence of metal ions, which is the time dependence of the primer strand partitioning between the sites. Half-lifes of transitions were slower with Mg<sup>2+</sup>, and Mg<sup>2+</sup> and Ca<sup>2+</sup> ions have opposite effects on the direction of the shift between the pol and exo sites. In the presence of Mg<sup>2+</sup>, the shift of the partitioning was from less to more exo site binding.
of the primer terminus and vice versa with Ca$^{2+}$.

**Effect of phosphorothioate linkage on the partitioning of the primer strand between the pol and exo sites**

Phosphorothioate linked DNA, in which one nonesterified oxygen atom of a phosphate is replaced by sulfur, has Rp and Sp isomers. A structural study for inhibition of the 3’-5’ exonuclease activity of wild type Klenow by phosphorothioates showed that the Rp isomer inhibits the activity by inducing small changes in the binding of substrate, nucleophile, and the catalytic cations and that the Sp isomer inhibits the reaction by prohibiting metal ions from binding in concert with the Sp isomer at the active site (20). Our phosphorothioate linked DNAs are composed of a mixture of Rp and Sp isomers. Figure 3.5 shows comparison of fluorescence intensity changes of matched DNA/KF-complex with the phosphorothioate linkage (S-0mmDNA/KF-) and the normal phosphodiester bond (0mmDNA/KF-) in Ca$^{2+}$ buffer. 0mmDNA and S-0mmDNA have the same sequence (Table 3.1). In the absence of the phosphorothioate linkage, 0mmDNA/KF-complex showed much greater time dependence of the intensity changes. The fluorescence intensity of 0mmDNA/KF-complex was much higher than that of S-0mmDNA/KF-, and the half-life of the intensity transition of 0mmDNA/KF- was slower ($t_{1/2} = 8.5$ min) than that of S-0mmDNA/KF- ($t_{1/2} = 3.5$ min reported in Table 3.2). The results indicate that the primer strand of the phosphodiester bonded 0mmDNA which does not compete with divalent ions at the exo site could participate more at the exo site resulting in the higher fluorescence intensity and the slower half-life of the transition. From the results, we can speculate that our phosphorothioate linked pt-DNAs compete binding with Ca$^{2+}$ (possibly with Mg$^{2+}$ as well) at the exo site of KF-.
57

Figure 3.5. Comparison of fluorescence intensity changes of matched DNA/KF- complex with the phosphorothioate linkage (S-0mmDNA/KF-) and the phosphodiester bond (0mmDNA/KF-). 0mmDNA and S-0mmDNA have the same sequence. In the absence of the phosphorothioate linkage, 0mmDNA/KF- complex showed much higher time dependence of the intensity changes than S-0mmDNA/KF- complex. The data were fit with the half-life equation.

**Effects of DNA sequence on the partitioning of the primer strand between the pol and exo sites**

The previous studies that determined the exo site binding of primer strand of correctly matched pt-DNA reported a wide range of partitioning at the exo site of D424A Klenow, KF-, from nearly 0% up to 43% (38-43, 136). Millar and co-workers determined partitioning constants which are equivalent to the ratio of the equilibrium fractions of DNA bound in the editing and polymerization modes of D424A Klenow for matched DNAs with different sequences by measuring time-resolved fluorescence polarization anisotropy of a dansyl fluorophore attached to a primer base located eight base pair upstream of the primer/template DNA in the presence of Mg$^{2+}$. The changes in the lifetime of the dansyl fluorophore appeared to depend mostly on the local environment occupied by the probe within the protein (i.e., buried versus partially exposed) and were correlated with specific binding conformations of the primer to provide an estimate of the fractional occupancy of the pol and the exo sites (38, 39, 42, 136). Among those studies, they used four different sets of matched DNA sequences. When they used a primer sequence of the last four bases 5’-AATG-3’ (the sequence of four bases of the primer
strands is shown), they estimated about 7% (42) and 14% (136) of the primer end to be at the exo site. Making their primer/template junction more GC-rich (5’-AGGG-3’) showed about 7% partitioning at the exo site (38). Making their primer-template junction more AT-rich (5’-GTNTT-3’) showed 3% occupancy at the exo site. On the other hand, another AT-rich sequence (5’-ATTT-3’) which is the same sequence as the previous one, except the 4th base is changed to A from G increased the exo site binding to 43%, and its single terminal mismatched DNA (T/T mismatch) made by the same primer sequence reduced the DNA binding at the exo site to 23% (39).

A smFRET study in the presence of Mg$^{2+}$ for measurements of real time polymerization dynamics with a Cy3 fluorophore labeled pt-DNAs and Cy5 labeled exonuclease deficient Klenow showed around 24% of the primer 5’-AGGA-3’ and 30% of the primer 5’-GGAA-3’ of matched DNAs bound at the exo site of D424A Klenow in the presence of Mg$^{2+}$ (40). Another smFRET study for nucleotide selection of D424A Klenow in the presence of Mg$^{2+}$ showed that D424A Klenow binding to a matched DNA with a 5’-AAGC-3’ primer sequence was bound exclusively at the pol site (43). These two smFRET studies utilized the same technique, including the assay buffers, Cy3 labeling on a template DNA strand, and the position of Cy5 attachment to KF-. Because researchers assign a smFRET efficiency based on their understanding of the molecules being studied, the interpretation of the FRET efficiency could vary from one researcher to another. One smFRET study assigned the lower FRET efficiency as representing the exo site binding of the primer when they attached Cy3 on a template base 11 base pairs away from the primer/template junction (40). On the other hand, the other study assigned FRET efficiency in an opposite way, even with the DNA being similarly Cy3-labeled on a template base 9 base pairs away from the primer/template junction (43). Their interpretations
seem to contradict each other.

Stability of the primer/template junction (the relative contributions of hydrogen bonding and base stacking) has been shown to be a major factor for formation of exonuclease complexes with phage T4 DNA polymerase which has a 3’-5’ exonuclease activity and also requires unannealing of the terminal 3 base pairs at the primer end (reviewed in (144) and (145)). In the case of KF-, it is hard to find the effects of DNA sequences at the primer/template junction from the above studies (38-43, 136). Compared to those, the results shown here with the phosphorothioate linked matched DNA binding (S-0mmDNA: 5’-TCTA-3’) by KF- showed the lower end of the partitioning at the exo site (Table 3.2).

**Effects of terminal 2AP-T base pairing on the partitioning of the primer strand between the pol and exo sites**

Datta et al. observed residual exonuclease activity of D424A Klenow during circular dichroism measurements of primers containing 2AP dimer at the 3’ terminus of pt-DNA within a period of hours. In order to maintain integrity of their DNA substrates, they replaced Mg$^{2+}$ with Ca$^{2+}$. Their fluorescence and CD measurements which monitor local conformation of a 2AP dimer at the 3’ end of the primer (5’-CCAA-3’; here A is 2AP) showed about a 43% occupancy at the exo site for matched pt-DNA (41). Their result showed the high end of the partitioning at the exo site even with the proper control of the residual exonuclease activity of KF-. Hence, we decided to use their DNA sequence in our experimental setup to evaluate the effect of 2AP-T base pair at the primer/template junction. The DNA sequences for matched (0mmDNA*) and triple mismatched (3mmDNA*) pt-DNAs are shown in Table 3.1, and they are the same sequences used by Datta et al. (41), except a single 2AP at the 3’ end of the primer strand was used instead of double 2AP probes. By using these pt-DNAs, we measured steady state fluorescence intensity of 2AP upon DNA/polymerase complex formation in the presence of Ca$^{2+}$.
The buffer condition used here is the same as used in reference (41), and the buffer is denoted as “*Ca2+ buffer”.

Figure 3.6; A shows time dependence of the fluorescence intensity change of 0mmDNA*/KF- and of 0mmDNA*/Klentaq complexes. Klentaq binding to 0mmDNA* was not time dependent and showed the least intensity. 0mmDNA*/KF- complex showed time dependence of intensity change as seen in Figure 3.3; B. KF- binding to 3mmDNA* did not show time dependence of fluorescence intensity change (data not shown).

Figure 3.6; B shows the emission spectra of 2AP embedded in 0mmDNA* and 3mmDNA* in *Ca2+ buffer. Steady state fluorescence intensity of 2AP in these DNA constructs is much higher than that of 0mmDNA and 3mmDNA (Figure 3.4; B), indicating that 2AP at the 3’ terminal end of the primer of pt-DNA has less of a base stacking effect since there is only one neighboring base next to it. Figure 3.6; C shows the emission spectra of 2AP of the pt-DNA/polymerase complexes, which were taken after the fluorescence intensities reached the plateaus.

The percent partitioning of the primer strand of 0mmDNA* at the exo site and the half-life of the transition were calculated as described earlier. For comparison with our DNA sequence, we measured KF- binding to S-1mmDNA in *Ca2+ buffer as well (data not shown), and the results are listed in Table 3.3.

2AP is capable of forming a Watson-Crick base pair with thymine (137). A previous study showed that polymerases are still able to discriminate between incorporation of 2AP and adenine opposite thymine although misincorporation frequencies for 2AP are much higher than for natural nucleotides (146). However, our results show that 2AP located at the primer/template junction of matched 0mmDNA* has a ~10 % higher occupancy at the exo site than 2AP located
in the 2nd base of the primer of single C-T terminal mismatched S-1mmDNA (Table 3.3). The determined 45.8-38.5 % exo site partitioning of 0mmDNA* having the 2AP-T base pair at the primer/template junction was similar to the exo site partitioning of 2AP dimer located at the 3’ end of the primer of the matched pt-DNA measured by Datta et al. (43 %) (41). These results indicate that DNA polymerase I can detect non-natural base pair (2AP-T) at the primer/template junction and send the 3’ primer termini to the proof reading site for removal of the non-natural base, 2AP. The higher exo site partitioning of 0mmDNA* would be unexpected if we consider the primer/template stability based only on the A-T or G-C richness. A misincorporation study of T4 DNA polymerase showed that five base pairs, both upstream and downstream from the site of 2AP misinsertion contributed to neighboring sequence stability (147). Hence, we also need to consider effects of sequence dependent base stacking on local DNA stability.

Figure 3.6. Time dependence of fluorescence intensity change of 0mmDNA*/KF- complex and of 0mmDNA*/KTQ complex. The 0mmDNA* sequence was shown as the sequence that gives the high primer strand partitioning (~43 %) at the exo site of KF- (41). The percent partitioning of the primer strand at the exo site was calculated in the same way described earlier. Here we assumed that the intensity of triple-mismatched 3mmDNA*/KF-complex shows 100 % exo mode binding. The obtained values were listed in Table 3.3. Panel A shows the time dependence of steady-state fluorescence change of 0mmDNA*/KF- and 0mmDNA/Klentaq complexes. Fluorescence emission spectra of 2AP buried within the matched and mismatched DNA substrates are shown in panel B. Fluorescence emission spectra of DNA/Polymerase complexes were taken after the fluorescence intensity reached the plateaus and are shown in panel C.
Table 3.3. Comparison of the primer strand partitioning at the exo site of 2AP positioned at the 3’ terminus of the matched and 2AP positioned at the 2\textsuperscript{nd} base from the 3’-end of the single mis-matched pt-DNA.

<table>
<thead>
<tr>
<th><em>\textsuperscript{2+}Ca\textsuperscript{</em>}</th>
<th>0mmDNA*</th>
<th>Time Dependent</th>
<th>45.8 % → 38.5 %</th>
<th>( t_{1/2} = 7.8 ) min</th>
</tr>
</thead>
<tbody>
<tr>
<td>S-1mmDNA</td>
<td>Time Dependent</td>
<td>32.0 % → 27.8 %</td>
<td>( t_{1/2} = 7 ) min</td>
<td></td>
</tr>
</tbody>
</table>

3.5 Concluding summary

The partitioning of the primer stand between the pol and exo sites of Klenow has been studied and debated for some time. It has been well established that the number of mismatched bases at the primer/template junction influences the apparent equilibrium partitioning between the active sites (37-39, 42, 136), but the initial relative partitioning between the sites has been controversial especially for binding of correctly matched DNA by D424A Klenow (38-43, 136). The Klenow D424A mutant has been preferentially used for DNA binding studies since D to A mutation largely diminishes the exonuclease activity by abandoning one of two metal bindings but retains proper substrate binding at the exonuclease catalytic site (14); however, this residual exonuclease activity is frequently ignored.

Our findings address much of the controversy in establishing the relative partitioning between the two sites by showing that it is extremely dependent on the time frame of the experiment and on exactly when and under what solution conditions partitioning at the pol and exo sites is measured. Using non-hydrolyzable phosphorothioate DNA constructs, the dynamics of partitioning between the two sites can be shown to be due to both an intrinsic kinetic shift in partitioning (Figure 3.3), and to residual exonuclease activity in D424A Klenow (Figure 3.1).
This study showed that the partitioning of matched and single mismatched pt-DNA between the two sites is a slow dynamic process in the presence of divalent cations (Figures 3.3, 3.5, and 3.6), and that the residual exonuclease activity causes the time dependent shift of apparent binding modes in the presence of Mg\(^{2+}\) ions (Figure 3.1)

Slow dynamic shifts in binding caused by presence of magnesium or calcium ions occur in opposite directions of the shift between the pol and exo sites. In the absence of divalent ions, the time dependence of the shift is absent. At this moment, we are unsure what causes these shifts in the presence of divalent ions when KF- binds to matched and single mismatched pt-DNA.
CHAPTER 4
DIFFERENT MODES OF INTERACTION BETWEEN NUCLEIC ACIDS
AND HIV-1 REVERSE TRANSCRIPTASE ARE DELINEATED BY THE
SALT DEPENDENCE OF PRIMER/TEMPLATE BINDING

4.1 Abstract

Direct binding equilibria and free energies have been characterized for the binding of
HIV-RT to several different primer/template nucleic acid constructs across a range of KCl
concentrations using a fluorescence anisotropy assay. At least two different binding modes
between nucleic acid duplexes and HIV-1 reverse transcriptase (HIV-RT) are revealed by the
thermodynamic salt linkages of the binding of the enzyme to different homo- and
heteroduplexes. When binding to 20/27mer primed-template duplexes, HIV-RT binds with
almost identical affinity to DNA/DNA and RNA/RNA duplexes at all salt concentrations
examined. Binding of either homoduplex results in the linked release of ~5.6 ions. In contrast,
DNA/RNA and RNA/DNA primer/template heteroduplexes bind to HIV-RT nearly 2 kcal/mole
tighter at any salt concentration, and binding of a heteroduplex results in the linked release of up
to 8.2 ions. While the affinities of the two homoduplexes are nearly identical, for the
heteroduplexes, RNA/DNA binds to HIV-RT 4-6X tighter at any salt concentration. The
different thermodynamic salt linkages clearly indicate that there are different interaction
footprints for HIV-RT binding to homoduplexes versus heteroduplexes. In addition, the differing
affinities of DNA/RNA versus RNA/DNA indicates that HIV-RT can readily sense whether
RNA is in the primer or the template strand in a heteroduplex, even though the enzyme binds
identically to DNA and RNA homoduplexes.

4.2 Introduction

HIV-reverse transcriptase (HIV-RT) is a multifunctional enzyme that performs a
complex, multi-step reaction to convert the retroviral single-stranded RNA genome into double-
stranded DNA during the HIV life cycle. HIV-RT performs both RNA- and DNA-dependent DNA polymerization, and removes the RNA template of a DNA/RNA hybrid through RNase H activity (52, 53). During the replication process, HIV-RT encounters several different combinations of nucleic acid substrates and must select between polymerization and RNase H activities.

HIV-RT is a heterodimer consisting of two polypeptide chains (p66 and p51). The p51 subunit is derived from p66 by proteolytic cleavage by the viral protease (97, 98). The two subunits have in common fingers, thumb, palm, and connection domains; however, only the p66 subunit has a C-terminal RNase H domain. Both the active DNA polymerase and RNase H domains reside in the p66 subunit, and a cleft runs between two active sites (102-104).

The kinetic preference of HIV-RT for DNA/RNA heteroduplexes over DNA/DNA homoduplexes has been established for some time (148, 149), but little or no equilibrium thermodynamic binding data on the interactions of the protein with different nucleic acids has previously been reported. In this study we utilize a fluorescence anisotropy based direct binding assay to examine the affinities and the thermodynamic salt linkages for HIV-RT binding to 20/27mer (primer/template) DNA/DNA and RNA/RNA homoduplexes and to both DNA/RNA (primer/template) and RNA/DNA heteroduplexes. Because the salt dependence of DNA binding is extremely sensitive to the DNA binding footprint, the thermodynamic linked ion releases upon binding indicate two distinct molecular binding modes for the HIV-RT to the different nucleic acid constructs. Heteroduplexes bind to HIV-RT with approximately 25X higher affinity than do homoduplexes, and RT-heteroduplex binding is linked to the release of approximately 2.5 more ions than is RT-homoduplex binding. These two different thermodynamic signatures directly imply two different nucleic acid binding topologies in solution dependent upon whether the
nucleic acid substrate is a heteroduplex or a homoduplex. Interestingly, these thermodynamic results concur with the recent crystallographic demonstration of a second nucleic acid binding topology by Yang and associates (111), although exact correspondence between the structurally and thermodynamically revealed binding modes cannot be established as yet. It is also notable that the affinity and salt linkage for HIV-RT binding to RNA/RNA and DNA/DNA constructs are nearly identical, and suggests their binding topologies are identical, as little structural or binding information on HIV-RT interactions with RNA/RNA constructs exists.

4.3 Materials and methods

Materials

Fluorescently labeled and unlabeled oligonucleotides were purchased from Integrated DNA Technologies Inc. Stoichiometric DNA binding experiments were performed with the following DNA/DNA-primer/template sets: D/D-20/27, and D/D-63/70. Their sequences are listed in Table 4.1. All DNA and RNA primer/template constructs were labeled at the 5’ end of the primer strand with rhodamine-X (ROX). The binding buffer is 10 mM Tris, 50 mM KCl, 2 mM EDTA, and 1 mM DTT (pH 7.9 at 25 °C).

Equilibrium binding experiments were performed with 20 oligonucleotide long primer and 27 oligonucleotide long template pairs of DNA/DNA (D/D), DNA/RNA (D/R), RNA/DNA (R/D), and RNA/RNA (R/R). All nucleic acid constructs are named in the order: "primer/template". DNA and RNA primer and template sequences are all the same sequences except for the replacement of T by U in RNA primer and template sequences (Table 4.1). Equimolar amounts of primer and template strands were annealed in TE buffer (pH 7) by heating at 95 °C for 5 minutes and gradually returning to room temperature. The buffer conditions for binding are 10 mM Tris, 2 mM EDTA, and 1 mM DTT (pH 7.9 at 25 °C) with varying KCl
concentrations. All buffers were prepared with 1% diethyl pyrocarbonate (Sigma Aldrich) treated water. Premade TE buffer (pH 7) was purchased from Ambion, Life Technologies, and all other chemicals were purchased from Amresco or Sigma Aldrich.

Table 4.1. Nomenclature and sequence of nucleic acids used for binding experiments.

<table>
<thead>
<tr>
<th>Nomenclature</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>D/D-63/70</td>
<td>5'-TACGCAAGCTACATGCTGACTGGGATAACGGTGCCGTTTGGCCAGCTTTCGGAAGCCGTCCA-3'</td>
</tr>
<tr>
<td></td>
<td>3'-ATCGCTGGCAATGCACTGACTGACCTTTGCAAGCTGACGTCCGTTCCCAA-5'</td>
</tr>
<tr>
<td>D/D-20/27, DNA/DNA (D/D)</td>
<td>5'-CCGACCTTTCGAGCGGAGGTCCCAA-3'</td>
</tr>
<tr>
<td></td>
<td>3'-GGCTGAAGCGCGGCAGGTCCCAA-5'</td>
</tr>
<tr>
<td>DNA/RNA (D/R)</td>
<td>5'-CCGACTTTTCGAGCGGTCCA-3'</td>
</tr>
<tr>
<td></td>
<td>3'-GGCUGAAAGCGGCAGGTCCCAA-5'</td>
</tr>
<tr>
<td>RNA/DNA (R/D)</td>
<td>5'-CCGACTTTTCGAGCGGTCCA-3'</td>
</tr>
<tr>
<td></td>
<td>3'-GGCUGAAAGCGGCAGGTCCCAA-5'</td>
</tr>
<tr>
<td>RNA/RNA (R/R)</td>
<td>5'-CCGACTTTTCGAGCGGTCCA-3'</td>
</tr>
<tr>
<td></td>
<td>3'-GGCUGAAAGCGGCAGGTCCCAA-5'</td>
</tr>
</tbody>
</table>

HIV-RT purification

The HIV-1 reverse transcriptase (HIV-RT) expression clone (His-p66/His-p51 RT) was a gift from Stuart Le Grice. Recombinant HIV-RT was expressed and purified based on the previously described procedures (150). The purification steps are described below:

IPTG induced cells were harvested by centrifugation (RCF = 8700 g for 20 minutes at 4 °C), and kept at -80 °C overnight. The cells were resuspended in two volumes (ml/mg) of lysis buffer composed of 50 mM sodium phosphate, 300 mM NaCl, and 10 mM imidazole (pH 7.8). PMSF and lysozyme were added to the resuspended cells with final concentrations of 1 mM PMSF and 4 mg/ml lysozyme. The suspension was stirred for 20 minutes on ice and sonicated 6 X 30 seconds with 30 second cooling periods. The lysate was centrifuged at RCF = 48,400 g for 2 hours at 4 °C. The supernatant was loaded onto the first column, Ni-NTA Superflow (QIAGEN), pre-equilibrated with 50 mM sodium phosphate pH 7.8 buffer containing 300 mM NaCl, 10 mM imidazole, and 10% glycerol. The column was washed with approximately 15 column volumes
of the same buffer, and a subsequent wash with 50 mM sodium phosphate pH 6 buffer containing 300 mM NaCl, 10 mM imidazole, and 10 % glycerol was performed until OD$_{280}$ of the flow-through reached less than 0.01. Then the His-tagged RT was eluted with 50 mM sodium phosphate pH 6 buffer containing 300 mM NaCl, 250 mM imidazole, and 10 % glycerol. The eluted protein was dialyzed overnight against 50 mM sodium phosphate pH 7 buffer containing 150 mM NaCl and 10 % glycerol. The dialyzed sample was then directly loaded onto a heparin-agarose (Thermo Scientific) column, equilibrated in the same buffer as used for the dialysis, and the column was washed with the same buffer until OD$_{280}$ reached less than 0.01. RT was eluted with 50 mM sodium phosphate pH 7 buffer containing 500 mM NaCl and 10 % glycerol. The heparin column eluate was dialyzed overnight against 50 mM sodium phosphate pH 7 buffer containing 25 mM NaCl and 10 % glycerol. The dialyzed sample was loaded onto a UNOsphere S (BIO-RAD) column, equilibrated with the previous dialysis buffer, and followed by a wash with 15 column volumes of the same buffer. RT was eluted with 50 mM sodium phosphate pH 7 buffer containing 500 mM NaCl and 10 % glycerol. The purified RT was stored in 10 mM Tris pH 8 buffer containing 50 mM potassium glutamate, 1 mM EDTA, 0.5 mM DTT, and 50 % glycerol at -20 °C. The purity of the protein was determined on a SDS-PAGE gel and achieved over 95 %. For the fluorescence anisotropy assay, RT was dialyzed overnight against an appropriate binding assay buffer before titrations.

**Fluorescence anisotropy assay**

ROX-labeled DNA was titrated with increasing concentrations of protein, and binding was monitored using the anisotropy signal change as protein-substrate complex is formed (54-56, 151). Fluorescence anisotropy measurements were performed at 25 °C using a FluoroMax-4 spectrofluorometer (Horiba Scientific) equipped with an automated polarizer and regulated at the
indicated temperature. The excitation and emission wavelengths were 583 and 605 nm respectively, with 8-nm band-pass and an integration time of 10 s. For the stoichiometric DNA binding titrations, a total of 100 nM DNA (composed of 1 nM fluorescently labeled and 99 nM unlabeled DNA) was titrated with the protein. For the equilibrium binding titrations, 1 nM fluorescently labeled primer/template was titrated with HIV-RT.

Data Analysis

Stoichiometric binding curves were fit to the equation,

$$
\Delta A = \Delta A_T/2D_T\left\{\left(E_T + D_T + K_d\right) - \left(E_T + D_T + K_d\right)^2 - 4E_TD_T\right\}^{1/2} + (a + bE_T)
$$

(Equation 1)

where $\Delta A$ is the change in anisotropy, $\Delta A_T$ is the total anisotropy change, $E_T$ is the total RT concentration at each point in the titration, $D_T$ is the total DNA concentration, $K_d$ is the dissociation constant (152, 153), $a$ is the y-intercept, and $b$ is the slope of the modified x-axis from the original x-axis. When it is used to analyze stoichiometric binding curves, $D_T$, $a$, and $b$ were allowed to vary during the non-linear regression, and $K_d$ was fixed to 2 nM, which was obtained from the salt linkage data at 50 mM KCl for D/D-20/27 substrate. The binding stoichiometry was determined as the ratio of the fitted $D_T$ and the known $D_T$.

Equilibrium binding curves were fit to the following equation composed of a standard single site isotherm usable when [primer/template] $<<$ $K_d$, and a linear plateau drift expression.

$$
\Delta A = \left\{\left(\Delta A_T(E_T/K_d)\right)/(1+E_T/K_d)\right\} + (a + bE_T)
$$

(Equation 2)

where $\Delta A$ is the change in anisotropy, $\Delta A_T$ is the total anisotropy change, $E_T$ is the total RT concentration at each point in the titration, $K_d$ is the dissociation constant, $a$ is the y-intercept, and $b$ is the slope of the plateau region. At high binding saturation (above 90 %), some protein begins to precipitate out of solution and interferes with the anisotropy signals by causing the signals to increase due to scattering effects. As a result, both types of binding curves
(stoichiometric and equilibrium) display slowly increasing plateaus. Therefore, the binding curves were fit to the above equations, Equation 1 and Equation 2, both of which contain addition of a linear plateau drift expression in order to compensate for this problem. In effect, this operation makes X and Y axes tilt leftward so that ascending plateaus become parallel to the adjusted X axis. The fits of the equilibrium titration data up to 90% saturation (~10X above the Kd value) to Equation 2 without the linear extension \( \Delta A = [(\Delta A_T(E_T/K_d)) / (1+E_T/K_d)] \) were essentially same (within the error ranges) to the Kd values obtained from fitting all of the data to Equation 2. In other words, unlike other Pol I polymerase binding titrations (54-58), binding titrations with HIV-RT often do not completely plateau at high protein concentrations, but continue to drift upward after saturation. To compensate for this we modified the standard binding isotherm by adding a linear drift expression (Equation 2). Fits to full data sets fit with the linear drift expression were compared to fits performed with the standard binding equation on data sets truncated at the beginning of the sloped plateau, and Kd values performed with either procedure were identical within error.

Linked ion release upon binding of RT to different primer/template sets was calculated using a basic linkage relationship (154-157):

\[
\frac{\partial \ln \frac{1}{K_d}}{\partial \ln [KCl]} = \Delta n_{\text{ions}} = \Delta nK^+ + \Delta nCl^- \quad \text{(Equation 3)}
\]

Thus, the slope of a plot of \( \ln \frac{1}{Kd} \) versus \( \ln [KCl] \) will be equivalent to the net number of ions \( \Delta n_{\text{ions}} \) that are bound or released when RT complexes with the primer/template substrate, where \( \Delta n_{\text{ions}} \) is the net sum of the binding and release of both the anions and cations \( \Delta nK^+ + \Delta nCl^- \) (154-156). All non-linear fitting was performed using the program KaleidaGraph (Synergy Software).
4.4 Results

Binding stoichiometry of HIV-RT

Because a variety of construct lengths have previously been used to examine HIV-RT nucleic acid interactions (148, 158-161), we performed stoichiometric titrations with several construct lengths to insure that our binding titrations could be done with 1:1 stoichiometry. Stoichiometric titrations of HIV-RT versus primer/template DNA/DNA were performed at 50 mM KCl and the results are shown in Figure 4.1. HIV-RT was titrated into DNA at high concentrations of DNA ([DNA] >> Kd). The anisotropy of the DNA increases as protein binds due to the decreasing rate of molecular rotation of the DNA in the complex. The fitted binding stoichiometries are 1.00 for D/D\textsubscript{-20/27} and 2.39 for D/D\textsubscript{-63/70}. Shorter (13/20mer) constructs also yielded 1:1 stoichiometry, but had significantly reduced affinity, indicating that they were shorter than the optimal binding length (data not shown). The 20/27mer length is slightly longer than the length of the nucleic acids that are shown to span the full-length of the binding interface of the molecule in the DNA/RNA co-crystal structures (110, 111).

HIV-RT binding to different nucleic acid substrates

Direct binding of HIV-RT to DNA/DNA, RNA/RNA, DNA/RNA, and RNA/DNA was measured across a range of different salt concentrations. All primer/template substrates are composed of a 20 oligonucleotide primer and a 27 oligonucleotide template with a 7 nucleotide overhang at the 5’ end. Nomenclatures and sequences of the nucleic acid substrates are listed in Table 4.1. The sequences of the different primer/template pairs are exactly equivalent, and are different from either the HIV polypurine tract (PPT) or the HIV primer binding site (PBS) sequences, and are the same as primer-template constructs previously used by our laboratory to characterize binding by other Pol I polymerases (54-58). Because we are using wild type HIV-
RT, no magnesium salts are included in any of the buffers in the current study in order to prevent
degradation of the RNA substrates by the HIV RNase H, and indeed, no degradation of any of
the substrates was observed in any of the experiments. These control experiments were done by
monitoring the anisotropy change of the Rox-labeled primer-template substrate incubated with
Kd concentration of RT over two hours. In the course of two hour incubation, no anisotropy
change was observed in the absence of Mg$^{2+}$. On the other hand, a large anisotropy drop was
observed in the presence of Mg$^{2+}$. Figure 4.2 shows representative fluorescence anisotropic
titrations of the binding of HIV-RT to RNA/RNA at several different KCl concentrations at 25
°C. Binding affinities (Kd values) were determined by fitting the data to Eq. 2 as described
under “Data Analysis,” and data for all constructs are listed in Table 4.2.

Figure 4.1. Determination of binding stoichiometry for HIV-RT binding to two different lengths
of DNA/DNA-primer/template. Shown are fluorescence anisotropy-monitored stoichiometric
titrations of HIV-RT into 100 nM DNA. HIV-RT binding to D/D-20/27 and D/D-63/70 are
shown as closed circles and closed squares, respectively. The titrations were performed in 10
mM Tris, 50 mM KCl, 2 mM EDTA, 1 mM DTT, pH 7.9, at 25 °C. Lines are the fit to Eq. 1
(Materials and methods).
Figure 4.2. Representative equilibrium binding isotherms for the binding of HIV-RT to RNA/RNA-primer/template at various KCl concentrations at 25 °C. Binding was monitored using fluorescence anisotropy. KCl concentrations used for RNA/RNA binding are listed in the legend. Binding affinity decreases (left to right) as salt concentration increases. Data are fit to Equation 2 (Materials and methods).

Table 4.2. KCl dependence of nucleic acids binding by HIV-RT.

<table>
<thead>
<tr>
<th>KCl (mM)</th>
<th>Kd (nM)</th>
<th>Ions released&lt;sup&gt;a&lt;/sup&gt;</th>
<th>ΔG at 1M KCl (kcal/mol)</th>
<th>KCl (mM)</th>
<th>Kd (nM)</th>
<th>Ions released&lt;sup&gt;a&lt;/sup&gt;</th>
<th>ΔG at 1M KCl (kcal/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>D/D</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>63</td>
<td>7.7 ± 0.6</td>
<td></td>
<td></td>
<td>113</td>
<td>7.5 ± 0.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>75</td>
<td>15.6 ± 1.2</td>
<td></td>
<td></td>
<td>125</td>
<td>19.9 ± 1.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>88</td>
<td>44.8 ± 2.6</td>
<td>5.7 ± 0.3</td>
<td>−1.8 ± 0.5</td>
<td>138</td>
<td>30.3 ± 2.0</td>
<td>7.8 ± 0.5</td>
<td>−1.0 ± 0.6</td>
</tr>
<tr>
<td>100</td>
<td>91.1 ± 3.8</td>
<td></td>
<td></td>
<td>150</td>
<td>74.7 ± 7.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>113</td>
<td>209.1 ± 13.1</td>
<td></td>
<td></td>
<td>163</td>
<td>140.0 ± 11.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R/R</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>63</td>
<td>7.3 ± 0.9</td>
<td></td>
<td></td>
<td>138</td>
<td>7.2 ± 0.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>75</td>
<td>14.0 ± 1.0</td>
<td></td>
<td></td>
<td>150</td>
<td>11.9 ± 0.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>88</td>
<td>32.4 ± 2.0</td>
<td>5.6 ± 0.4</td>
<td>−2.0 ± 0.6</td>
<td>163</td>
<td>25.1 ± 1.4</td>
<td>8.2 ± 0.5</td>
<td>−1.5 ± 0.5</td>
</tr>
<tr>
<td>100</td>
<td>90.8 ± 2.2</td>
<td></td>
<td></td>
<td>175</td>
<td>45.9 ± 3.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>113</td>
<td>184.3 ± 10.5</td>
<td></td>
<td></td>
<td>183</td>
<td>104.4 ± 13.1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup> Values determined from the slope of the $\partial \ln Kd$ versus $\partial \ln [\text{KCl}]$ linkage.

Errors are the parameter value errors returned from the fits to Equation 2 (Kd values) or Equation 3 (number of ion release and ΔG at 1 M KCl).
The results show that binding of HIV-RT to the DNA/DNA and RNA/RNA homoduplex primer/templates is nearly identical, despite the fact that there are differences in geometry of the substrates (B-form versus A-form) and the presence or absence of -2’OH on the sugar molecule (deoxyribose versus ribose). On the other hand, similar binding affinities of HIV-RT to heteroduplexes are displaced higher salt ranges, meaning that at any single specific salt concentration the binding of the hybrid duplexes is tighter than that of the homoduplex substrates. Figure 4.3 shows the equilibrium binding of DNA/DNA, RNA/RNA, and DNA/RNA at the same salt concentration (112.5 mM KCl). At this salt concentration, HIV-RT binding to DNA/RNA is 25-28x tighter (ΔΔG ~ 2.0 kcal/mol tighter) than binding to DNA/DNA or RNA/RNA.

Figure 4.3. The binding of HIV-RT to DNA/DNA, RNA/RNA, and DNA/RNA. Equilibrium titrations are shown at 112.5 mM KCl. HIV-RT binding to DNA/RNA, DNA/DNA, and RNA/RNA are shown as closed diamonds, closed circles, and closed squares, respectively. The titrations were performed in 10 mM Tris, 112.5 mM KCl, 2 mM EDTA, 1 mM DTT, pH 7.9, at 25 °C. Data are fit to Eq. 2 (Materials and methods).

Although HIV-RT binds DNA/DNA and RNA/RNA homoduplexes with nearly identical affinities, the affinities for DNA/RNA vs. RNA/DNA are quite different. At the three KCl
concentrations where both hybrids were directly measured (137.5 mM, 150 mM, and 162.5 mM), RNA/DNA binding is ~4-6× ($\Delta\Delta G = 0.9$-1.1 kcal/mol) tighter than DNA/RNA binding. Figure 4.4 shows side-by-side equilibrium titrations of DNA/RNA and RNA/DNA at 150 mM KCl where the binding differs by ~6× ($\Delta\Delta G = 1.1$ kcal/mol tighter binding to RNA/DNA than to DNA/RNA).

Figure 4.4. The binding of HIV-RT to DNA/RNA and RNA/DNA. Equilibrium titrations are shown at 150mM KCl. HIV-RT binding to DNA/RNA and RNA/DNA are shown as closed diamonds and closed triangles, respectively. The titrations were performed in 10mM Tris, 150mM KCl, 2mM EDTA, 1mM DTT, pH 7.9, at 25 °C. Data are fit to Eq. 2 (Materials and Methods).

**KCl linkages for the binding of HIV-RT to different nucleic acids**

Figure 4.5 shows thermodynamic linkage plots for nucleic acids binding as a function of KCl concentration ($\partial \ln K_d$ versus $\partial \ln [KCl]$). The negative slopes of the linkage plots are indicative of net ion release upon formation of the protein-nucleic acid complex (54, 57, 58, 154-157). As observed for Pol I DNA polymerases (54, 57, 58), nucleic acid binding by HIV-RT is linked to ion release for all the primer/template duplexes, but differs significantly between
homoduplexes and heteroduplexes. The exact linked ion releases are obtained from the slopes of the linkage plots and are reported in Table 4.2. HIV-RT binding to either homoduplex releases the same number of linked ions (5.7 ions for DNA/DNA and 5.6 ions for RNA/RNA), and the two linkage plots essentially overlap. The linkage plots for the heteroduplexes are nearly parallel, both showing much higher ion releases (7.8 – 8.2) upon binding of HIV-RT than do the homoduplexes. The heteroduplex linkages are both displaced to higher salt ranges, and at any particular salt concentration it can be seen that the binding affinity relationship will be DNA/DNA ≈ RNA/RNA < DNA/RNA < RNA/DNA. Extrapolation of the linkage plots to 1 M salt provides an estimate of the non-electrostatic components of the binding free energy (154-157) which is within error in the range of -1.0 to -2.0 kcal/mol for all of the duplexes (Table 4.2).

Figure 4.5. KCl linkages (δln1/Kd versus δln[KCl]) for primer/template binding of different nucleic acids by HIV-RT. The slopes of the plots give the thermodynamic net average number of ions released upon complex formation. As also indicated in the figure inset, the symbols for the different substrates are: closed circles for DNA/DNA, open squares for RNA/RNA, closed diamonds for DNA/RNA, and open triangles for RNA/DNA.
4.5 Discussion

The two catalytic sites of HIV-RT are known to accommodate binding of various DNA-RNA combinations during retroviral replication, and a number of different researchers have characterized numerous aspects of these interactions (96, 148, 149, 158-161). However, all previous estimates of binding affinities for the different nucleic acids have been extrapolations from different types of activity assays, an approach that is fraught with caveats, thus herein we performed direct equilibrium substrate binding experiments for different pairs of nucleic acids (DNA/DNA, RNA/RNA, DNA/RNA, and RNA/DNA) using a fluorescence anisotropy assay. Because all nucleic acid-protein interactions are salt sensitive, and because quantification of the salt dependence directly reflects properties of the binding footprint, we also characterized the thermodynamic salt linkages for each of the different duplex combinations, finding that these dependencies indicate different binding topologies for homoduplex versus heteroduplex nucleic acids.

In the current study we utilized a generic/random nucleic acid sequence, previously used to study the DNA binding of other Pol I polymerases, and avoided the HIV primer-binding site (PBS) and polypurine tract (PPT) sequences, as these genomic domains are likely to exhibit sequence dependent binding effects, since that is their physiological function. This approach allows us to ask, in this particular study, the more general question of structural selectivity of HIV-RT for DNA vs. RNA, rather than sequence selectivity. We utilized a 20/27mer construct as a 1:1 binding substrate to avoid multiple binding or end binding of HIV-RT, since HIV-RT is also capable of binding to blunt-ended DNA (162). Crystal structures of HIV-RT complexed with duplex DNA show that an 18-20 base-paired duplex region and 1-5 nucleotides of the 5’-
overhang of the template are visible (107-109, 163, 164), and our binding results confirmed that the DNA/DNA-20/27 binding stoichiometry is 1.00.

**Distinguishing homoduplexes from heteroduplexes / two binding modes for HIV-RT on nucleic acids**

Our results show that for the substrates examined, the binding hierarchy is RNA/DNA > DNA/RNA > DNA/DNA ≈ RNA/RNA from the tightest to the weakest binding (Table 4.2). The binding affinity for heteroduplex nucleic acids is over 25X tighter (ΔΔG > 1.9 kcal/mol) than the affinity for homoduplexes of DNA or RNA, and HIV-RT binding to heteroduplexes releases ~8 ions while homoduplex binding releases ~5.6 ions. Ion release upon complex formation is a direct reflection of the DNA binding footprint and often reflects the number of DNA phosphates buried in the binding interface. Typical ion releases can range from about 2-12 ions (54, 154-157), depending on the DNA binding protein, and a 2.5 ion change indicates a rather significant change in the binding footprint. If the ion releases primarily reflect phosphate burial, which is often but not always the case, the results would indicate the burial of approximately 3 additional phosphates upon heteroduplex binding of HIV-RT relative to homoduplex binding.

By measuring kinetic off-rates, DeStefano and associates first posited that HIV-RT binds more tightly to DNA/RNA heteroduplexes than to DNA/DNA duplexes (148), and through an elegant series of experiments have shown that relative to a DNA/DNA duplex, HIV-RT will dissociate more slowly from a construct that has even only a small number of RNA bases near the 3' end of the primer (148, 158, 159, 161). The thermodynamic results presented herein corroborate the kinetic conclusions, and provide the first precise Kd and free energy values for the binding of HIV-RT to several different nucleic acid constructs. It should be noted that using primer extension assays, DeStefano and associates also previously estimated what they acknowledge as "apparent Kd" values for HIV-RT binding, but that those estimates completely
disagree with both their own kinetically established findings (148, 158-161) and with the thermodynamic data in this study, and appear to be an example of the fact that estimating Kd values from activity data is often highly problematic. It should also be noted, however, that to avoid RNA degradation by the wild type HIV-RT, no magnesium salts are included in any of the direct binding experiments herein, while activity assays clearly require Mg$^{2+}$. Because the thermodynamic results herein qualitatively agree with the majority of earlier findings that HIV-RT prefers DNA/RNA over DNA/DNA, it seems unlikely that these buffer differences will dramatically alter the Kd values, but it is a formal possibility.

All of the four substrates examined herein are physiological substrates for HIV-RT. Initial tRNA priming of the ssRNA genome presents the reverse transcriptase with an RNA/RNA primer/template. The extension of the initial DNA copy of the genome presents a DNA/RNA primer/template substrate, as the processivity of HIV-RT averages approximately 50-100 nucleotides (but can range up to ~300) (96), thus, even at the high end of its range the enzyme will dissociate and rebind a number of times while making the first cDNA strand. It is the hybrid duplex; however, that is also the substrate for the polymerization-dependent and independent RNaseH activity of the enzyme. During cDNA synthesis, secondary structures in the RNA template such as hairpins can cause pausing of the polymerization activity, and such pauses promote RNase H cleavages. The mode of RNA cleavage is known as “DNA 3’ end-directed cleavage”. The RNA/DNA primer/template used here presents a substrate for the polymerization-independent RNase H activity, which is only a substrate for the RNase H activity of the enzyme since the RNA sequence used here is a generic, non-PPT sequence. This synthesis-independent RNase H activity participates in removal of genomic RNA from the DNA-RNA hybrid but retains the polypurine tract (PPT) primer for dsDNA synthesis. When the
RNA/DNA primer template substrate interacts with HIV-RT, its blunt end would be placed close to the polymerization active site, and its primer/template junction side would be placed closer to the RNase H active site, and this binding mode is considered the RNA 5’ end–directed (or DNA 3’ end-directed) cleavage mode (106). This orientation preference of HIV-RT binding to non-PPT RNA primed to DNA template was previously shown in a smFRET study (165). This is the only primer/template substrate used in this study to orient in this manner. The other primer/template substrates (RNA/RNA, DNA/RNA, and DNA/DNA) should bind to HIV-RT in an orientation placing the primer/template junction near the polymerization active site and the blunt end near the RNase H active site of the enzyme. A DNA/DNA primer/template junction is then the substrate the enzyme encounters while making the initial cDNA into dsDNA. The nearly identical affinities of the DNA/DNA and RNA/RNA substrates thus makes physiological sense as the enzyme should only interact with these substrates in polymerization mode, and never in RNase H mode. The DNA/RNA hybrid duplex, on the other hand, interacts with both active sites on the enzyme, and it is likely the additional interactions with the RNaseH active site that lead to the altered thermodynamic salt linkages that are observed (see also the discussion below about correlations between the thermodynamic data and available crystal structures). It is somewhat surprising that the RNA/DNA substrate, which is only a substrate for the RNase H activity of the enzyme, binds to HIV-RT with the highest affinity: an average of 5X tighter than the DNA/RNA heteroduplex. This observation might indicate that the binding of the DNA/RNA primer/template to HIV-RT might be fluctuating between polymerization mode and RNase H compatible mode, or that the 5’ overhang of the template strand in RNA/DNA substrates coming out from the RNase H active site might interact with another part of HIV-RT (besides the nucleic binding cleft), resulting in stronger stabilization of the binary complex.
Comparisons with other Pol I polymerases

Pol I DNA polymerases and HIV-RT have substantial differences in their overall molecular properties. However, there are equally striking similarities between the polymerase domains of the two enzymes. In both polymerases, the polymerase domain is subdivided into three subdomains which can be described as the fingers, palm, and thumb of a right hand. These subdomains together contribute to proper binding of nucleic acid at the polymerase active site (19, 107-109, 163, 164). When the salt dependence of HIV-RT binding is compared with that of other Pol I polymerases, it appears that electrostatic forces play a much more prominent role in the binding of HIV-RT. In comparison to the 5.6 to 8.2 ions released upon binding of HIV-RT, DNA binding by either full-length Taq DNA polymerase or its Klentaq large-fragment domain is linked to the release of only ~3 ions. When E. coli’s Klenow polymerase binds DNA, ~4.5 ions are released (54). Differences are likely due to HIV-RT’s more elongated binding surface which interacts with a longer stretch of nucleic acid. More striking, however, are comparisons between the relative electrostatic versus non-electrostatic binding energies.

Extrapolation of salt linkage plots to 1 M salt provides an estimate of the non-electrostatic components of the binding interactions (154-157). Non-electrostatic binding free energies of Taq, Klentaq, and Klenow range from about -5 to -8 kcal/mole, which is ≥ 50 % of their binding free energies at 100 mM KCl (54). In contrast, the -1 to -2 kcal/mole non-electrostatic binding free energies for HIV-RT represent < 20 % of their binding energies at 100 mM salt. This comparison indicates that the balance of non-covalent forces driving the binding of HIV-RT to nucleic acids is much more electrostatic in nature, and more focused on the phosphate backbone of the nucleic acids, than the binding of Taq or E. coli Pol I type DNA polymerases.
Comparison with Crystal Structures

A number of co-crystal structures of HIV-RT complexed with DNA/DNA exist (102, 107-109, 163, 164), along with two different co-crystal structures of HIV-RT bound to DNA/RNA (110, 111). How might the thermodynamic results correlate with the known structural details? The structures of HIV-RT complexed with a DNA/DNA substrate show that the 3’ end of the primer is bound at the polymerase catalytic center (Figure 4.6; A) (102, 107-109). The duplex near the polymerase active site adopts A-form, there is a kink adjacent to the p66 thumb, and the DNA has B-form geometry near the RNase H active site (Figure 4.6; A and D) (102, 107-109). The most numerous nucleic acid interactions with protein occur primarily along the sugar-phosphate backbone of the DNA, which is in accord with the high proportion of electrostatic binding free energy that stabilize the complexes. Although no structure of HIV-RT bound to RNA/RNA yet exists, the near identical thermodynamics of DNA/DNA and RNA/RNA binding predict that the structure of an RNA/RNA-HIV-RT co-crystal should be found to very similar to the DNA/DNA structures, despite the fact that RNA/RNA adopts an A-form helix (166). A biochemical study showed that the RNase H of HIV-RT does not cleave an RNA/RNA substrate (167), and a mass spectroscopic protein footprinting study showed that interactions of HIV-RT with tRNA<sup>125s3</sup> hybridized to the genomic PBS sequence occur within the primer-template binding cleft (168).

Two crystal structures of DNA/RNA-HIV-RT co-crystals exist, and show two quite different binding topologies (Figure 4.6; B and C). The structure of HIV-RT complexed with DNA/RNA containing a polypurine tract (PPT) shows an overall geometry of the heteroduplex that is quite similar in conformation to the bound DNA/DNA duplex, although the HIV-RT has numerous additional interactions with 2’-OH groups of the RNA template in resulting more
extensive contacts between HIV-RT and DNA/RNA versus DNA/DNA (Figure 4.6; B and E) (110). This DNA/RNA structure, like the DNA/DNA structures, shows the nucleic acid binding in the polymerase active site. The recently solved structure of HIV-RT containing an NNRTI and a non-PPT DNA/RNA sequence differs structurally from all previously reported HIV-RT-nucleic acid complexes and is compatible with RNA cleavage (110, 111). The bound heteroduplex geometry is mostly in the A-form and has an additional kink before entering the RNase H domain. The 3’ end of the DNA primer reaches into the palm but veers ~5 Å away from the polymerase active site (111). In addition, both p66 and p51 also showed conformational changes resulting in novel interactions with the heteroduplex (Figure 4.6; C and F) (111).

It is not immediately obvious which, if either, DNA/RNA co-crystal structure most directly corresponds with the differences in salt linkage found in solution, however, the PPT DNA/RNA structure (1HYS) (Figure 4.6; B and E), despite having the nucleic acid in a highly similar topology as in DNA/DNA complexes, does include additional protein-nucleic acid interactions, including interactions with at least two additional template phosphates (110), which would correlate quite well with the measured higher ion release upon binding in solution. In contrast, the RNase H compatible DNA/RNA structure (4B3O) (Figure 4.6; C and F), while displaying a significantly altered binding footprint on the nucleic acid which could alter the linked ion release, also includes a bound non-nucleoside reverse transcriptase inhibitor (NNRTI) in order to effectively block the nucleic acid from binding in polymerization mode (111). Additional thermodynamic linkage experiments in the presence of NNRTIs should prove informative about the relationships between these two different DNA/RNA-HIV-RT complexes.
Figure 4.6. Crystal structures of HIV-RT-nucleic acid complex showing the polymerization and RNase H compatible modes of nucleic acid binding. The p66 and p51 subunits are colored green and lime green, respectively. The DNA primer, DNA template, and RNA template are colored purple, blue, and red, respectively. Both polymerase catalytic residues (Asp 110, Asp 185, and Asp 186) and the RNase H catalytic residues (Glu 478, Asp 443, Asp 498, and Asp 549) are shown as spheres and colored orange and magenta, respectively. Panel A (PDB: 1R0A) shows the polymerization mode of DNA/DNA binding by HIV-RT (164). Panel B (PDB: 1HYS) shows the polymerization mode of DNA/RNA binding by HIV-RT (110). Panel C shows the RNase H compatible mode of DNA/RNA binding by HIV-RT (111). Panel C, D, and E showing the interaction of the nucleic acid and the p66 subunit of HIV-RT are created by removing the p51 subunit and turning 90° angle (clockwise) along the Z-axis from Panel A, B, and C, respectively.

4.6 Acknowledgements

We thank Stuart Le Grice and Jenny Miller for the HIV-RT expression clone and advice on expression and purification. We thank Chantal San Miguel for technical assistance. This work was supported by the National Science Foundation.
CHAPTER 5
CONCLUSIONS

In this Chapter, molecular models for potentially explaining the thermodynamic results obtained in this study are discussed.

5.1 Contribution of the negative enthalpy of DNA binding to the catalytic activity of Klenow and Klentaq polymerases

In Chapter 2, I showed that the nucleotide incorporation activity of Klenow and Klentaq was only observed when $\Delta H$ of DNA binding is negative. This raises the question of how the negative enthalpy of substrate binding contributes to the enzymatic activity.

The $\Delta H$ is positive in endothermic reactions which absorb heat, and negative in heat-releasing exothermic processes. This study showed the enzymatic activities of both polymerases were enhanced when DNA binding was exothermic. How could this heat released by DNA binding be used by the polymerases to become catalytically competent?

In this scenario, I hypothesize that the heat deposited near the catalytic site would foster local heat transfer to the atoms of the catalytic residues, metal ions, water molecules, dNTP, and pt-DNA. As a result, these atoms might gain vibrational energy and wiggle more. The increased wiggling would increase the chance of attaining the proper transition state orientation (shown in Figure 1.3), thus, enhancing the catalytic rate. The percent change in the average temperature factor of the atoms of the three catalytic residues of Klentaq (Asp 610, Asp 785, and Glu 786 (21)) in the crystals of without (PDB: 1KTQ (73)) and with DNA bound (PDB: 4KTQ (169)) has shown to be $\sim$60 $\%$, which indicates that these catalytic residues become more flexible upon complex formation.

Or this heat energy could be used to speed up the conformational changes that happen after DNA/dNTP substrate binding (Step 3 in Figure 1.2) and before the chemical reaction
occurs. Such a conformational change has been observed crystallographically with Klentaq polymerase (19) where the “fingers” subdomain moves closer to the polymerase catalytic site in the “palm” subdomain to help the incoming dNTP base-pair with the template base. Tyr 671 in Klentaq, at the bottom of the O helix in the “finger” subdomain, has been shown to occupy the insertion site where the next template base will base pair with the incoming dNTP. The O helix movement is associated with the side chain of this tyrosine residue moving away from the insertion site. This tyrosine is conserved in the Type I DNA polymerase family (170), Hence, the heat released upon DNA binding near this residue might make this residue more flexible and help relocate the side chain of Tyr 671 away from the insertion site so that the catalytically competent conformation could form.

Another possibility is that the heat released upon DNA binding might change the strength or orientation of the non-covalent bonding network between DNA and the polymerases. Some non-covalent bonds might become stronger, and some might become weaker among the forty amino acid residues that have shown to interact with DNA (19). Such changes in the non-covalent bonding network associated with the change in ∆H favor strengthening of catalytically important bonds as temperature increases. Such changes in bond strength would not necessarily alter the active site structure, but could affect the catalytic rate. The negative enthalpy of substrate binding might have an effect on some combination of any or all of the possibilities listed here.

5.2 Networking of the polymerization and proofreading sites in Klenow for DNA replication fidelity

In chapter 3 of this dissertation, the primer binding of matched and single terminal mismatched pt-DNA to these active sites of D424A Klenow exo minus was shown to be time dependent in the presence of divalent cations. In the case of the presence of magnesium ions,
which are cofactors for both polymerization and exonuclease, the primer strand of both matched and single mismatched pt-DNA initially partitions more at the polymerization domain, and then some strands shift to the exonuclease/proofreading domain.

A previous study showed that Klenow binding to the three terminal mismatched pt-DNA is 2-3 times tighter than the correctly matched pt-DNA (59). As shown in Table 3.2, the perfectly matched pt-DNA binding is almost completely maintained in the polymerization domain, and the three mismatched pt-DNA binding is in the proofreading domain (41). In this section I describe possible models for 1) why Klenow binds mismatched DNA more tightly, and 2) why the DNA moves after initial binding.

**Why does Klenow bind mismatched DNA more tightly?**

DNA polymerases, including Klenow, prevent mismatched base incorporation by reducing the rate of the phosphodiester bond formation. However, if it does happen, then they also reduce the rate of mismatch extension (76). The exonuclease reaction rate is slow compared to the incorporation rate of a correctly matched nucleotide (13), but the rate of incorporation onto a primer strand with a terminal single mismatched base (mismatch extension) is slower. Hence, the slow kinetic mechanism of the mismatch extension would allow more time for the exonuclease activity to excise the incorrectly incorporated base from the 3’ end of the primer strand (22). Taking the above information into account, I can speculate that the reason why Klenow has a proofreading site with a tighter DNA binding affinity than the polymerization site is that in order to compensate the slow exonuclease reaction, it is necessary for Klenow to hold DNA more tightly at the proofreading site.
Why does the primer strand move after initial binding?

The time dependence of the primer partitioning between the polymerization site and the proofreading site in the presence of magnesium ions showed a shift from the polymerization site to the proofreading site after initial binding (Figure 3.3; A). A prior mutational study has suggested that a mismatched pt-DNA is first recognized within the polymerization domain of Klenow (171). The primer strand of the stalled (not elongating) single mismatched pt-DNA in the polymerization domain is then transferred to the proofreading site intramolecularly (without dissociating from DNA) (35). The data in this dissertation appear to document the kinetics of this intramolecular transfer.

In the single terminal mismatched pt-DNA, two more base pairs still need to be unwound in order to bind at the proofreading site (17, 41). The time dependence of the binding shift observed in this study might indicate that the polymerization mode of DNA binding somewhat destabilizes base paring near the primer-template junction. Then the primer strand is transferred to the proofreading site. Such an opportunity might depend on thermodynamic stability of the hydrogen bonding between the base-paired bases next to the terminal mismatch base pair. Base substitution mutations do tend to occur more frequently in more stable (GC rich) regions and less frequently in less stable (AT rich) regions (145, 146). The weakening of the hydrogen bonding might be achieved by interaction of the fraying primer with some amino acid residues in the polymerization domain, and those favorable or unfavorable interactions might trigger the primer strand transfer to the proofreading site. Such amino acid residues in the polymerization domain may even have been previously identified (35, 171). Mutation in those residues changes the percent partitioning of the primer strand between the two active sites of Klenow (35, 171).
The time dependence of the shift was observed only when divalent ions were present. Metal ions bound at the exonuclease active site increase the partitioning of the primer strand at the site. Also, the presence of metal ions in the bulk solution might more effectively destabilize the bonding to the polymerization site, and this destabilization might be the slow kinetic process that was observed in this study.

**Klenow molecular strategies for the proofreading activity**

This study and previous studies together have illustrated strategies that are used to excise a misincorporated base from the primer strand by Klenow. If an incorrect base is incorporated, the polymerase activity stalls, and some of the amino acid residues in the polymerization domain recognize the misincorporation which triggers sending the mismatch primer to the proofreading site intramolecularly, establishing a new pol:exo site equilibrium. Now this fixed pt-DNA is a substrate for polymerase activity. Because the rate of the 3’-5’ exonuclease activity is slow, it will not remove bases processively (144). This slow rate of the exonuclease activity prevents unnecessary nucleotide removals. These two kinetically and thermodynamically unbalanced polymerization and exonuclease binding reactions and catalytic activities work together to make faithful replication happen.

**5.3 Two different modes of nucleic acid binding of HIV-RT**

HIV-RT has also two different catalytic activities: DNA polymerization and RNase H activity. The task of HIV-RT is to convert single-stranded genomic viral RNA into double-stranded DNA so that the viral integrase may insert the double-stranded DNA into host genomic DNA. After integration, the viral genes are transcribed into mRNA and translated into viral protein particles by the host transcriptional and translational machinery. Then those viral proteins
and two copies of mRNA assemble together, bud off from the host membrane, and mature into an infectious virion (53).

In Chapter 4 of this study, it is shown that HIV-RT can sense differences in nucleic acid composition. In other words, HIV-RT has at least two different binding modes; one is for homoduplex (DNA/DNA and RNA/RNA) binding, and the other is for heteroduplex (DNA/RNA and RNA/DNA) binding. The heteroduplex binding to HIV-RT is tighter and releases more ions than the homoduplex binding to HIV-RT (Table 4.2). In this section I describe possible models for 1) why HIV-RT binds heteroduplexes more tightly, 2) why HIV-RT binds RNA/RNA the same as DNA/DNA, and 3) why HIV-RT does not bind RNA/RNA and DNA/RNA equivalently.

**Why does HIV-RT bind heteroduplexes more tightly?**

I hypothesize that the tighter bind mode seen in thermodynamic studies in Chapter 4 indicates the RNase H mode of heteroduplex binding based on the following reasoning: First, no dNTP was included in the assay buffer so that the heteroduplex is a primary substrate for the RNase H activity (106). Second, the sequence of our primer/template heteroduplexes is different from the polypurine tract (PPT) sequence which has been show to exhibit an unusual heterduplex conformation and which does bind in polymerization mode (110). Third, a smFRET study showed that HIV-RT can sense the difference between DNA and RNA strands and locate the RNA strand close to the RNase H active site, which agrees with our results (165).

**Why does HIV-RT bind DNA/DNA and RNA/RNA equivalently?**

The fact that DNA/DNA and RNA/RNA bind HIV-RT with identical thermodynamics suggests that they bind with similar topologies, yet RNA is A-form (166). The binary complex of HIV-RT and DNA homoduplex shows that the major interaction between the molecules happens within the polymerization domain where the DNA duplex adopts A-form (102, 107-109, 163,
The DNA sugar-phosphate backbone interacts with amino acid residues in the polymerization domain through hydrogen bonding and van der Waals interaction. Hence, it is possible for the RNA/RNA to interact with HIV-RT in the polymerization domain in the same way as the DNA/DNA interacts with HIV-RT.

**Why does HIV-RT bind RNA/RNA and DNA/RNA differently?**

A single HIV virion composes approximately 50 RT molecules (172) with two single-stranded genomic RNA molecules (52, 86). The genomic RNA forms secondary structures, such as hairpins with duplex stems (173). Hence, it is important for HIV-RT not to cleave its own genomic RNA. Although an RNA/RNA homoduplex adopts a similar configuration as a DNA/RNA heteroduplex, which is A-form, the RNA/RNA does not bind in the RNase H mode.

The crystal structure of the recent RNase H mode HIV-RT-heteroduplex complex shows that there are conformational changes in the connection domains of both subunits and the p51 “thumb”, resulting in new interactions with the DNA strand beyond the polymerization domain (111). Hence, it could be possible that those interactions favor a DNA molecule composed of deoxyribose sugars and disfavors a RNA molecule composed of ribose sugars. If so, RNA/RNA homoduplex binding would be biased toward the interaction within the polymerization domain. I.e. it may be the DNA part of the heteroduplex that allows binding of the heteroduplex in RNase H mode.
REFERENCES


# APPENDIX

## ELSEVIER LICENSE TERMS AND CONDITIONS

Sep 30, 2013

This is a License Agreement between Hiromi S Brown ("You") and Elsevier ("Elsevier") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Elsevier, and the payment terms and conditions.

**All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.**

<table>
<thead>
<tr>
<th>Supplier</th>
<th>Elsevier Limited</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>The Boulevard,Langford Lane</td>
</tr>
<tr>
<td></td>
<td>Kidlington,Oxford,OX5 1GB,UK</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Registered Company Number</th>
<th>1982084</th>
</tr>
</thead>
<tbody>
<tr>
<td>Customer name</td>
<td>Hiromi S Brown</td>
</tr>
<tr>
<td>Customer address</td>
<td>Louisiana State University</td>
</tr>
<tr>
<td></td>
<td>BATON ROUGE, LA 70803</td>
</tr>
</tbody>
</table>

| License number             | 3238800235368              |
| License date               | Sep 30, 2013               |
| Licensed content publisher | Elsevier                  |
| Licensed content publication | Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics |
| Licensed content title     | Enthalpic switch-points and temperature dependencies of DNA binding and nucleotide incorporation by Pol I DNA polymerases |
| Licensed content author    | Hiromi S. Brown, Vince J. LiCata |
| Licensed content date      | October 2013               |
| Licensed content volume number | 1834                     |
| Licensed content issue number | 10                       |
| Number of pages            | 6                         |
| Start Page                 | 2133                      |
| End Page                   | 2138                      |
INTRODUCTION

1. The publisher for this copyrighted material is Elsevier. By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the Billing and Payment terms and conditions established by Copyright Clearance Center, Inc. ("CCC"), at the time that you opened your Rightslink account and that are available at any time at http://myaccount.copyright.com).

GENERAL TERMS

2. Elsevier hereby grants you permission to reproduce the aforementioned material subject to the terms and conditions indicated.

3. Acknowledgement: If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source, permission must also be sought from that source. If such permission is not obtained then that material may not be included in your publication/copies. Suitable acknowledgement to the source must be made, either as a footnote or in a reference list at the end of your publication, as follows:
“Reprinted from Publication title, Vol /edition number, Author(s), Title of article / title of chapter, Pages No., Copyright (Year), with permission from Elsevier [OR APPLICABLE SOCIETY COPYRIGHT OWNER].” Also Lancet special credit - “Reprinted from The Lancet, Vol. number, Author(s), Title of article, Pages No., Copyright (Year), with permission from Elsevier.”

4. Reproduction of this material is confined to the purpose and/or media for which permission is hereby given.

5. Altering/Modifying Material: Not Permitted. However figures and illustrations may be altered/adapted minimally to serve your work. Any other abbreviations, additions, deletions and/or any other alterations shall be made only with prior written authorization of Elsevier Ltd. (Please contact Elsevier at permissions@elsevier.com)

6. If the permission fee for the requested use of our material is waived in this instance, please be advised that your future requests for Elsevier materials may attract a fee.

7. Reservation of Rights: Publisher reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

8. License Contingent Upon Payment: While you may exercise the rights licensed immediately upon issuance of the license at the end of the licensing process for the transaction, provided that you have disclosed complete and accurate details of your proposed use, no license is finally effective unless and until full payment is received from you (either by publisher or by CCC) as provided in CCC's Billing and Payment terms and conditions. If full payment is not received on a timely basis, then any license preliminarily granted shall be deemed automatically revoked and shall be void as if never granted. Further, in the event that you breach any of these terms and conditions or any of CCC's Billing and Payment terms and conditions, the license is automatically revoked and shall be void as if never granted. Use of materials as described in a revoked license, as well as any use of the materials beyond the scope of an unrevoked license, may constitute copyright infringement and publisher reserves the right to take any and all action to protect its copyright in the materials.

9. Warranties: Publisher makes no representations or warranties with respect to the licensed material.

10. Indemnity: You hereby indemnify and agree to hold harmless publisher and CCC, and their respective officers, directors, employees and agents, from and against any and all claims arising out of your use of the licensed material other than as specifically authorized pursuant to this license.

11. No Transfer of License: This license is personal to you and may not be sublicensed,
assigned, or transferred by you to any other person without publisher's written permission.

12. No Amendment Except in Writing: This license may not be amended except in a writing signed by both parties (or, in the case of publisher, by CCC on publisher's behalf).

13. Objection to Contrary Terms: Publisher hereby objects to any terms contained in any purchase order, acknowledgment, check endorsement or other writing prepared by you, which terms are inconsistent with these terms and conditions or CCC's Billing and Payment terms and conditions. These terms and conditions, together with CCC's Billing and Payment terms and conditions (which are incorporated herein), comprise the entire agreement between you and publisher (and CCC) concerning this licensing transaction. In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall control.

14. Revocation: Elsevier or Copyright Clearance Center may deny the permissions described in this License at their sole discretion, for any reason or no reason, with a full refund payable to you. Notice of such denial will be made using the contact information provided by you. Failure to receive such notice will not alter or invalidate the denial. In no event will Elsevier or Copyright Clearance Center be responsible or liable for any costs, expenses or damage incurred by you as a result of a denial of your permission request, other than a refund of the amount(s) paid by you to Elsevier and/or Copyright Clearance Center for denied permissions.

LIMITED LICENSE

The following terms and conditions apply only to specific license types:

15. Translation: This permission is granted for non-exclusive world English rights only unless your license was granted for translation rights. If you licensed translation rights you may only translate this content into the languages you requested. A professional translator must perform all translations and reproduce the content word for word preserving the integrity of the article. If this license is to re-use 1 or 2 figures then permission is granted for non-exclusive world rights in all languages.

16. Website: The following terms and conditions apply to electronic reserve and author websites:

Electronic reserve: If licensed material is to be posted to website, the web site is to be password-protected and made available only to bona fide students registered on a relevant course if:
This license was made in connection with a course,
This permission is granted for 1 year only. You may obtain a license for future website posting,
All content posted to the web site must maintain the copyright information line on the bottom
of each image,
A hyper-text must be included to the Homepage of the journal from which you are licensing
at http://www.sciencedirect.com/science/journal/xxxxx or the Elsevier homepage for books
at http://www.elsevier.com, and
Central Storage: This license does not include permission for a scanned version of the
material to be stored in a central repository such as that provided by Heron/XanEdu.

17. **Author website** for journals with the following additional clauses:

All content posted to the web site must maintain the copyright information line on the bottom
of each image, and the permission granted is limited to the personal version of your paper.
You are not allowed to download and post the published electronic version of your article
(whether PDF or HTML, proof or final version), nor may you scan the printed edition to
create an electronic version. A hyper-text must be included to the Homepage of the journal
from which you are licensing at http://www.sciencedirect.com/science/journal/xxxxx. As part
of our normal production process, you will receive an e-mail notice when your article
appears on Elsevier’s online service ScienceDirect (www.sciencedirect.com). That e-mail
will include the article’s Digital Object Identifier (DOI). This number provides the electronic
link to the published article and should be included in the posting of your personal version.
We ask that you wait until you receive this e-mail and have the DOI to do any posting.

Central Storage: This license does not include permission for a scanned version of the
material to be stored in a central repository such as that provided by Heron/XanEdu.

18. **Author website** for books with the following additional clauses:
Authors are permitted to place a brief summary of their work online only.
A hyper-text must be included to the Elsevier homepage at http://www.elsevier.com. All
content posted to the web site must maintain the copyright information line on the bottom
of each image. You are not allowed to download and post the published electronic version
of your chapter, nor may you scan the printed edition to create an electronic version.

Central Storage: This license does not include permission for a scanned version of the
material to be stored in a central repository such as that provided by Heron/XanEdu.

19. **Website** (regular and for author): A hyper-text must be included to the Homepage of the
journal from which you are licensing
at http://www.sciencedirect.com/science/journal/xxxxx, or for books to the Elsevier
homepage at http://www.elsevier.com

20. **Thesis/Dissertation**: If your license is for use in a thesis/dissertation your thesis may be
submitted to your institution in either print or electronic form. Should your thesis be
published commercially, please reapply for permission. These requirements include
permission for the Library and Archives of Canada to supply single copies, on demand, of
the complete thesis and include permission for UMI to supply single copies, on demand, of
the complete thesis. Should your thesis be published commercially, please reapply for
permission.

21. Other Conditions:

v1.6

If you would like to pay for this license now, please remit this license along with your payment made payable to "COPYRIGHT CLEARANCE CENTER" otherwise you will be invoiced within 48 hours of the license date. Payment should be in the form of a check or money order referencing your account number and this invoice number RLNK501124341.

Once you receive your invoice for this order, you may pay your invoice by credit card. Please follow instructions provided at that time.

Make Payment To:
Copyright Clearance Center
Dept 001
P.O. Box 843006
Boston, MA 02284-3006

For suggestions or comments regarding this order, contact RightsLink Customer Support: customercare@copyright.com or +1-877-622-5543 (toll free in the US) or +1-978-646-2777.

Gratis licenses (referencing $0 in the Total field) are free. Please retain this printable license for your reference. No payment is required.
THE VITA

Hiromi Shimasaki Brown was born in Kochi, Japan. Hiromi received her Bachelor of Science in biochemistry from Louisiana State University, Baton Rouge, Louisiana, in December 2006. She entered the graduate program at Louisiana State University in the Department of Biological Sciences in the Spring of 2007 and became a doctoral student in the laboratory of Dr. Vince J. LiCata. Hiromi anticipates her doctor of philosophy degree in biochemistry in December 2013.