UC Santa Cruz
UC Santa Cruz Electronic Theses and Dissertations

Title
Single Molecule Studies of DNA Polymerase Fidelity

Permalink
https://escholarship.org/uc/item/8c15v9br

Author
Dahl, Joseph Michael

Publication Date
2016

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA
SANTA CRUZ

SINGLE MOLECULE STUDIES OF DNA POLYMERASE FIDELITY

A dissertation submitted in partial satisfaction of the
requirements for the degree of

DOCTOR OF PHILOSOPHY

in

CHEMISTRY

by

Joseph M. Dahl

September 2016

The dissertation of Joseph M. Dahl
is approved:

_____
Professor Michael Stone, chair

_____
Professor Mark Akeson

_____
Professor Ted Holman

_____
Tyrus Miller
Vice Provost and Dean of Graduate Studies

Table of Contents

List of Figures

List of Tables

Joseph Dahl

SINGLE MOLECULE STUDIES OF DNA POLYMERASE FIDELITY

Abstract

Replicative DNA polymerases (DNAPs) are molecular motors responsible for high fidelity replication of the genome prior to each cellular division. These enzymes require two divalent cations (Me^{2+}) to catalyze nucleotidyl transfer reactions. Many DNAPs have the ability to perform this chemical transformation in two active sites; (i) the polymerase active site where deoxyribonucleoside-triphosphate (dNTP) is selected and checked for template directed Watson-Crick base pairing prior to phosphodiester bond formation, and (ii) the exonuclease active site, where the 3′ primer terminal nucleotide can be excised in a manner consistent with proofreading the nascent DNA strand. Replicative DNAPs must choose the dNTP complementary to the template from a mixture of both dNTPs and ribonucleoside-triphosphates (rNTPs). A balance exists between substrate selection and proofreading, the primary mechanisms that ensure DNAP fidelity. Decades of biochemical research have identified several conserved residues associated with the activities of these two catalytic centers, using the highly processive DNAP from the bacteriophage phi29, structural studies implicated a residues Y254, Y390 & Y226, in the mechanism of translocation.

Translocation occurs after nucleotide addition in the polymerase active site and is essential for processive DNAP replication. The translocation is a single nucleotide step that is required to move the DNAP along its DNA substrate, thus resetting the active site with the next template nucleotide poised for subsequent dNTP selection.

The kinetic complexity of replicative DNAPs requires an experimental method capable of distinguishing the multiple kinetic steps associated with DNAP activity. We developed the first such method for direct measurement of translocation, uniquely capable of making robust quantitative measurements with single nucleotide spatial precision and submillisecond temporal resolution. This method utilizes an α-Hemolysin nanopore embedded in a lipid bilayer which separates two wells of ionic solution to capture phi29 DNAP-DNA complexes in an electric field applied by a patch clamp amplifier. We have modeled the noncovalent kinetic transitions of translocation, nucleotide binding and primer strand transfer between the polymerase and exonuclease active sites. We demonstrated that in the pre-translocation state, which is analogous to the state the DNAP-DNA complex is in immediately following phosphodiester bond formation, a kinetic checkpoint exists where the polymerase can either send the primer terminus from the polymerase active site, in the pre-translocation state, to the exonuclease active site or it can retain the primer strand in the polymerase active site and translocate to the post-translocation state. If the polymerase transfers the primer strand to the

exonuclease active site and the primer terminus is unedited then it can be returned from the exonuclease active site back to the polymerase active site in the pre-translocation state. If the polymerase translocates from the pre-translocation state to the post-translocation state then it is capable of binding to complimentary dNTP. Bound dNTP must dissociate before the polymerase can reverse translocate, from the post to the pre-translocation state.

We applied this experimental method to test several hypotheses regarding the mechanisms of high fidelity DNA replication. In the second chapter we describe the role of two conserved residues, Y226 & Y390, suggested by structural studies to be critical to the mechanism of translocation. Our data indicate that rather than being major determinants in the kinetic mechanism of translocation they are instead associated with active site assembly during dNTP substrate selection. In the third chapter we describe the kinetic mechanisms contributing to the stable incorporation of rNTPs into DNA by replicative DNAPs. In this study we systematically show that while the affinity for correctly base paired rNTP is considerably lower than it is for complementary dNTP, after incorporation of rNTP into the primer terminus the kinetic decision to edit or translocate is no more probable for complementary rNTP than dNTP. In chapter four we describe the role of divalent cations (Me^{2+}) on the noncovalent transitions associated with DNAPs. Here we determined at submillimolar concentrations of Me^{2+} the kinetic pathway for primer strand transfer, from the pre-translocation state in the

polymerase active site to and from the exonuclease active site, is composed of more than the two kinetic states identified in prior experiments with > 1mM Me^{2+}. We also showed that across five orders of magnitude Me^{2+} causes a concentration dependent decrease in the rate of translocation from the pre to the post-translocation state and a concentration dependent increase in the rate of translocation from the post to the pre-translocation state. Also, we demonstrated that, in the presence of Ca^{2+}, the presence of the primer terminal 3′OH does not contribute to the ground state stabilization of dNTP binding.

Dedication

I am honored to

dedicate this

dissertation to my

graduate mentor

and dear friend,

Dr. Kate Lieberman

Acknowledgments

This dissertation is the product of many hours of scientific practice. The capability to produce such a document was fostered over many years of academic diligence and research training. During this time I have had many peers and mentors who I would like to thank. While there are more individual names that come to mind as I reflect on my journey I will keep it brief. I would like to start by thanking Dr. Kate Lieberman, my graduate mentor. She has changed my life and guided me to be the scientist I am today. The impression she has made will persist throughout my career. I would also like to thank Professors Mark Akeson and Hongyun Wang for years of guidance on my journey. I consider my self lucky to have worked so closely with these three outstanding scientists.

I began my 7 year journey in the Nanopore Lab as a community college student, participating as a summer intern with the ACCESS program. The instructors I had in community college, undergraduate studies and through graduate school have all played their parts in directing me to this place and time. Along the way there have been many class mates, lab mates and non-instructional staff members that have kept me grounded in pursuit of my doctoral degree. I would like to acknowledge professors Stone, Akeson and Holman for there mentorship and participation as my committee. I have also accumulated several non-committee academic mentors whose guidance has been perfectly delivered, I am so grateful to each of you. I would also like to thank my family, biological and chosen, you are a beautiful collection of individuals, with divergent perspectives, who have been more instrumental in this endeavor than you may realize. Cheers to each of you.

The text of this dissertation includes reprints of the following previously published material:

[Chapter 2] Dahl, J. M., Wang, H., Lázaro, J. M., Salas, M., and Lieberman, K. R. (2014) Dynamics of Translocation and Substrate Binding in Individual Complexes Formed with Active Site Mutants of phi29 DNA Polymerase. J Biol Chem 289, 6350–6361. [25]

[Chapter 3] Dahl, J. M., Wang, H., Lázaro, J. M., Salas, M., and Lieberman, K. R. (2014) Kinetic Mechanisms Governing Stable Ribonucleotide Incorporation in Individual DNA Polymerase Complexes. Biochemistry 53, 8061–8076. [26]

[Chapter 4] Dahl, J.M., Wang, H., Lieberman K.R., (2016) Modulation of DNA Polymerase Noncovalent Kinetic Transitions by Divalent Cations. J Biol Chem 291, 6456-6470. [27]

K.R.L. listed in these publications directed and supervised the research which forms the basis for the dissertation. All nanopore experiments were conducted by the author (J.M.D) and were devised and analyzed by K.R.L, H.W. and J.M.D with the exception of some analysis that was conducted by Ms. Ashley Cox for *Dynamics of Translocation and Substrate Binding in Individual Complexes Formed with Active Site Mutants of phi29 DNA Polymerase*. In this same paper, Processive DNA Synthesis gel assays were conducted and analyzed by J.M.L. and M.S. In *Kinetic Mechanisms Governing Stable Ribonucleotide Incorporation in Individual DNA Polymerase Complexes,* supplemental Polymerase-Exonuclease Equilibrium gel assays were conducted and analyzed by J.M.L and M.S.

Chapter 1

Introduction

DNA Replication is carried out by replicative DNA Polymerases. Replicative DNA polymerases (DNAPs) are molecular motors that catalyze $5'$ to $3'$ template-dependent DNA replication prior to each cell division. Thus, DNAPs are responsible first and foremost with maintaining the genome. DNA replication requires a primer strand bearing a terminal $3'$ hydroxyl ($3'$OH) to act as the nucleophile for the nucleotidyl transfer reaction. This reaction utilizes two essential divalent metal cations which are coordinated by a pair of highly conserved aspartic acid residues [109, 92]. The first metal ion, termed metal-A, activates the nucleophile. The second metal ion, termed metal-B, stabilizes charge build up in the transition state and coordinates the substrate and leaving group of the reaction [109, 92]. A model for the transition state of nucleotidyl transfer has been proposed and describes the mechanistic roles of these two essential cofactors [92].

Fidelity describes the accuracy of template directed DNA replication and is primarily achieved through two tightly coordinate processes; (i) selection of complementary deoxyribonucleoside-triphosphate (dNTP) from a pool of

dNTPs, and (ii) through 3′ to 5′ exonucleolytic editing of non-complementary deoxyribonucleoside-monophosphates (dNMPs) that escaped initial substrate selection and were incorporated into the primer terminus. Both the polymerase and exonuclease active sites use the two divalent cation mechanism [10, 91]. The coordinated activity of these two active sites imparts replicative DNAPs with an error rate of $\sim 10^8$-10^{10} [62, 54].

Recently it has been demonstrated that ribonucleoside-triphosphates (rNTPs) can be stably incorporated into the genome by replicative DNAPs [76, 104, 44, 83, 72, 107]. A mechanistic description of DNAP fidelity must then account for discrimination among pools of complementary and non-complementary dNTPs and rNTPs. DNA synthesis and editing activities are therefore tasked not only with selection of correctly paired nucleobases in the polymerase active site but also selection of the appropriate sugar moiety.

After addition of a nucleotide to the primer terminus DNAPs must translocate the distance of a single nucleotide to reset the active site for subsequent dNTP selection. Errors in translocation can cause deletion and frameshift mutations. Translocation is thus an essential feature of DNAP fidelity because it maintains the register for nucleotide incorporation. Errors that evade the polymerases mechanisms for fidelity can cause persistent mutations and disease. Absence of a technique capable of distinguishing the kinetic steps of substrate selection, exonucleolytic editing and translocation has limited our understanding of the mechanisms that regulate fidelity during DNA replication.

In eukaryotes leading and lagging strand DNA replication is carried out by Pol-ε and Pol-δ, respectively, which belong to the B-family of replicative DNA polymerases [76]. There is a high degree of structural conservation in the B-family of DNAPs which have a structure likened to a right hand; with thumb, fingers and palm subdomains [6, 32, 48, 57, 40] (Fig 1.1). Figure 1.1 depicts the orientation of the conserved subdomain architecture in a DNAP complexed with DNA and dNTP substrates [6]. The thumb subdomain interacts with the primer-template duplex, orienting the primer terminus and template base (n=0) in the active site in the post translocation state [6, 32, 48, 57, 40]. The fingers subdomain is associated with substrate selection. Structural studies indicate the fingers make an open to closed transition upon complementary nucleotide binding [6, 32, 48, 57, 40]. The palm subdomain contains a set of highly conserved residues which provide the architecture for dNTP substrate coordination and catalysis [6, 32, 48, 57, 40].

Prior to the work presented here data relating the mechanisms of DNAP substrate selection and proofreading to translocation were limited. The most insightful data were provided by structural studies. Using the highly processive DNAP from the bacteriophage phi29 a structural mechanism for translocation has been suggested for B-family DNAPs [6]. A series of DNAP-DNA binary and DNAP-DNA-dNTP ternary complexes show snapshots of the fingers closed post-translocation state ternary complex, which is a structural proxy for the n=0 fingers closed pre-translocation state binary complex, the

3
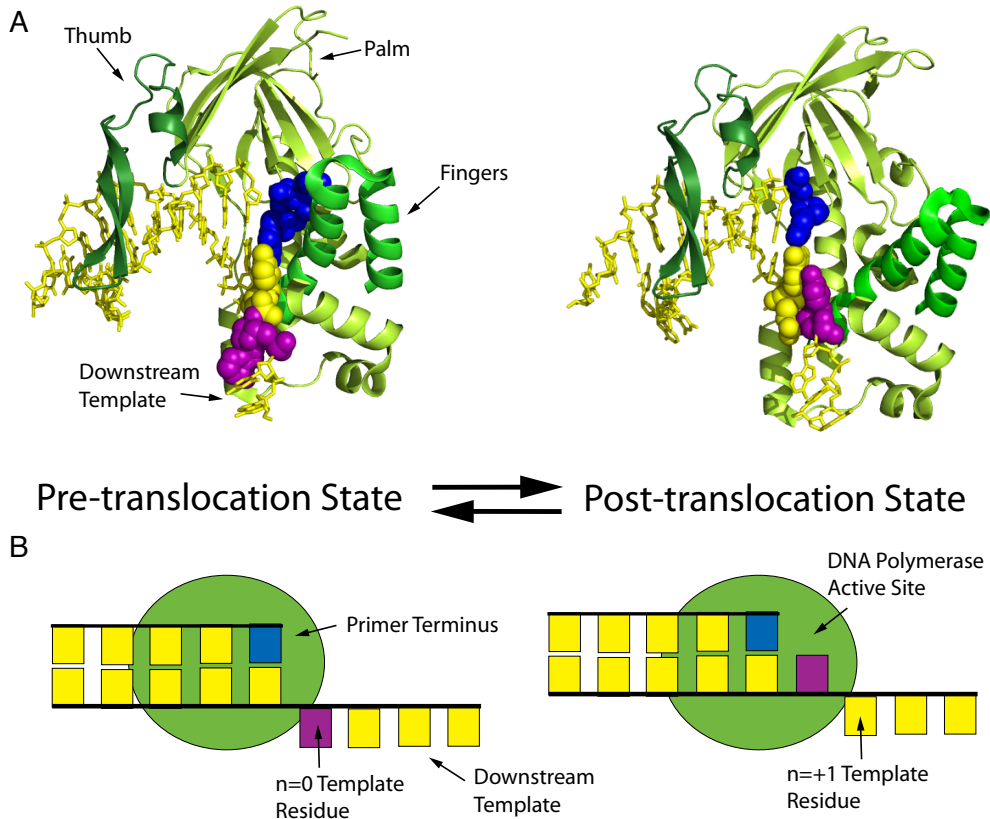
Figure 1.1: Phi29 DNAP-DNA complex architecture across the translocation step. A: Cartoon model of phi29 DNAP crystal structures in the DNAP-DNA-dNTP fingers closed ternary complex, a structural proxy for the pre-translocation state, and the DNAP-DNA fingers open post-translocation state binary complex (adapted from ref. 6, PDBs 2PYL & 2PZS, respectively). Residues of phi29 DNAP not associated with the canonical fingers (bright green), palm (light green) and thumb (dark green) subdomains are omitted to clearly show the conserved subdomain architecture. **B:** Cartoon schematic of DNA movement in the polymerase active site across the translocation step. The DNAP (green) moves the distance of a single nucleotide on its DNA substrate (yellow in A & B). Across the translocation step the nascent base pair, formed between the primer terminus (blue in A & B) and n=-1 template residue, vacates the polymerase active site and the n=0 template residue (purple in A & B) rotates into the polymerase active site assembling the active site for subsequent dNTP substrate binding.

n=+1 translocated fingers open binary and fingers closed ternary complexes.

These data suggested the structural role of a set of highly conserved residues

that were previously determined to be associated with DNAP fidelity by *in vitro*

ensemble experiments [11, 95, 15, 6, 69, 14, 86]. Most notable were residues Y226,

Y254 and Y390. Y254 has been described as a steric gate residue associated with ground state stabilization of dNTP binding and providing selectivity for dNTPs over NTPs [14,15, 25, 86]. Y226 and Y390 have phenolic oxygens and planar ring systems which were shown structurally to interact with one another and Y254 in distinct ways across the fingers open to closed transition [6]. Apart from their role in substrate binding these residues were implicated in the structural mechanism for translocation. These structural data were unique in the mechanistic description of translocation by B-family DNAPs, however, the mechanism proposed was limited to the static images determined from the X-ray crystallographic data. Thus, there was no other evidence implicating these residues directly in the mechanism for translocation. The related ensemble experiments, being a population average, could not reconcile the role of these residues in translocation [11, 95, 15, 6, 69, 14, 86]. To determine the mechanistic role of these residues in B-family DNAP fidelity a tool that can directly measure translocation is required.

For exonucleolytic editing to occur the primer strand must be transferred ~30-40 Å [6, 32, 37, 58, 98, 93, 47, 48]. This process can be intramolecular or intermolecular depending on the polymerase [50, 18]. For phi29 DNAP the exonuclease domain is composed of N-terminal amino acids of the polymerase primary structure, thus primer strand transfer is an intramolecular process for this DNAP [6, 9, 42 58]. Very little was known about the kinetic pathway of primer strand transfer between the polymerase and exonuclease

active sites. It has been inferred that incorporation of a mismatch at the primer terminus inhibits the pre to post-translocation state transition and that proofreading is consequently initiated in the pre-translocation state [62, 82].

Another B-family DNAP, from the bacteriophage RB69, bears both polymerase and exonuclease active sites in a single primary structure [47, 98, 99, 100]. Crystallographic data of RB69 DNAP complexed with DNA bearing a dAMP-furan mismatch at the primer-template junction have suggested a structural mechanism for primer strand occupancy of the exonuclease active site [47]. These data resolved two structures of the DNAP-DNA complex in the pre-translocation state and two other structures, in the same asymmetric subunit, with the primer strand in the exonuclease active site [47]. This further supports a model where primer strand transfer between the two active sites is initiated in the pre-translocation state. However, these structures are again snapshots of complexes and provide limited evidence for the biophysical mechanism of primer strand transfer between the polymerase and exonuclease active sites in terms of the translocation state. Another caveat to the RB29 DNAP data is that these structures were determined with an unnatural substrate inducing a mismatch-like conformation in the polymerase active site [47]. In these structures the DNA template was structurally distinct from structures determined with the natural DNA template [47, 98, 99, 100, 106].

Ensemble studies of phi29 DNAP had identified the amino acid residues essential for catalysis in the exonuclease active site (D12 and D66) [9, 42] and

another residue associated with facilitating primer strand transfer between the two active sites (N62) [58]. Prior to the work described below the kinetic mechanisms for the process of primer strand transfer between the polymerase and the exonuclease active sites, with regard to the translocation state, had never been described. Thus, the translocation state that the primer strand transfer initiates from and returns to had not been directly measured.

Developing a tool capable of resolving the kinetic steps involved in a single nucleotide addition cycle. We described the first direct observation of DNAP translocation in 2012 [24]. This study utilized a single α-Hemolysin (α-HL) nanopore to detect phi29 DNAP bound to DNA in both binary and ternary complexes. In the last ~20 years the nanopore has emerged as a single molecule technique to study DNA and DNAPs [2, 5, 30, 41, 65]. Initial nanopore studies using phi29 DNAP demonstrated the ability to monitor and control replication of individual molecules of DNA [65, 20]. This approach allows for robust data sets measuring hundreds of individual DNAP-DNA complexes with submillisecond temporal resolution.

In our work, the device utilizes a single isolated α-HL nanopore embedded in a lipid bilayer which separates two wells of ionic buffered solution (Fig 1.2 A). A patch clamp amplifier applies a voltage across the bilayer and records the current. A single nanopore inserted into the bilayer allows for the flow of ions (K^+ and Cl^-) across the channel it creates. In this assay DNA is driven to the nanopore by the electric field. The dimensions of an α-HL nanopore are such

that a single strand of DNA can pass through yet double stranded DNA is too

large and in turn is captured in the pore. When phi29 DNAP-DNA1-H_H

complexes are captured by the pore the enzyme is too large to to enter the

channel and instead perches atop the pore (Fig 1.2 B, C). These DNAP-DNA
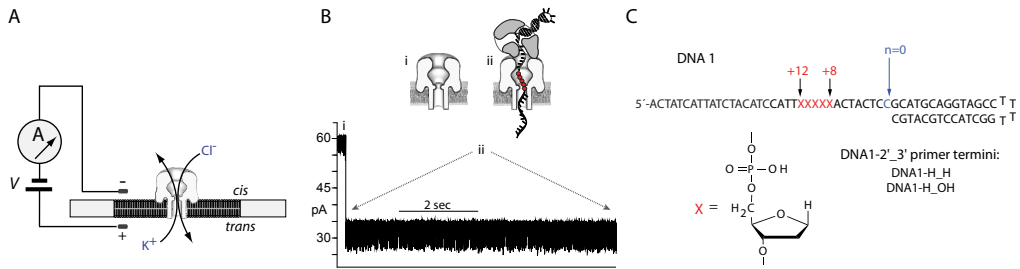
Figure 1.2: Capture of individual DNAP-DNA complexes atop an α-HL nanopore. A: A single α-HL nanopore is inserted in a ~25 µm-diameter lipid bilayer separating two chambers (cis and trans) containing buffer solution. A patch clamp amplifier applies voltage across the bilayer and measures ionic current carried through the pore by K^+ and Cl^- ions. **B:** Ionic current time trace for a binary complex formed between wild type phi29 DNAP and a DNA substrate (DNA1-H_H, in panel C) captured at 180 mV applied potential in buffer containing 10 mM K-Hepes, pH 8.0, 0.3 M KCl, 1 mM EDTA, 1 mM DTT, and 11 mM $MgCl_2$. phi29 DNAP and DNA1-H_H were added to the nanopore cis chamber to final concentrations of 0.75 and 1 µM, respectively. Cartoons above the current trace illustrate the sequence of events: **i)** the ionic current through the open nanopore; **ii)** the current drops rapidly when a complex is captured. The enzyme is too large to enter the nanopore, and therefore the complex, with the enzyme bound at the primer–template junction of the DNA substrate, perches atop the pore. The DNA template strand of the complex is suspended through the nanopore lumen, which is just wide enough to accommodate a single strand of DNA. Individual complexes reside atop the nanopore for tens of seconds. Amplitude fluctuations correspond to movement of the DNA substrate relative to the enzyme and the nanopore by the distance of a single nucleotide [65, 66]. This spatial displacement of the DNA substrate that occurs during the translocation is detected in the time traces of ionic current by the use of a reporter group of five consecutive abasic (1′-H, 2′-H) residues in the template strand (red circles in the cartoon). The fluctuations between the two amplitudes continue until complexes dissociate or are manually ejected, after which another complex can be captured. **C:** DNA hairpin substrates comprise a 14-base pair duplex region and a single-stranded template region of 35 nucleotides. The reporter group of five abasic (1′, 2′-H) residues spans positions +8 to +12 (red Xs in the sequence). DNA1-H_OH bears a 2′-H, 3′-OH primer terminus and DNA1-H_H bears a 2′-H, 3′-H primer terminus.

complexes reside atop the pore for 10s - 100s of seconds. During this time, for complexes captured in the absence of nucleotide substrate, DNAP-DNA1-H_H complexes rapidly fluctuate between two discreet states: a high amplitude state centered at ~32pA and another lower amplitude state centered at ~26pA (Fig 1.2 B ii, & 1.4 A). These states were identified by the use of an abasic reporter group composed of 5 consecutive abasic residues (1′, 2′-H) (Fig 1.2 C). Prior nanopore experiments measuring replication by phi29 DNAP identified the ideal size and location for the abasic reporter to maximize the signal resolution for non-replicating complexes [65, 20] . The absence of nucleobases in the reporter region allows for an increase in the flow of ions for complexes that are captured across the channel. Since these foundational studies other nanopores and approaches have been developed for analysis of DNA and its associated proteins, however, we chose to implement this particular assay for biophysical analysis of replicative DNAPs. We demonstrated that DNAP-DNA complexes captured atop the nanopore fluctuate the distance of a single nucleotide across the translocation step through Brownian thermal motion [24]. In this study we assigned the identity of the translocation states in the ionic current time traces for phi29 DNAP-DNA complexes, with the upper amplitude corresponding to the pre-translocation state and the lower amplitude corresponding to the post-translocation state. The data demonstrated that the post-translocation state could be stabilized by dNTP in a concentration dependent manner, without saturation, up to 1200uM

9

[24]. Additional ensemble assays showed a strong correlation between the probability of both exonucleoytic editing and pyrophosphorylysis with the probability of pre-translocation state occupancy for complexes measured using the nanopore. Pyrophosphorylysis is the reverse reaction of nucleotidyl transfer in the polymerase active site, where pyrophosphate (PPi), the leaving group of DNA polymerization, reacts with the terminal nucleotide in the pre-translocation state to reduce the primer length by a single nucleotide. Considered with the nanopore data the ensemble experiments allowed us to propose a model for the process of translocation, dNTP binding and exonucleolytic editing during a single dNTP incorporation cycle (Fig 1.3) [24]. However, we did not have the capability for robust quantification to verify the model at this early stage in the project.

pre-translocation   post-translocation

polymerase active site

exonuclease active site
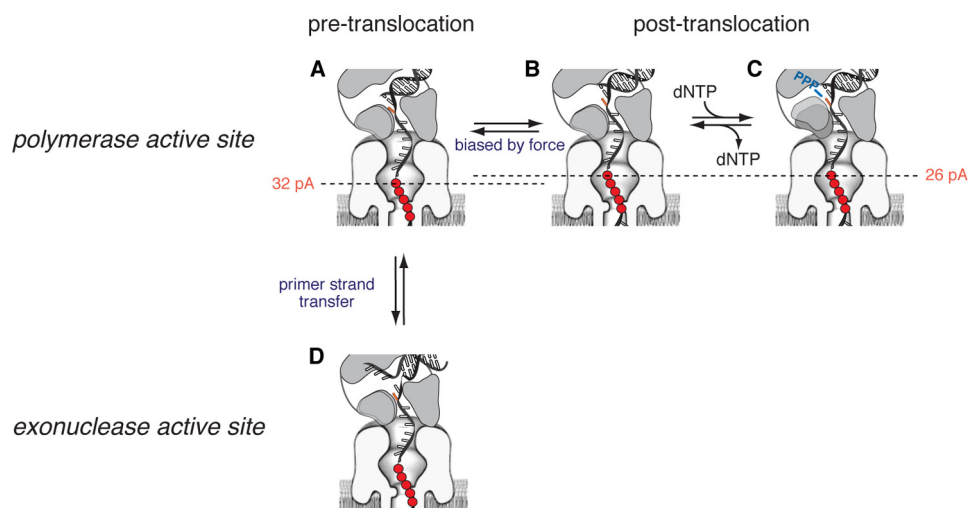
Figure 1.3: Schematic illustration of a model for the states of phi29 DNAP-DNA complexes observed on the nanopore. Immediately following phosphodiester bond formation, phi29 DNAP-DNA complexes are in the pre-translocation state (**A**) with the primer-template junction in the polymerase active site. Driven by Brownian thermal motion, the binary complexes fluctuate between the pre-translocation (A) and post-translocation (**B**) states in the polymerase active site. The translocation equilibrium is biased toward the pre-translocation state by force on the template strand. Complexes can be rectified in the post-translocation state by the binding of complementary dNTP (**C**). When the primer strand of the DNA substrate bears a 3′-H terminus, complementary dNTP binds to and unbinds from the post-translocation complex. During processive synthesis, when the DNA substrate bears a 3′-OH terminus, complementary dNTP binding drives the complex forward through the nucleotide addition cycle. The pre-translocation state (A) is also a branch point from which the primer strand can be transferred to and from the exonuclease active site (**D**). In these schematics, the abasic residues spanning positions 8 to 12 of the template strand are shown as red circles, and the nucleotide at position n=0 is highlighted is in orange. The fingers-closing transition that accompanies complementary dNTP binding is illustrated in C. The horizontal dashed lines indicate the displacement of the template strand in the nanopore lumen by a distance of 1 nucleotide that reports the changes in register of the DNA substrate with respect to the enzyme during the transitions between the pre-translocation and post-translocation states.

Measuring the rates of translocation for DNAP-DNA complexes. Subsequently

we modeled the energetic contributions of DNA sequence on the rates of

translocation [66]. In this work we developed a mathematical method to

quantify the specific rates across the translocation step for complexes

captured atop an α-HL nanopore. To determine the rates we utilized the half amplitude threshold method [23] to extract the dwell time samples of each of the discreet amplitude associated translocation state populations from the measured ionic current time traces (Fig 1.4 A). We applied the auto-correlation function to quantify the kinetic rates of transition from these dwell time samples. These experiments demonstrated that while DNAPs are molecular motors responsible for replication of the genome, in all its diversity of sequence, the active site proximal sequence is a determinant for the path,
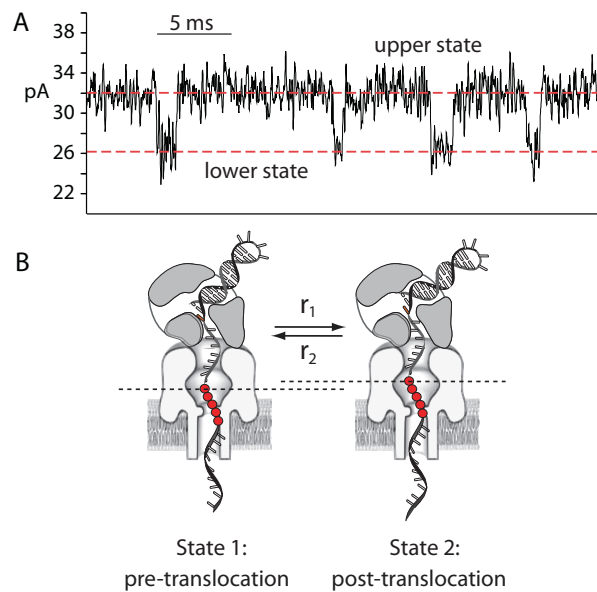
Figure 1.4: A two-state model for the dynamics of the phi29 DNAP translocation step. A: Current trace segment for a captured phi29 DNAP–DNA1 complex. The ionic current fluctuates between the pre-translocation state ~32 pA and the post-translocation state ~26 pA on the millisecond time scale as complexes are held atop the nanopore at 180 mV applied potential. **B:** The two-state kinetic model for the phi29 DNAP translocation step. The transitions between the two states are observed as fluctuations in the measured ionic current as the template strand, and its embedded abasic reporter, are displaced in the nanopore lumen.

and thus the probability, of escape from both the pre and post-translocation states [66].

This established the first technique to directly measure the rates of translocation in single DNAP-DNA complexes. To date, this technique is uniquely capable of measuring translocation in DNAP-DNA complexes with single nucleotide spatial precision and submillisecond temporal resolution (Fig 1.4 B). In addition to measuring the rates of translocation the nanopore can also manipulate the rates of translocation by varying the voltage applied to captured DNAP-DNA complexes. With increasing voltage decreasing the rate of forward translocation (*r1*), from the pre to post-translocation state, and simultaneously increasing the rates of reverse translocation (*r2*) (Fig 1.4 B). Manipulation of voltage in nanopore experiments provides a unique variable that has allowed us to further leverage the kinetic mechanisms of replicative DNAPs and explore the mechanisms of DNAP fidelity.

Determining the kinetic pathway for translocation and nucleotide binding. Having established a technique capable of quantifying the kinetic rates across the translocation step in single DNAP-DNA complexes we further developed the technology to probe more complex experimental questions regarding the kinetic mechanisms that contribute to DNAP fidelity. In 2013 we developed a three state model to describe the relationship between translocation and nucleotide binding [67]. Translocation is believed to operate by two distinct, yet not necessarily mutually exclusive, pathways [46, 88]; (i) The Brownian Ratchet,

where thermal energy of DNA diffusion in the DNA polymerase drives the process of translocation, and (ii) The Power Stroke Model, where dissociation of pyrophosphate (PPi) and or binding of dNTP provide the mechanical energy to translocate the enzyme on its DNA substrate. We developed and tested several models to determine the kinetic pathway(s) for translocation and nucleotide binding.

Using a half amplitude threshold method we again compared the dwell time samples for ionic current time trances of complexes captured in the absence (Fig 1.5 A i) or presence of complementary dNTP (Fig 1.5 A ii) and saw the emergence of a cluster of complexes at the lower amplitude, post-translocation state, in the presence of complementary nucleotide (Fig 1.5 B). As above, using the autocorrelation method we could estimate the kinetic rates associated with dNTP binding and translocation simultaneously. The dwell time observed for complexes captured in the absence of nucleotide is centered around 0.4ms (Fig 1.5 B i). For complexes captured in the presence of 10μM complementary dNTP a cluster centered around 30ms emerges that is indicative of a ternary complex (Fig 1.5 B ii). When complementary nucleotide is titrated to higher concentrations the shorter-duration lower amplitude binary complex population decreased in frequency indicating that the ternary complex becomes the dominant species in the lower amplitude (Fig1.5 C).
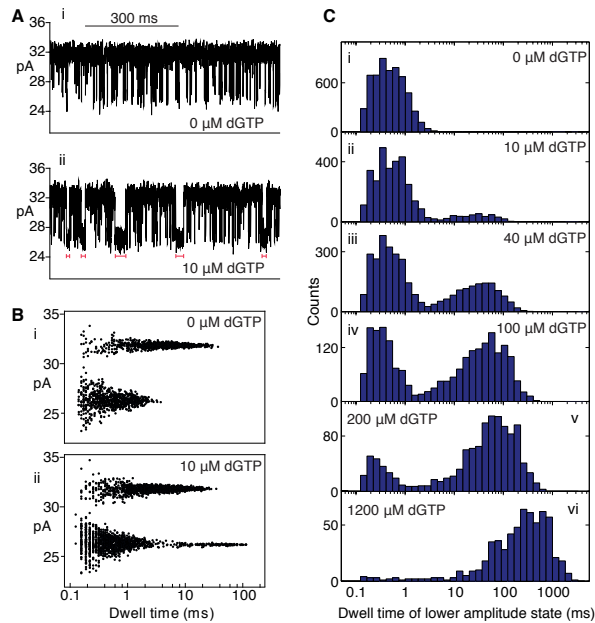
Figure 1.5: Complementary dNTP binding to phi29 DNAP complexes. A: Ionic current traces for phi29 DNAP–DNA1 complexes, captured at 180 mV in the presence of 0 µM (**i**) or 10 µM (**ii**) dGTP. In the presence of dGTP, some of the transitions to the lower amplitude, post-translocation state have dwell times that are significantly longer than those typically observed in the absence of dGTP (indicated by red lines under the trace in ii). **B:** Plots of log dwell time vs amplitude for dwell time samples extracted from complexes captured at 180 mV in the presence of 0 µM (**i**) or 10 µM (**ii**) dGTP. In the absence of dNTP (**i**) there are two clusters that are well separated in amplitude: the pre-translocation state dwell time samples centered at ~32 pA, and a cluster of post-translocation state samples centered at ~26 pA. Upon addition of dGTP (**ii**) a third cluster, with longer dwell times, emerges at the post-translocation state amplitude. **C:** Histograms of log dwell time samples of the post-translocation, lower amplitude state centered at ~26 pA, extracted from ionic current traces for complexes captured at the indicated dGTP concentrations.

We proposed and tested three models for nucleotide binding based on the nanopore data; (i) a four state model where no assumptions are made and nucleotide can bind to both the pre and post-translocation states with transitions between all states possible (Fig 1.6 A), (ii) a three state model where nucleotide binds to the post-translocation state and can facilitate translocation (Fig 1.6 B) and (iii) a three state model where nucleotide binds
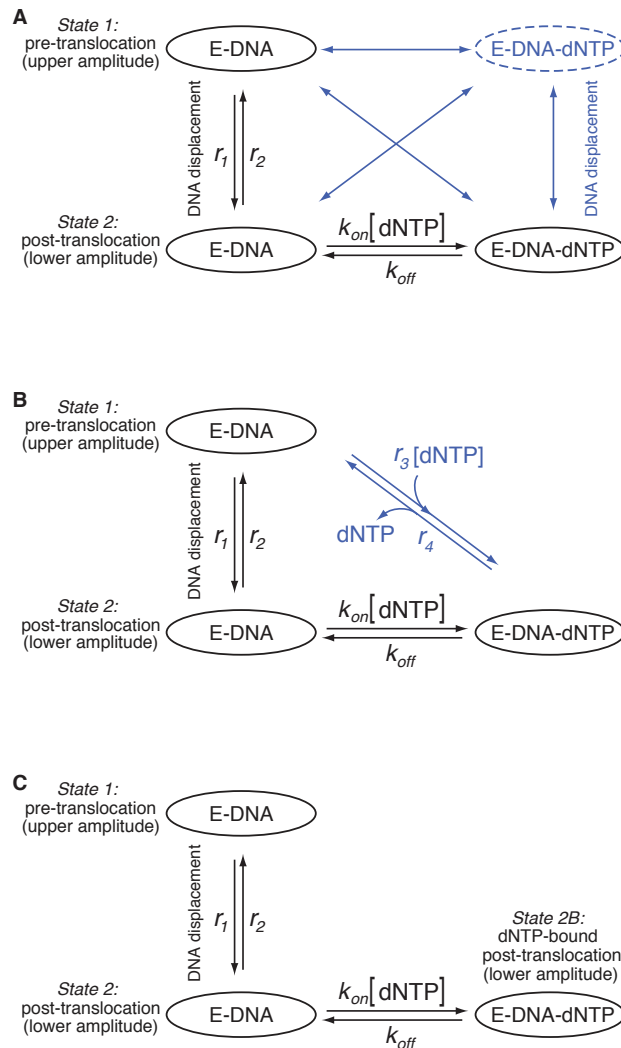
Figure 1.6: Potential models for the kinetic relationship between translocation and dNTP binding. Phi29 DNAP–DNA binary complexes fluctuate between the pre-translocation and post-translocation states with the forward rate r1 and the reverse rate r2. **A**: The four-state model used for studying the equilibrium properties of translocation state transitions in phi29 DNAP–DNA complexes. **B**: A three-state model in which dNTP can influence the transition from the pre-translocation to the post-translocation state (r3[dNTP]) and can bind to complexes in the post-translocation state (kon[dNTP]). Dissociation of dNTP can occur prior to (koff) and concurrent with (r4) the transition from the post-translocation to the pre-translocation state. **C**: A three-state model in which translocation and dNTP binding are sequential: dNTP can bind to complexes only after the transition from the pre-translocation to the post-translocation state (kon[dNTP]); the transition from the post-translocation to the pre-translocation state cannot occur before the dissociation of dNTP (koff).

to the post-translocation state and does not facilitate translocation (Fig 1.6 C). In earlier studies we demonstrated that nucleotide stabilizes the post-translocation state in a concentration dependent manner without saturation up to and above ~1200µM and that binding of nucleotide to the pre-translocation state was negligible [24]. These data support the two three state models (Fig 1.6 B & C), where nucleotide binding stabilizes the post-translocation ternary complex, and is inconsistent with the 4 state model proposed (Fig 1.6 A).

To determine which of the two three state models describes the data measured with the nanopore we had to determine if dNTP binding facilitates translocation or if translocation and nucleotide binding are sequential. The nanopore proved particularly well adept for this purpose, the ability to manipulate voltage to apply varying amounts of force to the DNAP-DNA complex changes the rates of translocation for complexes captured in the absence of nucleotide. If dNTP binding facilitated the translocation step, dNTP should change the rates of translocation, thus complexes captured at a specific concentration of nucleotide across a range of voltages should have a distinct set of rates when compared to complexes captured absent dNTP.

A simple kinetic explanation of this is that the rate describing the transition to the post-translocation dNTP bound state is the product of the binding constant k_{on} and the concentration of dNTP, k_{on}[dNTP]. If the transition from the pre-translocation state to the dNTP bound post-translocation state is possible then at high concentrations of dNTP the transition from the pre-

translocation state to the dNTP bound post-translocation state should dominate the path for escape from the pre-translocation state. However, at high concentrations of complementary dNTP we determined a high degree of experimental and statistical uncertainty for the rate describing this transition, nominally (*r3*[dNTP]), and no significant change in the rates of forward translocation (*r1*) when comparing complexes captured in the absence or presence of varying concentrations of dNTP [67]. Thus, we concluded that the rate of transition form the pre-translocation state to the post-translocation dNTP bound state (*r3*[dNTP]) (Fig 1.6 B) is negligible and that the path of escape from the pre-translocation state directly to the dNTP bound post-translocation state may not exist.

The simplest explanation for the reverse transition, from the dNTP bound post-translocation state to the pre-translocation state (*r4*) (Fig 1.6 B), not existing is the principle of microscopic reversibility; simply stated, if the forward transition is possible then the reverse transition must be possible and conversely if the transition is not possible then neither are possible. We directly tested the potential for the existence of the reverse transition (*r4*) none the less. At high concentrations of complementary dNTP the dominant species for the post-translocation state is the dNTP bound state. If the kinetic transition from the post-translocation dNTP bound state to the pre-translocation state were possible then escape to the pre-translocation state from the dNTP bound post-translocation state should be quantifiable.

18

However, experimentally the kinetic rates were considerably smaller than the error associated with the proposed transition [67]. The statistical uncertainty again supports the theory that a direct transition from the dNTP bound post-translocation state to the pre-translocation state ($r4$) does not exist. These data exclude model ii (Fig 1.6 B) as a potential kinetic model for translocation and nucleotide binding.

If, in fact, model iii (Fig 1.6 C) were the most accurate model then both forward ($r1$) and reverse ($r2$) translocation rates and nucleotide association (k_{on}) and dissociation (k_{off}) rates should be independent and the rate associated with nucleotide binding should be unaffected by voltage. Indeed the data demonstrate this (Fig 1.7) [67]. In the presence of specific concentrations of complementary dNTP and across a range of voltages the rates of translocation ($r1$ & $r2$) are not affected by the nucleotide concentration (Fig 1.7 A & B), having the same response to voltage irrespective of the complementary nucleotide concentration each complex is captured in the presence of. As well, the rates for nucleotide binding and release, k_{on} and k_{off} respectively, show no systematic trend across the range of voltages tested (Fig 1.7 C & D).

It stands to reason that within the range of voltages tested the potential for a direct transition between the pre-translocation and post-translocation dNTP bound states should be sensitive to voltage because within this regime of voltages (140-210mV) the translocation step is sensitive to voltage. Thus
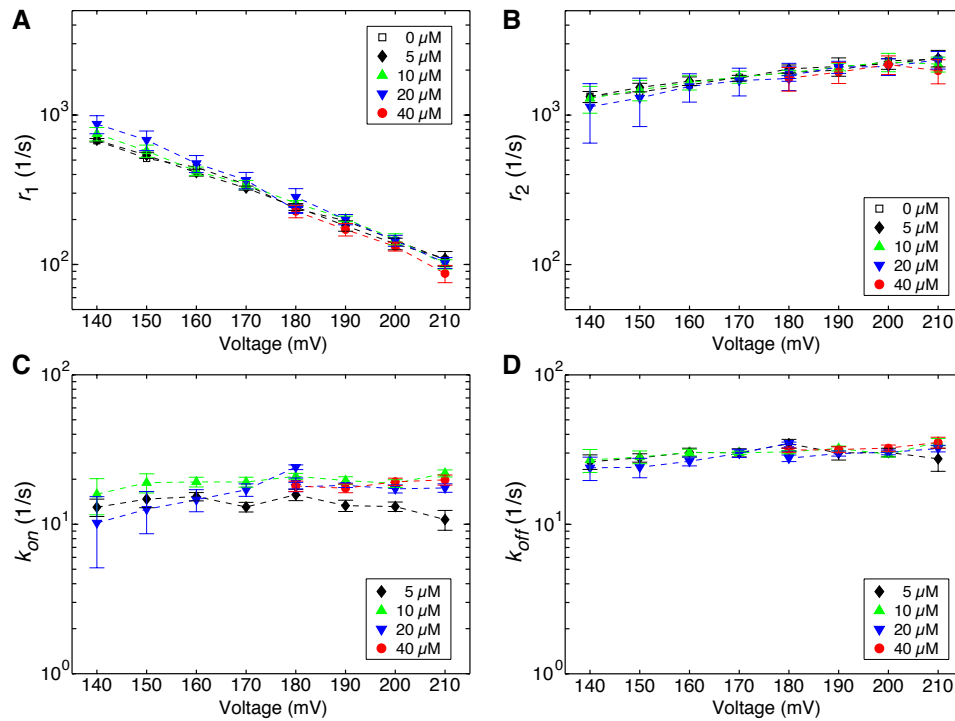
Figure 1.7: Translocation rates and dNTP association and dissociation rates determined from the three-state model. Plots of (**A**) $\log(r_1)$ vs voltage and (**B**) $\log(r_2)$ vs voltage for phi29 DNAP–DNA1 complexes captured in the presence of 0, 5, 10, 20, or 40 µM dGTP. Plots of (**C**) k_{on} vs voltage and (**D**) k_{off} vs voltage for complexes captured in the presence of 5, 10, 20, or 40µM dGTP. Rates were extracted from ionic current traces using the autocorrelation function and the three-state model in Figure 6C. Because k_{on} is calculated by dividing the apparent rate by the concentration of dGTP added to the experiments, the variability in the values for this rate may be attributable to small differences in the actual dGTP concentration between experiments due to pipetting inaccuracies. Errors bars indicate the standard error.

both translocation and nucleotide binding rates should be affected by voltage if dNTP facilitated translocation. When considered in its entirety the data support a sequential model for translocation and dNTP binding; from the pre-translocation sate the complex can proceed to the post-translocation state where a kinetic decision is made for escape from the post-translocation state; to either take the path to the pre-translocation state or bind dNTP. When the

latter is chosen dNTP must first dissociate before the complex can reverse translocate to the pre-translocation state (Fig 1.6 C). Establishing this model provided the experimental capability to simultaneously quantify the rates of translocation and nucleotide binding in DNAP-DNA-dNTP ternary complexes.

Determining the kinetic pathway for translocation and primer strand transfer. The most recent technological advance to study fidelity in DNAP-DNA complexes captured atop a nanopore was the development of a model that could measure the rates of primer strand transfer between the polymerase and exonuclease active sites [68]. A description of this process requires discriminating which translocation state the DNAP is in when the primer strand is transferred between the two polymerase and exonuclease active sites. Measuring the rates of primer strand transfer and translocation allows for a description of the probability of editing during a single nucleotide addition cycle.

Ionic current time traces of phi29 DNAP (D12A/D66A) bound to our standard DNA sequence (Fig 1.2 C) bearing a dideoxy-terminated (2′H, 3′H denoted H_H) primer terminus, captured at 180mV applied voltage, fluctuate rapidly across the translocation step, between the pre and post-translocation states, with a bias for the upper amplitude pre-translocation state (Fig 1.8 A). The D12A/D66A mutation substitutes non-polar alanine residues for the two essential aspartic acid residues that are ligands for the metal ions that catalyze exonucleolysis [9, 42]. This mutant is the ideal version of the enzyme

21
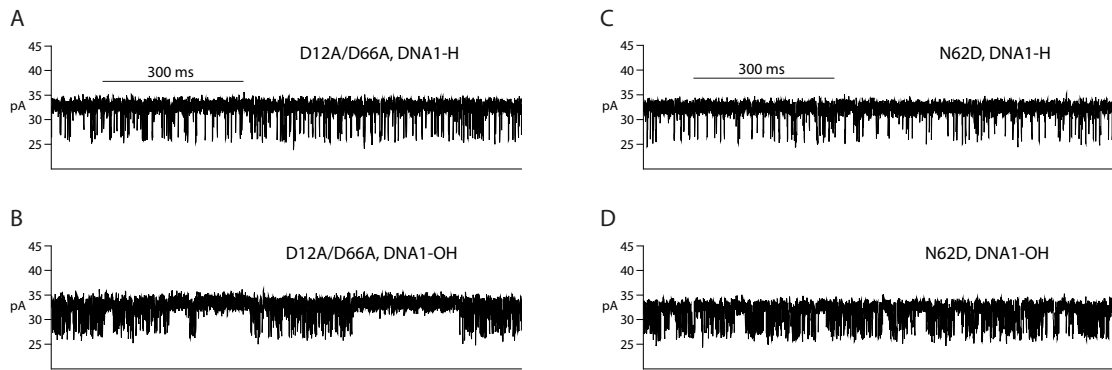
Figure 1.8: Upper amplitude pauses in phi29 DNAP-DNA complexes. A: Ionic current time trace from a complex formed between D12A/D66A phi29 DNAP and DNA1-H_H, captured at 180 mV. **B:** Ionic current time trace from a complex formed between D12A/D66A phi29 DNAP and DNA1-H_OH, captured at 180 mV. **C:** Ionic current time trace from a complex formed between N62D phi29 DNAP and DNA1-H_H, captured at 180 mV. **D:** Ionic current time trace from a complex formed between N62D phi29 DNAP and DNA1-H_OH, captured at 180 mV.

to test primer strand transfer between the polymerase and exonuclease active sites because it does not allow for covalent modification of the primer terminus, elucidating the mechanism of exonucleoltic editing by determining the translocation state from which primer strand transfer initiates and returns. When phi29 DNAP (D12A/D66A) is captured with a DNA substrate bearing the natural primer terminus 2′H, 3′OH (H_OH), rather than (H_H) in otherwise identical sequence context, there is a dramatic difference in the appearance of the ionic current time traces (Fig 1.8 A & B). In contrast to the rapid fluctuations between the two amplitudes observed with the H_H primer terminated substrate the H_OH substrate pauses in the upper amplitude. Prior ensemble and nanopore experiments indicated that the H_H primer terminus was a poor substrate for exonucleolysis when compared to the

H_OH primer terminated substrate [24]. We proposed that the pauses observed in the upper amplitude for (D12A/D66A)DNAP-DNA1-H_OH complexes may be periods where the primer terminus has been transferred from the polymerase active site to the exonuclease active site (Fig 1.8 B).

To test this we employed the use of another mutant of phi29 DNAP, N62D. Biochemical studies conducted under a collaborator, Dr. Margarita Salas, indicated that the N62D mutation impaired the primer strand transfer process [58]. We hypothesized that if the observed pauses for DNAP-DNA1-H_OH complexes were associated with primer strand transfer between the two active sites that a mutation that affects this process should perturb the upper amplitude pauses observed in the ionic current time traces. Indeed, this was what we observed. Qualitative comparison of traces for DNAP-DNA1-H_OH complexes, formed with phi29 DNAP D12A/D66A (Fig 1.8 B) or N62D (Fig 1.8 D), indicate that the duration and frequency of the upper amplitude pauses are decreased for the N62D mutant when compared with the D12A/D66A mutant. The ionic current trace data with the N62D mutant of phi29 DNAP are the most compelling evidence supporting the hypothesis, yet several other experimental observations indicate that pauses in the upper amplitude for DNAP-DNA-H_OH complexes correspond to primer strand transfer between the polymerase and exonuclease active sites, including experiments monitoring iterative rounds of nucleotide excision and addition with wild type DNAP [68]. Confident that pauses in the upper amplitude, uniquely observed

23
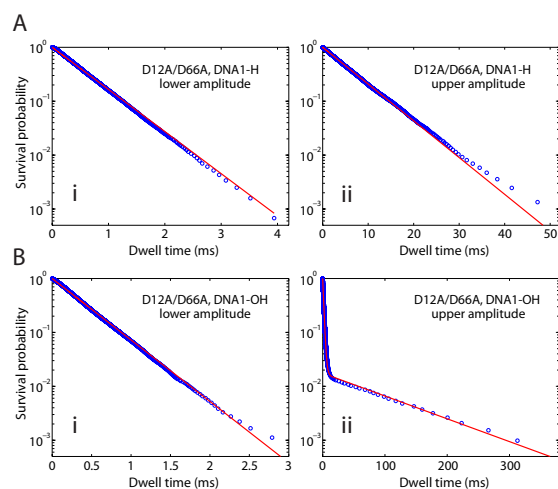
Figure 1.9: A third kinetic state in phi29 DNAP complexes formed with DNA1-OH. A: Survival probability vs dwell time plots for dwell time samples from (**A, i**) the lower amplitude state, or (**A, ii**) the upper amplitude state, extracted from ionic current traces for complexes formed between D12A/D66A phi29 DNAP and DNA1-H_H. **B:** Survival probability vs dwell time plots for dwell time samples from (**B, i**) the lower amplitude state, or (**B, ii**) the upper amplitude state, extracted from ionic current traces for complexes formed between D12A/D66A phi29 DNAP and DNA1-H_OH. In panels **A, i; A, ii; B, i;** the survival probability data points were fitted to an exponential distribution. In panel **B, ii;** the survival probability data points were fitted to a model of two exponential modes. The dwell time samples were extracted from data files in which complexes were captured at 180 mV; each file yields ~20,000–40,000 dwell time samples for each amplitude level. In the plots, while 1 out of every 20 points is shown, the curves are fit to the full set of ~20,000–40,000 dwell time samples.

for DNAP-DNA1-H_OH complexes, are associated with the primer terminus occupying the exonuclease active site we set out to model the kinetic mechanism for primer strand transfer between the two active sites.

The upper and lower amplitude ionic current trace data for complexes of (D12A/D66A)DNAP-DNA1-H_H can be plotted on a semi-log plot of survival probability as a function of dwell time (Fig 1.9 A) [68]. This plot describes the probability of a state continuing for a specific duration, *t*, prior to transitioning to the other amplitude associated state. Such a plot is useful for describing

the exponential decay of each state. The upper and lower amplitudes of complexes of (D12A/D66A)DNAP-DNA1-H_H each fit well to a single exponential decay function (Fig 1.9 A i & ii), indicating one kinetic state at each amplitude for these complexes. Complexes of (D12A/D66A)DNAP-DNA1-H_OH behave differently (Fig 1.9 B). A plot of the lower amplitude survival probability vs dwell time fit to a single exponential decay function, again indicating a single kinetic state (Fig 1.9 B i). However, for this complex survival probability vs dwell time plots of the upper amplitude show multiple modes of exponential decay. Rather than fitting to a single exponential decay function the upper amplitude for (D12A/D66A)DNAP-DNA1-H_OH complexes fit a model of two exponential modes, indicating the presence of a second kinetic state in the upper amplitude (Fig 1.9 B ii) [68].

This second kinetic state describes periods when the primer strand occupies the exonuclease active site. In the survival probability vs dwell time plots of the upper amplitude for (D12A/D66A)DNAP-DNA1-H_OH complexes (Fig 1.9 B ii) the initial decay mode corresponds to the the escape from the upper amplitude to the lower amplitude, transitioning from the pre to the post-translocation state, while the later decay mode describes the pauses observed in the upper amplitude ionic current time trace data (Fig 1.9 B), which correspond to the kinetic process of primer strand transfer between the polymerase and exonuclease active sites.
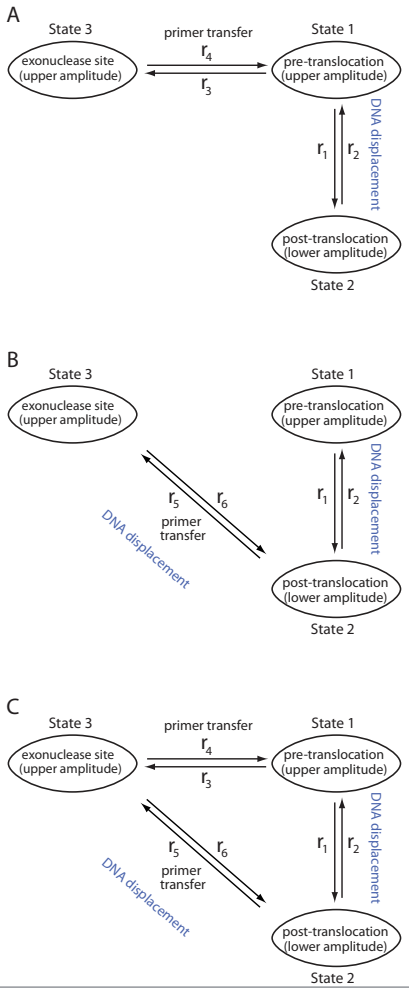
Figure 1.10: Potential models for the kinetic relationship between the translocation step and primer strand transfer between the polymerase and exonuclease sites. In all of the models, phi29 DNAP-DNA binary complexes fluctuate between the pre-translocation and post-translocation states with the forward rate r_1 and the reverse rate r_2. **A**: A three-state model in which transfer of the primer strand from the polymerase site to the exonuclease site initiates from the pre- translocation state (r_3), and the uncleaved primer strand returns from the exonuclease site to the pre-translocation state in the polymerase site (r_4). **B**: A three-state model in which transfer of the primer strand from the polymerase site to the exonuclease site initiates from the post- translocation state (r_5), and the uncleaved primer strand returns from the exonuclease site to the post-translocation state in the polymerase site (r_6). **C**: A three-state model in which transfer of the primer strand from the polymerase site to the exonuclease site can initiate from both the pre-translocation state and post-translocation states ($r_3 + r_5$), and the uncleaved primer strand can return from the exonuclease site to both the pre-translocation and post-translocation states in the polymerase site ($r_4 + r_6$).

In modeling the state observed in the upper amplitude for complexes of phi29 (D12A/D66A)DNAP-DNA1-H_OH it was essential to determine the kinetic pathway for primer strand transfer between the polymerase and exonuclease active sites, for complexes that do not undergo exonucleolysis, in relation to the translocation states in the polymerase active site. Again, we tested three distinct kinetic models to determine the pathway for primer strand transfer between the two active sites, with the specific aim of distinguishing whether primer strand transfer initiated in the lower amplitude, post-translocation state, or the upper amplitude, pre-translocation state. In the first kinetic model we proposed that primer strand transfer initiates in the polymerase active site in the pre-translocation state and that the uncleaved primer strand returns from the exonuclease active site to the pre-translocation state (Fig 1.10 A). In the second kinetic model we proposed that primer strand transfer initiates in the post-translocation state and that the primer terminus is returned from the exonuclease active site to the polymerase active site in the post-translocation state (Fig 1.10 B). The third model can be considered a combination of the two models, where primer strand transfer can initiate in the polymerase active site in either the pre or post-translocation states and the uncleaved primer strand can be returned to either state (Fig1.10 C)

To determine which of these kinetic models best fit the data we considered the rate describing escape from the exonuclease active site. In the first model the probability of escape from the exonuclease active site is described by *r4*

(Fig 1.10 A). In the second model the rate is described as $r6$, however rather than the primer strand being transferred to the pre-translocation state it is transferred to the post-translocation state (Fig 1.10 B). In the third model the path for escape is the sum of these two rates, $r4 + r6$ (Fig 1.10 C). Recall, the nanopore captures complexes and applies a force that opposes forward translocation (decreasing $r1$) and promotes reverse translocation (increasing $r2$). In the first kinetic model for primer strand transfer (Fig 1.10 A) the pathway is independent the direction of DNA displacement and should be unaffected by varying the voltage, and thus force, applied to the captured DNAP-DNA complex. Alternatively, the other two models (Fig 1.10 B & C) include pathways that are along the direction of DNA displacement across the translocation step, in these models the rates describing the kinetic pathway for primer strand transfer from the exonuclease active site, $r6$ and $r4 + r6$, should be sensitive to changes in the applied voltage. Each model proposed is consistent with two modes of exponential decay determined from dwell time samples of the upper amplitude (Fig. 1.9 B ii). Using a two exponential probability density function the values for the probability of escape from the state in which the primer strand occupies the exonuclease active site can be determined [68]. Here we established the theoretical value Q, which describes the rate(s) of primer strand transfer from the exonuclease active site to polymerase active site for each of the three models [68]. A plot of Q as a function of voltage would determine if the first model (Fig 1.10 A) fit the data

28

or if one of the other two models (Fig 1.10 B, C) were more accurate. For

(D12A/D66A)DNAP-DNA1-H_OH complexes the plot of Q as a function of

voltage (Fig 1. 11) indicated that the kinetic pathway for escape from the

exonuclease active site shows no systematic dependence on applied voltage,

across the voltages tested. Thus, the first model describing the kinetic

pathway for primer strand transfer between the two active sites best fit the

data [68].
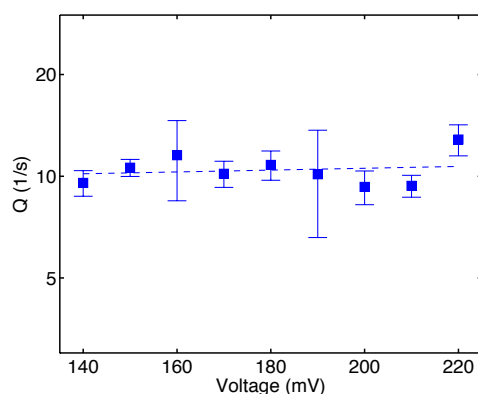
Figure 1.11: Distinguishing among the exonuclease kinetic models. Plot of Q vs voltage for complexes formed between the phi29 DNAP D12A/D66A mutant and DNA1H_OH. The quantity Q (derived in ref. 68) is calculated by fitting a model of two exponential modes to the dwell time samples of the upper amplitude state. The model of two exponential modes is consistent with all three kinetic models illustrated in Figure 10. The behavior of Q vs voltage, however, is expected to differ among the models, yielding the possibility of determining which model best explains the experimental observations. For the model in Figure 10A, Q is expected to be independent of the voltage, while for the models in Figure 10B, C; Q is expected to increase as voltage is reduced. The plot of Q vs voltage for the D12A/D66A complexes shows that Q is independent of the voltage, leading to the selection of the kinetic model in Figure 10A.

Using this kinetic model (Fig 1.10 A), where primer strand transfer between

the exonuclease and polymerase active sites initiates and ends in the

polymerase active site in the pre-translocation state, we developed a

mathematical method to extract the rates of translocation and primer strand transfer simultaneously [68]. We applied this method to extract the rates for complexes formed with (D12A/D66A)DNAP-DNA1-H_OH and (N62D)DNAP-DNA1-H_OH. The forward ($r1$) and reverse ($r2$) rates of translocation are, as expected, sensitive to the voltage being applied (Fig 1.12 A & B). However,
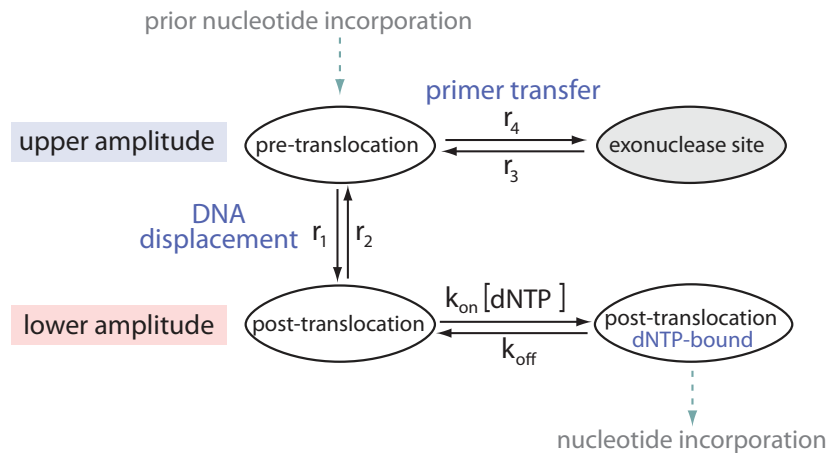
Figure 1.13: Kinetic transitions resolved in individual phi29 DNAP complexes captured atop a nanoscale pore. Four-state model for the reversible, noncovalent transitions that govern the fate of replicative DNAP complexes following each covalent nucleotide addition. The model shows the kinetic relationships among the steps of translocation, primer strand transfer between the polymerase and exonuclease active sites, and dNTP binding. Transfer of the primer strand from the polymerase to the exonuclease site precedes the forward translocation [68], and the forward translocation precedes dNTP binding; the translocation is rectified but not driven by dNTP binding, and dNTP has no influence on the rates across the translocation step [67]. The kinetic model comprises six transition rates: the rates of translocation (r_1, r_2), the rates of primer stand transfer between the pre-translocation state polymerase site and the exonuclease site (r_3, r_4), and the rates of dNTP binding to post-translocation state complexes (k_{on}[dNTP] and k_{off}). The mathematical framework based on the model allows these rates to be determined from experimental measurements [66, 67, 68]. Each of the three unshaded ellipses corresponds to a state (pre-translocation state, post-translocation state, or dNTP-bound post-translocation state) in which the primer strand, in duplex with the template strand, resides in the polymerase site; the ellipse that is shaded gray corresponds the state in which the primer terminus resides in the exonuclease site.
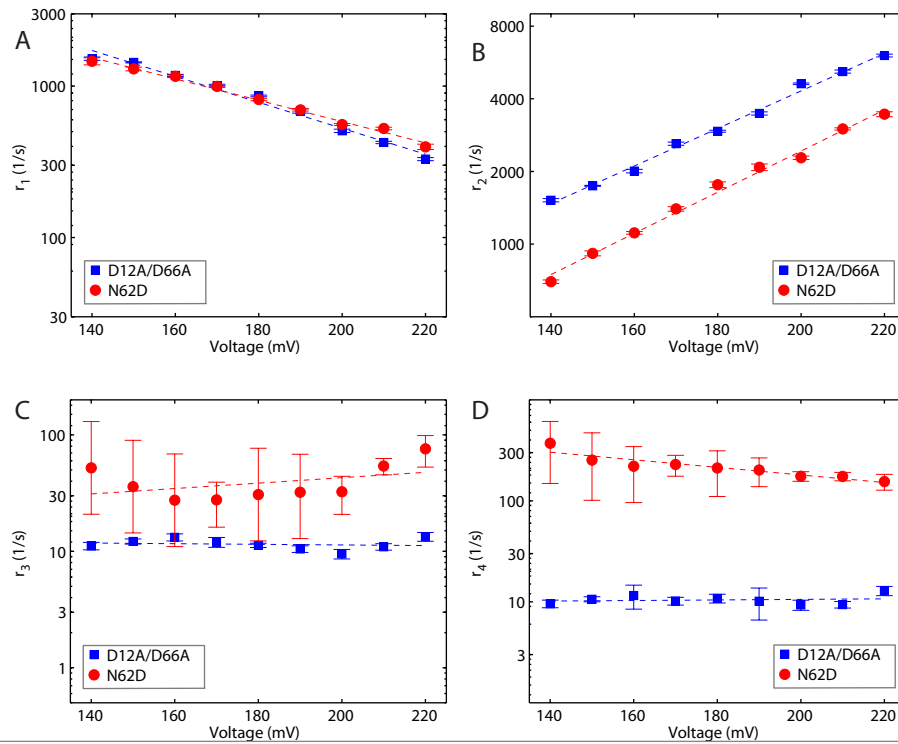
Figure 1.12: Rates of translocation and primer strand transfer between the polymerase and exonuclease sites. Plots of (**A**) $\log(r_1)$ vs voltage, (**B**) $\log(r_2)$ vs voltage, (**C**) $\log(r_3)$ vs voltage, and (**D**) $\log(r_4)$ vs voltage for complexes formed between DNA1-OH and the D12A/D66A (blue squares) and N62D (red circles) mutants of phi29 DNAP. Rates were determined from dwell time samples extracted from ionic current traces and the three-state model in Figure 10A. Errors bars indicate the standard error.

there is no systematic change in the rates of primer strand transfer from the polymerase to the exonuclease active site (*r3*) and from the exonuclease active site to the polymerase active site (*r4*) (Fig 1.12 C and D).

The kinetic models agree with the mechanism for editing, translocation and substrate selection we proposed in 2012 for complexes captured atop the α-HL nanopore (Fig 1.3) [24]. Considering the kinetic modeling described above we have developed a technology capable of measuring the kinetic mechanisms for primer strand transfer and nucleotide binding with regard to

the translocation step, something that has yet to be achieved by other technologies. Establishing that these kinetic mechanisms are separated by the translocation step (Fig 1.13) enabled us to measure the non-covalent kinetic rates associated with proofreading, translocation and substrate selection. This kinetic framework allowed us to ask mechanistic questions regarding the role of the translocation step as the branch point between DNA synthesis and editing activities.

Application of nanopore technology to probe the mechanisms of DNAP fidelity. The single molecule technique we developed was utilized to describe several mechanisms of DNAP fidelity. After developing the ability to measure translocation and substrate binding we applied this method to explore the role of highly conserved residues identified by ensemble and structural studies as essential in the mechanism for translocation and replication fidelity, Y226 and Y390 of phi29 DNAP [25]. The findings of this study are presented in chapter two. After developing the ability to additionally measure the rates of primer strand transfer between the polymerase and exonuclease active sites we applied this technique to determine the mechanisms for stable incorporation of ribonucleotides into DNA by replicative DNAPs during DNA replication, using phi29 DNAP as a model enzyme [26]. The finding of this study are presented in chapter three. Finally, we described the role of divalent metal cations in the non-covalent kinetic transitions that reside at the branch point between DNA synthesis and editing in replicative DNAPs [27]. We specifically

tested Mg^{2+} and Mn^{2+}, which promote chemical transformation of the DNA substrate by the polymerase and exonuclease active sites, and Ca^{2+}, which promotes nucleotide binding but not chemical transformation of the DNA substrate. The findings of this study are presented in chapter four.

Chapter 2

Dynamics of Translocation and Substrate Binding in Individual Complexes Formed with Active Site Mutants of Phi29 DNA Polymerase

2.1 Abstract

The phi29 DNA polymerase (DNAP) is a processive B-family replicative DNAP. Fluctuations between the pre-translocation and post-translocation states can be quantified from ionic current traces, when individual phi29 DNAP-DNA complexes are held atop a nanopore in an electric field. Based upon crystal structures of the phi29 DNAP-DNA binary complex and the phi29 DNAP-DNA-dNTP ternary complex, residues Y226 and Y390 in the polymerase active site were implicated in the structural basis of translocation. Here, we have examined the dynamics of translocation and substrate binding in complexes formed with the Y226F and Y390F mutants. The Y226F mutation diminished the forward and reverse rates of translocation, increased the affinity for dNTP in the post-translocation state by decreasing the dNTP dissociation rate, and increased the affinity for pyrophosphate in the pre-translocation state. The Y390F mutation significantly decreased the affinity for

dNTP in the post-translocation state by decreasing the association rate $\sim$ 2-fold and increasing the dissociation rate $\sim$ 10-fold, implicating this as a mechanism by which this mutation impedes DNA synthesis. The Y390F dissociation rate increase is suppressed when complexes are examined in the presence of Mn^{2+} rather than Mg^{2+}. The same effects of the Y226F or Y390F mutations were observed in the background of the D12A/D66A mutations, located in the exonuclease active site, $\sim$ 30 Angstroms from the polymerase active site. While translocation rates were unaffected in the D12A/D66A mutant, these exonuclease site mutations caused a decrease in the dNTP dissociation rate, suggesting that they perturb phi29 DNAP interdomain architecture.

2.2 Introduction

Replicative DNA polymerases (DNAPs) are molecular motors that translocate along their DNA substrates in single nucleotide increments as they catalyze template-directed DNA replication. The DNAP from the bacteriophage phi29 is a B-family polymerase that catalyzes highly processive DNA synthesis [12, 13, 85], without the need for accessory proteins such as sliding clamps or helicases, because it remains tightly associated with its DNA substrate and promotes downstream strand displacement during replication [12, 6, 73]. In addition to its 5'-3' polymerase active site, phi29 DNAP has a 3'-5' exonuclease active site, located in a separate domain of the protein, $\sim$ 30 Å from the polymerase active site [13, 85, 6, 58].

35
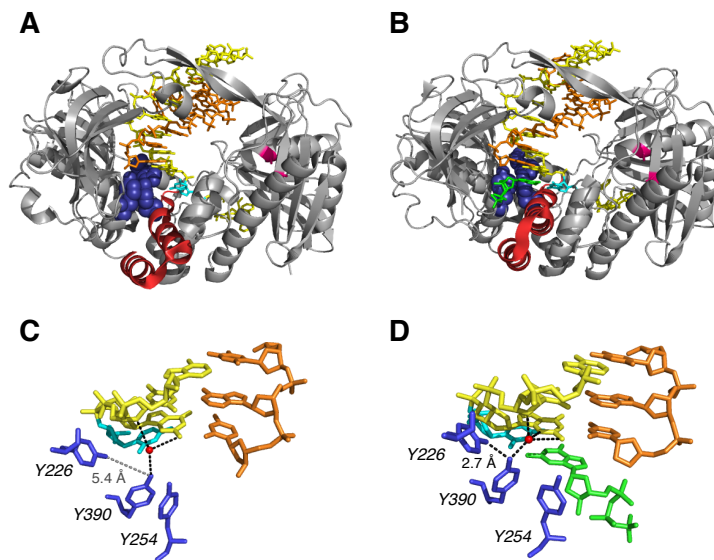
Figure 2.1: Structural transitions in phi29 DNAP-DNA complexes critical to the translocation step and to dNTP binding. Crystal structure models for (**A**) the phi29 DNAP-DNA post-translocation state binary complex, in the fingers-open conformation (PDBIS: 2PZS), and (**B**) the phi29 DNAP-DNA-dNTP, post-translocation state ternary complex, in the fingers-closed conformation (PDBID: 2PYJ). Panels (**C**) and (**D**) are close-up views of the polymerase active site, from the structures shown in panels A and B, respectively. The structures are from reference (4), and were determined using the D12A/D66A mutant of phi29 DNAP. In (A and B), the protein backbone is rendered as a grey ribbon, with residues 359-395 in the fingers domain in red ribbon to highlight the conformation difference between the open binary complex and the closed ternary complex. The backbone positions of the D12 and D66 residues in the exonuclease domain are colored magenta. In (A-D), the DNA primer strand is displayed in orange, the DNA template strand is yellow, with the templating base at n=0 in cyan. Residues Y254, Y226, and Y390 are rendered in blue (space-filling in A and B, sticks in C and D). In (B and D) the incoming dNTP is shown in green. In (A and C) the side chains of Y254 and Y390 are stacked, obstructing the dNTP binding site; in (B and D) both tyrosine side chains are rotated out of the stacking interaction, removing the steric impediment to the incoming dNTP. In (C and D), the water molecule that mediates the interaction of the hydroxyl group of Y390 with the -1 and -2 residues of the template strand of the duplex is shown as a red sphere. This water is part of an extensive network of water-mediated interactions with the minor groove of the active-site proximal duplex, a network that is precisely conserved between phi29 DNAP and the B-family DNAP from bacteriophage RB69 [100]. The black dashed lines indicate potential H-bonding interactions for the hydroxyl groups of the Y226 or Y390 side chains, including the H-bond between the two side chains (labelled 2.7 Å in panel D). In (C), the dashed grey line between the hydroxyl groups of the Y226 and Y390 side chains in the binary complex illustrates the increased distance between the hydroxyl groups of Y226 and Y390 (> 5 Å) when the fingers are in the open conformation.

Crystal structures of the phi29 DNAP binary complex with a primer-template DNA substrate bound in the polymerase active site (Fig. 2.1 A) and of the phi29 DNAP-DNA ternary complex with dNTP complementary to the templating base in the active site (Fig. 2.1 B) have been determined [6]. The architecture of the DNA polymerase domain is highly conserved and resembles a partially closed right hand. The palm sub-domain contains residues that participate in the chemistry of catalysis, whereas the thumb subdomain positions the primer-template duplex in the active site. The fingers subdomain contains residues essential for binding incoming nucleotide substrates. In crystal structures of complexes containing complementary dNTP, the position of the fingers subdomain differs from its position in the binary complex structures; elements of this sub-domain move in toward the active site cleft to achieve a tight steric fit with the nascent base pair (Fig. 2.1 A and 2.1 B).

In the fingers-open, post-translocation state binary complex, the side chains of Y254 and Y390 in the polymerase active site are stacked, in a conformation that sterically occludes dNTP binding (Fig. 2.1 C). In the fingers-closed ternary complex, the side chains of Y254 and Y390 both rotate relative to their positions in the fingers-open binary complex, disrupting the stacking interaction between them and allowing the incoming dNTP to bind. The deoxyribose sugar of the dNTP stacks on Y254, and the rotation of Y390 brings its hydroxyl group into H-bonding distance of the hydroxyl group of
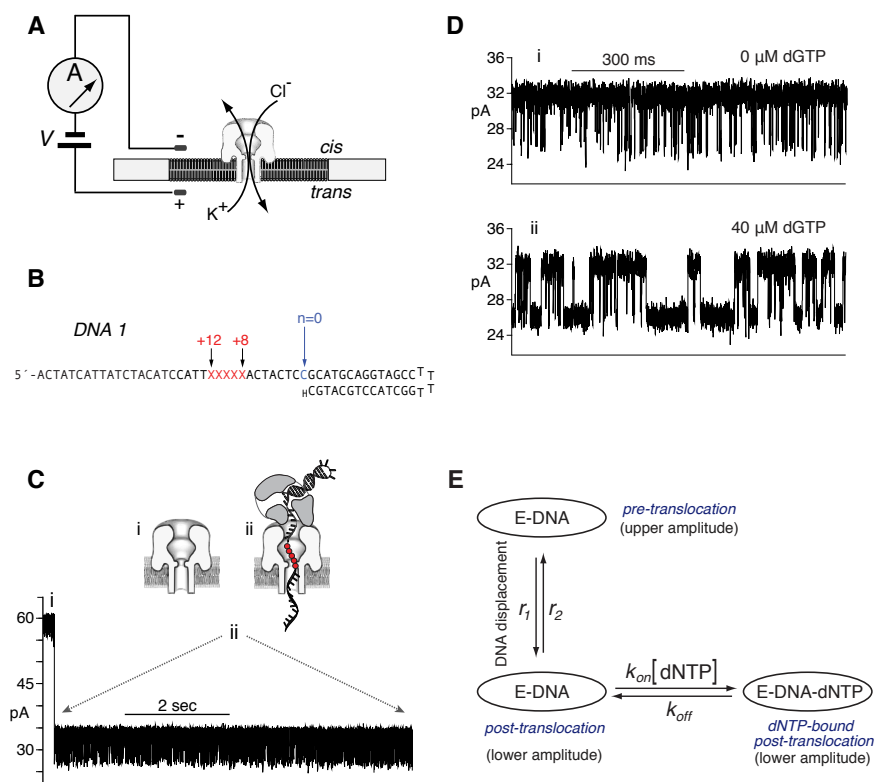
Figure 2.2: Capture of phi29 DNAP-DNA complexes on the α-HL nanopore. In the nanopore device (**A**), a single α-HL nanopore is inserted in a ∼ 25 µm diameter lipid bilayer separating two chambers (cis and trans) that contain buffer solution. A patch clamp amplifier applies voltage across the bilayer and measures ionic current, which is carried through the nanopore by K^+ and Cl^- ions. (**B**) DNA1 is a hairpin, featuring a 14-base pair duplex and a single-stranded template region of 35 nucleotides. The primer strand is terminated with a 2′-H, 3′-H CMP residue, and the template strand contains a reporter group of five consecutive abasic (1′-H, 2′-H) residues spanning positions +8 to +12 (indicated as red X's in the sequence). (**C**) Representative current trace for a binary complex formed between phi29 DNAP and the DNA1 substrate, captured at 180 mV applied potential in buffer containing 10 mM K-Hepes, pH 8.0, 0.3 M KCl, 1 mM EDTA, 1 mM DTT, and 11 mM $MgCl_2$. DNA and phi29 DNAP were added to the nanopore cis chamber to final concentrations of 1 and 0.75 µM, respectively. Cartoons above the current trace illustrate the sequence of events, which is described in the text. In the cartoons, the five consecutive abasic (1′, 2′-H) residues spanning positions +8 to +12 of the template strand, which serve as a reporter group, are shown as red circles. (**D**) Ionic current traces for phi29 DNAP=DNA1 complexes, captured at 180 mV in the presence of 0 µM (i) or 40 µM (ii) dGTP. (E) A three-state model in which translocation and dNTP binding are sequential: dNTP can bind to complexes (k_{on} [dNTP]) only after the transition from the pre-translocation to the post-translocation state (r_1); the transition from the post-translocation to the pre-translocation state (r_2) cannot occur before the dissociation of dNTP (k_{off}).

Y226 (Fig. 1D).

These key structural differences between the open and closed complexes prompted the proposal of an elegant model for the structural mechanism of translocation [6], in which the spatial displacement of translocation is directly linked to the structural transition of fingers opening. In this view, the fingers-closed post-translocation state ternary complex serves as a model for the structure of the pre-translocation state complex, in which the nascent base pair between the templating base at n=0 and the incoming complementary dNTP in the post-translocation state ternary complex occupies the site that would be occupied by the terminal base pair of the primer-template duplex in the pre-translocation state complex. The structure of the phi29 DNAP binary complex in the fingers-open, post-translocation state indicates that the pre-translocation state in the fingers-open conformation is sterically precluded. Specifically, the orientation of Y390 and Y254 would clash with the terminal base-pair of the duplex. Hence, fingers opening was proposed to compel the forward translocation [6].

Fluctuations between the pre-translocation and post-translocation states can be directly observed and quantified from ionic current time traces recorded when individual phi29 DNAP-DNA complexes are held atop a nanoscale pore in an electric field [24, 66, 67]. A single α-hemolysin (α-HL) nanopore is inserted into a lipid bilayer that separates two chambers (termed cis and trans) containing buffer solution (Fig. 2.2 A). A patch clamp amplifier applies voltage

across the bilayer, and measures the ionic current that flows through the nanopore, which is carried by K+ and Cl- ions in the buffer. A typical ionic current trace that results when a binary complex between phi29 DNAP and a DNA substrate (DNA1; Fig. 2.2 B) is captured atop the nanopore at 180 mV applied potential is shown in Fig. 2.2 C. The ionic current through the open pore (Fig. 2.2 C, i) drops rapidly when a complex is captured (Fig. 2.2 C, ii). The enzyme is too large to enter the nanopore. Thus, the phi29 DNAP-DNA complex, with the enzyme bound at the primer-template junction of the DNA substrate, perches atop the pore. The DNA template strand of the captured complex is suspended through the nanopore lumen, which is just wide enough to accommodate a single strand of DNA (Fig. 2.2 C, ii).

Captured phi29 DNAP-DNA complexes reside atop the nanopore for several seconds, during which the measured ionic current fluctuates on the millisecond time scale between two amplitude levels (Fig. 2.2 C, ii). Transition between the two amplitudes corresponds to movement of the DNA substrate relative to the enzyme and the nanopore; the distance of this displacement is ~ one nucleotide [24, 65]. Detection of the DNA displacement is achieved by the use of a reporter group comprising five consecutive abasic (1′-H, 2′-H) residues in the template strand (red Xs or red circles, in Fig. 2.2 B and 2.2 C, ii, respectively); a displacement of the reporter group in the nanopore lumen is manifested as a change in measured ionic current [24, 65]. In the upper amplitude, the primer-template junction of the DNA substrate is bound in the

40

polymerase active site, in the pre-translocation state. At 180 mV, the pre-translocation state amplitude is centered at ~ 32 pA (Fig. 2.2 C, ii; Fig. 2.2 D, i). In the lower amplitude, the primer-template junction of the DNA substrate resides in the polymerase active site, in the post-translocation state. The post-translocation state amplitude is centered at ~ 26 pA at 180 mV (Fig. 2.2 C, ii; Fig. 2.2 D, i). The amplitude fluctuations continue until complexes dissociate or are ejected, after which another complex can be captured. We have shown that the pre-translocation and post-translocation states are discrete kinetic states [66].

The primer strand of DNA1 bears a $2'$-H, $3'$-H terminus (Fig. 2.2 B) and thus DNA1 supports the formation of phi29 DNAP-DNA-dNTP ternary complexes but not the chemical step of phosphodiester bond formation. Binding of dGTP (complementary to the template dCMP residue at n=0) to phi29 DNAP-DNA1 complexes stabilizes the post-translocation state. In the absence of dNTP, complexes fluctuate rapidly between the two states (Fig. 2.2 D, i); the addition of dGTP (40 μM; Fig. 2.2 D, ii) causes the average dwell time in the lower amplitude, post-translocation state to increase, as a subpopulation emerges with longer dwell times. The kinetic mechanism of translocation and dNTP binding in individual phi29 DNAP-DNA complexes [67] is described by a three-state model with four transition rates (Fig. 2.2 E). In the three-state model, translocation and dNTP binding are sequential: dNTP can bind to complexes (*kon*[dNTP]) only after the transition from the pre-translocation to the post-

translocation state (*r1*); the transition from the post-translocation to the pre-translocation state (*r2*) cannot occur before the dissociation of dNTP (*koff*) (Fig. 2.2 E).

In the current study, we have examined the dynamics of translocation and substrate binding in individual complexes formed with the Y226F and Y390F mutants of phi29 DNAP. Both Y226 and Y390 are highly conserved residues in B-family DNAPs [11, 95]. Changing either Y226 or Y390 to phenylalanine disrupts the H-bond between their hydroxyl groups, an interaction which may have a role in stabilizing the orientation of Y390 in the fingers-closed, post-translocation ternary complex, or in the proposed fingers-closed, pre-translocation state complex [6]. This H-bond is the only structural interaction predicted to be directly affected by the Y226F mutation, although indirect effects on active site structure, including perturbations of other interactions in which Y390 is a partner, cannot be excluded. The Y390F mutation directly disrupts the H-bonding potential with Y226 in the closed complex, as well as the water-mediated interactions of Y390 with the template strand in the DNA duplex in both the open and closed complexes. Earlier biochemical studies of the Y390F mutant have shown that it is severely impaired in DNA synthesis relative to the wild type phi29 DNAP [11, 14, 86]. In contrast to Y390F, biochemical studies of the Y226F mutant showed that while it is impaired in transfer of the primer strand from the polymerase to exonuclease active site for DNA substrates bearing fully-paired duplexes [95], it is not impaired in DNA
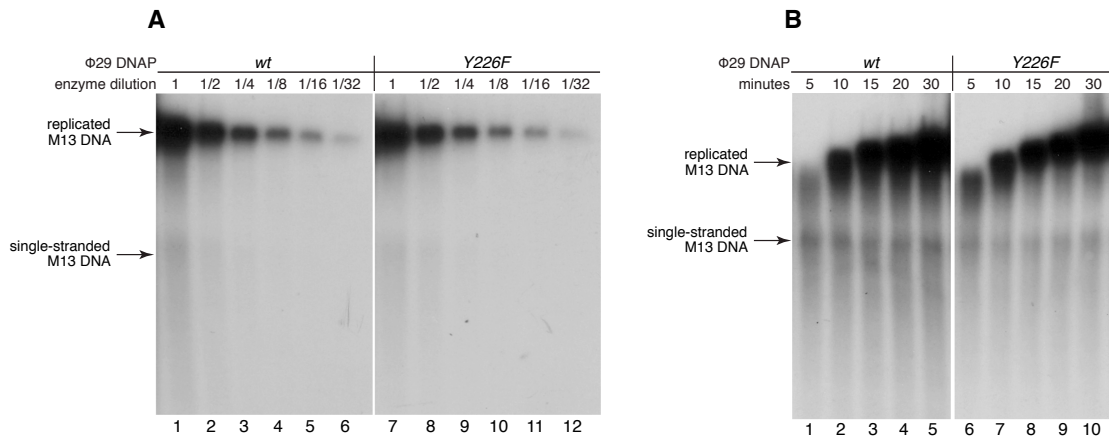
Figure 2.3: Processive DNA synthesis catalyzed by the Y226F mutant. (**A**) DNA synthesis catalyzed by the wild type phi29 DNAP (lanes 1-6) or the Y226F mutant (lanes 7-12) as a function of enzyme concentration. A series of two-fold serial dilutions of each enzyme was tested, in which the highest concentration (lanes 1 and 7, indicated by enzyme dilution = 1) was 30 pM; the reactions were conducted at 30°C for 30 minutes. (**B**) DNA synthesis as function of time, for reactions catalyzed by the wild type ø29 DNAP (lanes 1-5) or the Y226F mutant (lanes 6-10). Reactions were conducted using an enzyme concentration of 60 pM at 30°C for the indicated times. In both (A) and (B), the replication substrate was oligonucleotide-primed bacteriophage M13 single-stranded DNA (~ 3.35 pM). The reaction products were resolved by electrophoresis in alkaline agarose gels.

synthesis (Fig. 2.3 and ref. 95). Thus, the Y390F and Y226F mutations yield very different biochemical properties in phi29 DNAP.

2.3 Experimental Methods

DNA and enzymes. DNA1 was synthesized at Stanford Protein and Nucleic Acid Facility and purified by denaturing PAGE. The DNA1 hairpin was annealed by heating at 90°C for 4 min followed by snap cooling in ice water.

Wild type phi29 DNAP was obtained from Enzymatics (Beverly, MA). The D12A/D66A mutant was obtained from XPol Biotech (Madrid, Spain). The Y226F, Y390F, Y226F/D12A/D66A, and Y390F/D12A/D66A mutants were expressed in *Escherichia coli* BL21(DE3) cells and purified as described for

the wild type phi29 DNAP [69].

Nanopore methods. Nanopore experiments were conducted as described [24, 65, 66, 67, 68, 5, 2, 41]. Briefly, a single α-HL nanopore is inserted in a ~ 25 µm-diameter lipid bilayer that separates two chambers (*cis* and *trans*) containing buffer solution (10 mM K-Hepes, pH 8.0, 0.3 M KCl, and 1 mM EDTA). DTT was added to the nanopore *cis* chamber to a final concentration of 1 mM. $MgCl_2$ and ddCTP were added to final concentrations of 11 mM and 400 µM, respectively, except in those experiments in Fig. 2.7, in which the effects of Mn^{2+} were assayed. In those experiments, $MgCl_2$ and ddCTP were omitted and $MnCl_2$ was added to the *cis* chamber to a final concentration of 2 mM. DNA and phi29 DNAP were added to the *cis* chamber to final concentrations of 1 and 0.75 µM, respectively; and dGTP or pyrophosphate were added as indicated in the text and figures. Ionic current was measured with an integrating patch clamp amplifier (Axopatch 200B, Molecular Devices) in voltage clamp mode. Data were sampled using an analog-to-digital converter (Digidata 1440A, Molecular Devices) at 100 kHz in whole-cell configuration and filtered at 5 kHz using a low pass Bessel filter.

Data analysis. The value of p, the probability of the lower amplitude state, was determined from histograms of all sampled amplitude data points, generated with Clampfit software (Molecular Devices) at 0.2 pA bin width [24]. Histograms were fit to a two-term Gaussian function using the Levenberg-Marquardt search algorithm provided in Clampfit.

44

In the methods for extracting translocation rates and dNTP binding rates, segments of ionic current time traces were exported from Clampfit to Matlab (MathWorks); the centers of the two amplitude clusters and their probabilities were accurately calculated using the maximum likelihood estimation on samples of measured amplitudes. For each ionic current time trace, the autocorrelation was calculated and fitted to its expected expression derived from the mathematical model. For experiments conducted in the absence of dNTP, a two-state kinetic model was used to describe the transitions between the two translocation states [66]. In the presence of dNTP, a three-state kinetic model was employed to accommodate the additional dNTP bound state [67]. Kinetic rates were reconstructed by combining the results of maximum likelihood estimation on amplitude samples and the results of fitting the autocorrelation of the time trace to the corresponding model [66, 67]. At each experimental condition, we used 15 ~ 30 time traces. For each kinetic rate, the reported value and standard error were calculated based on values for these individual time traces.

Extraction of dwell time samples. The dwell time samples in each of the two amplitude states used in the analyses in Fig. 2.8 B and Table 2.3 were extracted using the single-channel detection function in Clampfit 10 (Molecular Devices). This software uses a half-amplitude threshold method to assign transitions between two user-defined amplitude levels [23]. Amplitude levels for each of the two states were determined for the single-channel

searches from histograms of all sampled amplitude data points. Complexes were captured at 160 mV.

Processive DNA synthesis assays. The incubation mixture contained 50 mM Tris-HCl, pH 7.5, 10 mM MgCl$_2$, 1 mM DTT, 4% (v/v) glycerol, 0.05% (v/v) Tween 20, 0.1 mg/ml BSA, 40 μM each dNTP and [α-^{32}P]dATP (0.3 μCi), 100 ng of oligonucleotide-primed M13 ssDNA (~ 3.35 pM in reaction mixture) and 30 ng wild type or mutant Y226F phi29 DNAP (dilution 1; 30 pM in reaction mixture), in a final volume of 15 μl. Two-fold serial dilutions of the enzymes were carried out as indicated. After incubation for the indicated times at 30°C, reactions were terminated by the addition of 30 mM EDTA/0.5% SDS. The DNA was denatured by addition of 0.4 M NaOH and subjected to alkaline electrophoresis in 0.7% agarose gels. After electrophoresis the gels were dried and autoradiographed.

2.4 Results and Discussion

Residues Y226 and Y390 in the polymerase active site of phi29 DNAP are implicated in structural transitions essential to both the translocation step and to dNTP binding [6, 95, 14, 86]. In this study, we examined the effects of introducing the Y226F or Y390F mutations on these DNAP functions, in two contexts: 1) as single amino acid changes in the otherwise wild type phi29 DNAP, and 2) in combination with the D12A/D66A mutations in the exonuclease domain active site. We first determined the effects of the Y226F and Y390F mutations on the dynamics of the translocation step (Fig. 2.4). We

46

used a robust method that employs autocorrelation and a two-state model (8) to extract the forward (r_1; Fig. 2.4 A) and reverse (r_2; Fig. 2.4 B) translocation rates from ionic current traces recorded when individual phi29 DNAP-DNA binary complexes reside atop the nanopore.
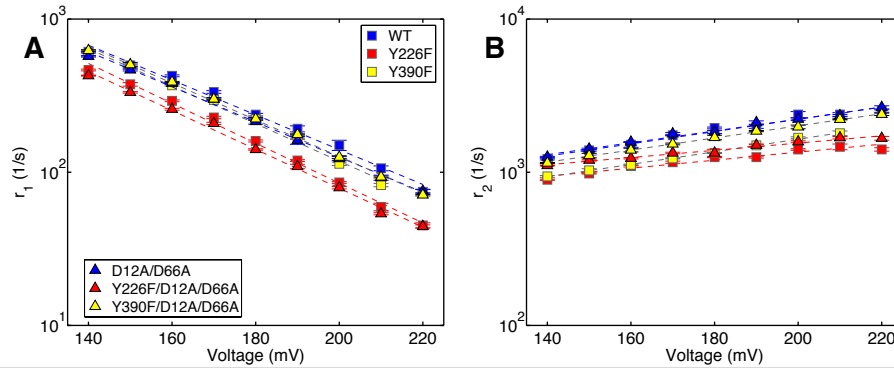
Figure 2.4. Transition rates of the phi29 DNAP translocation step extracted from ionic current traces measured in absence of dNTP. Plots of (**A**) log(r_1) versus voltage, and (**B**) log(r_2) versus voltage, for binary complexes formed between the wild type (blue squares), Y226F (red squares), Y390F (yellow squares), D12A/D66A (blue triangles), Y226F/D12A/D66A (red triangles) or Y390F/ D12A/D66A (yellow triangles) phi29 DNAP and DNA1. Each plotted point shows the mean ± standard error. In the absence of dNTP, the fluctuation rates between the pre-translocation and post-translocation states are fully described by a two-state model with two transition rates (8). Each plotted data point shows the mean ± standard error, determined from 15-30 ionic current time traces for individual captured complexes; each time trace had a duration of 5-10 seconds.

We compared the translocation rates for wild type, Y226F, Y390F, D12A/ D66A, Y226F/D12A/D66A, and Y390F/D12A/D66A phi29 DNAP enzymes, as a function of applied force (voltage). The force opposes the forward translocation and thus decreases its rate (r_1), while promoting the reverse translocation and increasing its rate (r_2). The slope of log(r_1) vs. voltage is negative and proportional to the distance between the pre-translocation state and the transition state; the slope of log(r_2) vs. voltage is positive and

proportional to the distance between the transition state and the post-translocation state. Plots of log(r_1) vs. voltage and log(r_2) vs. voltage show that the Y226F and Y390F mutations have modest effects on the translocation rates, in the context of either the wild type or D12A/D66A backgrounds (Fig. 2.4); neither the vertical intercepts nor the slopes of log(rate) vs. voltage exhibit large differences from those for the wild type enzyme, indicating that the mutations do not significantly change the energy landscape for the translocation step [66]. The rates for complexes formed with each of the six enzymes and captured at 180 mV are given in Table 2.1.

Table 2.1: Translocation rates for wild type phi29 DNAP and mutants at 180 mV.

Enzyme [a]	r_1 (s^{-1}) [b]	r_2 (s^{-1}) [c]
Wild type	237.57 ± 3.88	1945 ± 21.5
Y226F	159.95 ± 2.06	1253.9 ± 9.7
Y390F	221.43 ± 2.33	1334.4 ± 9.0
D12A/D66A	214.36 ± 8.56	1830.7 ± 32.2
D12A/D66A (in **Mn^{2+}**)	40.05 ± 0.88	1741.9 ± 23.8
Y226F/D12A/D66A	140.71 ± 1.81	1334.6 ± 77.3
Y390F/D12A/D66A	223.08 ± 4.6	1686.6 ± 58.8
Y390F/D12A/D66A (in **Mn^{2+}**)	40.43 ± 1.74	1252.6 ± 27.6

[a] All enzymes are examined in Mg^{2+} unless otherwise specified.
[b] The forward translocation rate.
[c] The reverse translocation rate.
All values are reported for data collected at 180 mV and are given with the standard error.

Across the range of voltages, the rates of the forward and reverse fluctuations

across the translocation step for the D12A/D66A mutant are almost indistinguishable from those of the wild type enzyme (Fig. 2.4). The most notable effects on the translocation rates are caused by the Y226F mutation. Across the range of voltages, the forward translocation rates for the Y226F and Y226F/D12A/D66A mutants are ~ 35% slower than those of the wild type and D12A/D66A enzymes, respectively (Fig. 2.4). The reverse translocation rates are also reduced by the Y226F mutation; r_2 for the Y226F enzyme is ~ 35% slower than it is for the wild type, and r_2 for the Y226F/D12A/D66A is ~ 15 % slower than for the D12A/D66A enzyme (Fig. 2.4). The moderate effects on r_1 and r_2 of introducing the Y226F or Y390F mutations can be contrasted with the significant effects of active-site proximal DNA substrate sequences on rates r_1 and r_2 in binary complexes formed with the wild type enzyme. For example, sequence changes at n=0 of the template strand, or in the -2 and -3 base-pairs of the duplex can yield an ~ 10-fold increase in r_1, and an ~ 15-fold decrease in r_2, relative to DNA1 [66].

Complementary dNTP binding to the mutant enzymes. We determined the effects of introducing the Y226F or Y390F mutations on complementary dNTP binding affinity in titration experiments, using complexes formed between each of the six phi29 DNAP enzymes and DNA1, captured at 180 mV. To display these data, we plotted the normalized p/(1-p), where p is the probability of post-translocation state occupancy, and the normalized p/(1-p) is defined as the value of p/(1-p) in the presence of a given concentration of

dNTP, divided by the value of p/(1-p) for the same phi29 DNAP-DNA complex at 0 µM dNTP [24]. The normalized p/(1-p) is solely determined by the binding affinity of dNTP; it is independent of the transitions between the two translocation states in the absence of dNTP (and thus independent of any differences in the translocation rates among the six enzymes); the effect of these transitions is eliminated when p/(1-p) is normalized by its value measured in the absence of dNTP. The normalized p/(1-p) thus permits direct comparison of the post-translocation state dNTP binding affinities among the enzymes.

The Y226F and Y390F mutants differ dramatically in dNTP binding affinity. Relative to the wild type enzyme, the Y266F mutation increases the affinity for dNTP, while the Y390F mutation significantly decreases it (Fig. 2.5 A). We also examined the dNTP binding affinities for Y226F and Y390F mutations in the D12A/D66A background. Fig. 2.5 B shows the results of dGTP titration experiments, for complexes formed with the wild type, D12A/D66A, Y226F, Y390F, Y226/D12A/D66A and Y390F/D12A/D66A enzymes. Because of the wide range of dGTP concentrations used in the experiments, and the differences in dNTP affinities among the enzymes, the normalized p/(1-p) as a function of [dGTP] for the six enzymes is most clearly compared on a log scale plot of (normalized p/(1-p)) -1 (Fig. 2.5 B). Interestingly, the D12A/D66A mutations themselves increase the affinity for dNTP relative to that of the wild type phi29 DNAP (Fig. 2.5 B). Residues D12 and D66 are located in the

exonuclease domain, ~ 30 Å from the polymerase active site (Fig. 2.1 A and 2.1 B). Nonetheless, these mutations in the exonuclease active site have been shown to disrupt the ability of phi29 DNAP to perform downstream strand displacement during DNA synthesis [73, 36].
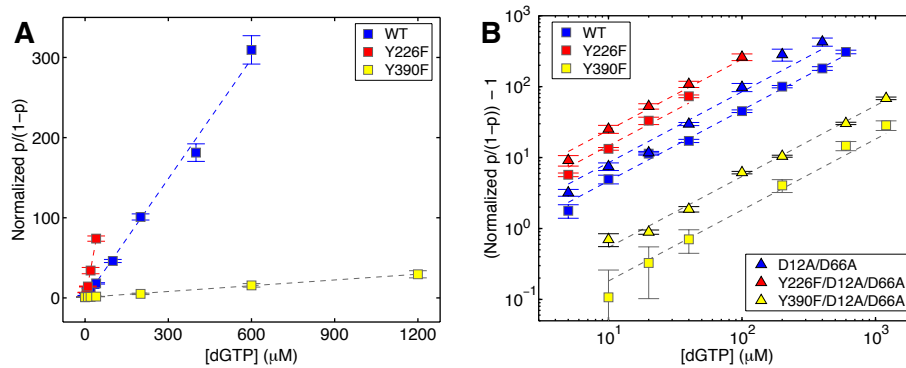
Figure 2.5: Complementary dNTP binding affinities of phi29 DNAP mutants. The normalized p/(1-p) (where p is the probability of post-translocation state occupancy, and the normalized p/(1-p) is defined as the value of p/(1-p) in the presence of a given concentration of dNTP, divided by the value for p/(1-p) for the same phi29 DNAP-DNA complex at 0 µM dGTP (7)) is plotted (**A**) as a function of dGTP concentration for complexes formed between wild type, Y226F, and Y390F phi29 DNAP. In (**B**), the (normalized p/(1-p)) -1 is plotted on a log scale as a function of dGTP concentration, for complexes formed between the wild type, Y226F, Y390F, D12A/D66A, Y226F/D12A/D66A or Y390F/D12A/D66A phi29 DNAP and DNA1. Complexes were captured at 180 mV. Plot symbols for each of the enzymes are given in the legend for Fig. 4. Errors bars indicate the standard error. Each data point was determined from 15-30 ionic current time traces for individual captured complexes; each time trace had a duration of 5-10 seconds.

The effect of the D12A/D66A mutations on dNTP binding, taken together with the effects of these mutations on strand displacement, suggest that the exonuclease active site mutations perturb phi29 DNAP interdomain architecture, yielding pleiotropic effects on enzyme function. When combined with either the Y226F or the Y390F mutations, the D12A/D66A mutations yield
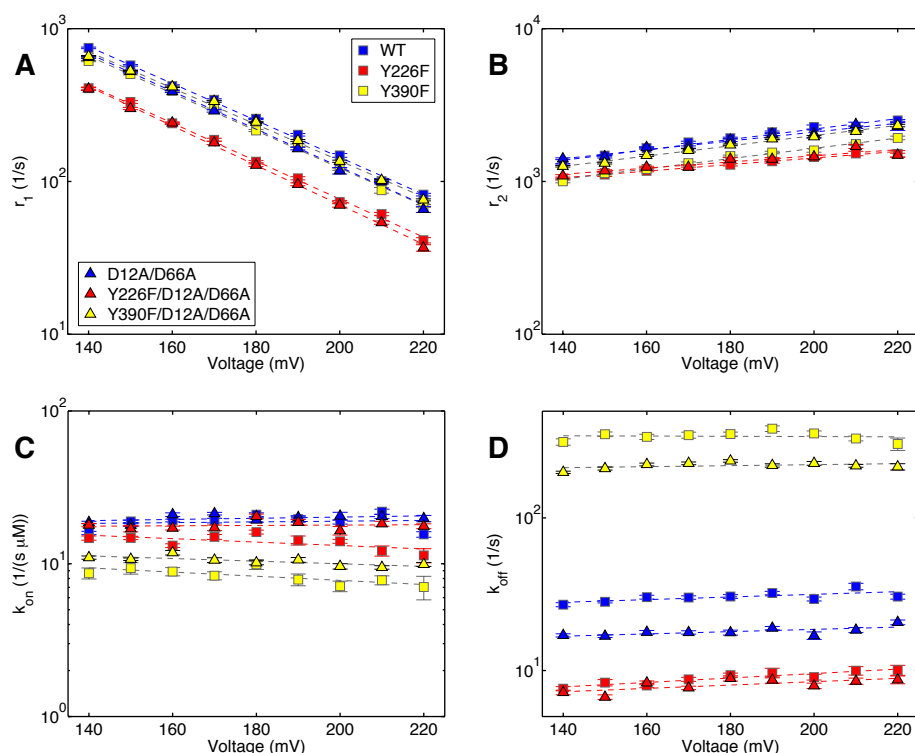
Figure 2.6: Translocation rates and dNTP association and dissociation rates determined simultaneously from ionic current traces measured in the presence of dNTP. Plots of (**A**) $\log(r_1)$ vs. voltage and (**B**) $\log(r_2)$ vs. voltage for complexes formed between the wild type, Y226F, Y390F, D12A/D66A, Y226F/D12A/D66A and Y390F/D12A/D66A phi29 DNAP with DNA1, captured in the presence of dGTP. Plots of (**C**) k_{on} vs. voltage and (**D**) k_{off} vs. voltage for complexes formed between the wild type, Y226F, Y390F, D12A/D66A, Y226F/D12A/D66A and Y390F/D12A/D66A phi29 DNAP with DNA1, captured in the presence of dGTP. Rates were extracted from ionic current traces using the autocorrelation function and the three-state model shown in Fig. 2E. Plot symbols for each of the enzymes are given in the legend for Fig. 4. Errors bars indicate the standard error. Each data point was determined from 15-30 ionic current time traces for individual captured complexes; each time trace had a duration of 5-10 seconds. The data plotted are for complexes captured in the presence of the following dGTP concentrations: wild type, 10 μM; D12A/D66A, 10 μM; Y226F, 5 μM; Y390F, 20 μM; Y226F/D12A/D66A, 5 μM; and Y390F/D12A/D66A, 20 μM. While we have shown that all four transition rates (r_1, r_2, k_{on}, and k_{off}) are independent of [dNTP] [67], the method of extracting the rates from the ionic current traces using autocorrelation and the three-state model (Fig. 2E) is most robust when using data collected under conditions where all three states are well sampled. For example, when the dNTP concentration is very low, the dGTP bound state is not well sampled; when the dNTP concentration is very high, only the dGTP bound state is well sampled. The dNTP concentrations optimal for the analysis vary with the dNTP binding affinity of each mutant.

an increase in affinity over that of the polymerase site mutation alone (Fig. 5B).

Dynamics of dNTP binding to the mutant enzymes. To uncover the kinetic mechanisms by which the mutations in the phi29 DNAP polymerase and exonuclease active sites perturb dNTP binding, we used autocorrelation and a three-state model (Fig. 2.2 E) to simultaneously extract the forward (r_1) and reverse (r_2) translocation rates, and the dNTP association (k_{on}) and dissociation (k_{off}) rates, from ionic current traces for complexes captured in the presence of complementary dNTP [66]. For each of the mutant enzymes, the plots of log(r_1) vs. voltage (Fig. 2.6 A) and log(r_2) vs. voltage (Fig. 2.6 B) obtained for complexes captured in the presence of dNTP are indistinguishable from those of the same enzyme in the absence of dNTP (Fig. 2.4 A and 2.4 B). Thus, as we have shown for wild type phi29 DNAP [66], the rates of fluctuation across the translocation step for the mutants are independent of [dNTP]; dNTP can only bind onto phi29 DNAP complexes after they transition from the pre-translocation state to the post-translocation state, and must dissociate from complexes prior to transition from the post-translocation state to the pre-translocation state.

We plotted the dGTP association rates (k_{on}; Fig. 2.6 C) and dissociation rates (k_{off}; Fig. 2.6 D) for complexes captured as a function of applied voltage. Neither k_{on} nor k_{off} for the mutant enzymes displays a systematic trend with the applied voltage (Fig. 2.6 C and 2.6 D); as we have previously shown for

the wild

type enzyme [66], the dNTP binding rates are independent of applied force

(Fig. 2.6 C and 2.6 D). Therefore, to compare dNTP binding rates among the

different enzymes, for each enzyme we treat all data points for k_{on} or k_{off} as

independent samples and calculate the mean and standard error for each of

the two rates (Table 2.2).

Table 2.2: Complementary dNTP binding rates for wild type phi29 DNAP and
mutants.

Enzyme [a]	$k_{on}(s^{-1}\mu M^{-1})$ [b]	$k_{off}(s^{-1})$ [c]	K_d (μM) [d]
Wild type	18.88 ± 0.5	30.41 ± 0.59	1.621 ± 0.049
Y226F	13.94 ± 0.36	8.97 ± 0.21	0.654 ± 0.041
Y390F	8.33 ± 0.24	343.01 ± 5.80	41.61 ± 1.74
D12A/D66A	19.87 ± 0.27	18.02 ± 0.30	0.908 ± 0.023
D12A/D66A (in **Mn^{2+}**)	11.45 ± 0.86	0.44 ± 0.02	0.040 ± 0.005
Y226F/D12A/D66A	17.84 ± 0.29	8.07 ± 0.18	0.453 ± 0.012
Y390F/D12A/D66A	10.42 ± 0.18	220.09 ± 2.63	21.23 ± 0.66
Y390F/D12A/D66A (in **Mn^{2+}**)	13.64 ± 0.25	11.35 ± 0.25	0.835 ± 0.023

[a] All enzymes are examined in Mg^{2+} unless otherwise specified.
[b] The dGTP association rate constant.
[c] The dGTP dissociation rate.
[d] K_d values calculated from the ratio $k_{off}(s^{-1})/k_{on}(s^{-1}\mu M^{-1})$.
All values are reported with the standard error.

The increase in dNTP binding affinity caused by the Y226F mutation (Fig. 2.5)

is due to a significant decrease in k_{off} (8.97 ± 0.21 s^{-1} for the Y226F mutant vs.

30.41 ± 0.59 s^{-1} for the wild type enzyme; Table 2.2, Fig. 2.6 D). The Y226F

mutation also causes a decrease in the dNTP association rate ($k_{on} = 13.94 \pm$

0.36 μM^{-1}s^{-1} for Y226F vs. 18.88 ± 0.5 μM^{-1}s^{-1} for the wild type enzyme; Table 2.2, Fig. 2.6 C), but the effect of this decrease in k_{on} on the binding equilibrium is more than offset by the decrease in the dissociation rate. For the Y390F mutant, k_{on} is slower than it is for the wild type phi29 DNAP (8.33 ± 0.24 μM^{-1}s^{-1} for Y390F vs. 18.88 ± 0.5 μM^{-1}s^{-1} for the wild type enzyme; Table 2.2, Fig. 2.6 C), and k_{off} is > 10 times faster than it is for wild type phi29 DNAP (343.01 ± 5.80 s^{-1} for Y390F vs. 30.41 ± 0.59 s^{-1} for the wild type; Table 2.2, Fig. 2.6 D). Therefore, changes in both dNTP binding rates contribute to the large decrease in dNTP binding affinity caused by the Y390F mutation.

The D12A/D66A mutations in the exonuclease active site have negligible effect on the dNTP association rate (Table 2.2, Fig. 2.6 C); they increase the dNTP binding affinity by decreasing k_{off} (18.02 ± 0.30 s^{-1} for the D12A/D66A enzyme vs. 30.41 ± 0.59 s^{-1} for the wild type enzyme; Table 2.2, Fig. 2.6 D). The increase in dNTP binding affinity of the Y226F/D12A/D66A mutant relative to the Y226F mutant (Fig. 2.5 B) is attributable to the combination of a modest increase in the association rate (17.84 ± 0.29 μM^{-1}s^{-1} for the Y226F/ D12A/D66A mutant vs. 13.94 ± 0.36 μM^{-1}s^{-1} for the Y226F mutant; Table 2.2, Fig. 2.6 C) and a modest decrease in the dissociation rate (8.07 ± 0.18 s^{-1} for Y226F/D12A/D66A vs. 8.97 ± 0.21 s^{-1} for Y226F; Table 2.2, Fig. 2.6 D). The increase in dNTP binding affinity for the Y390F/D12A/D66A mutant relative to the Y390F mutant (Fig. 2.5 B) arises due to an increase in the association rate (10.42 ± 0.18 μM^{-1}s^{-1} for Y390F/D12A/D66A vs. 8.33 ± 0.24 μM^{-1}s^{-1} for

Y390F; Table 2.2, Fig. 2.6 C) and a more substantial decrease in the dissociation rate (220.09 ± 2.63 s^{-1} for Y390F/D12A/D66A vs. 343.01 ± 5.80 s^{-1} for Y390F; Table 2.2, Fig. 2.6 D). Nonetheless, the dNTP dissociation rate for the Y390F/D12A/D66A mutant is still dramatically higher than it is for either the wild type enzyme (30.41 ± 0.59 s^{-1}) or the D12A/D66A mutant (18.02 ± 0.30 s^{-1}); in both backgrounds, the introduction of the Y390F mutation increases k_{off} by > 10-fold.

Rescue by Mn^{2+} of the impairment in dNTP binding kinetics caused by the Y390F mutation. In the presence of Mg^{2+}, the Y390F mutant is severely compromised in processive DNA synthesis relative to the wild type phi29 DNAP, even at high concentrations of dNTPs [14]. However, in the presence of Mn^{2+}, the level of processive synthesis catalyzed by the Y390F mutant as a function of dNTP concentration is very similar to the level catalyzed by the wild type enzyme [14]. We considered that the decrease in the dNTP association rate and the large increase in the dNTP dissociation rate caused by the Y390F mutation (Fig. 2.6 D; Table 2.1) might contribute to the synthesis impairment observed in the presence of Mg^{2+}, and whether the binding rates might be rescued in the presence of Mn^{2+}. In particular, the large dissociation rate of Mg^{2+}-dNTPs in the presence of the Y390F mutation could significantly decrease the probability of progressing to the chemical step upon dNTP binding. If so, the rescue of processive synthesis in the presence of Mn^{2+}-dNTPs might be explained, at least in part, if the
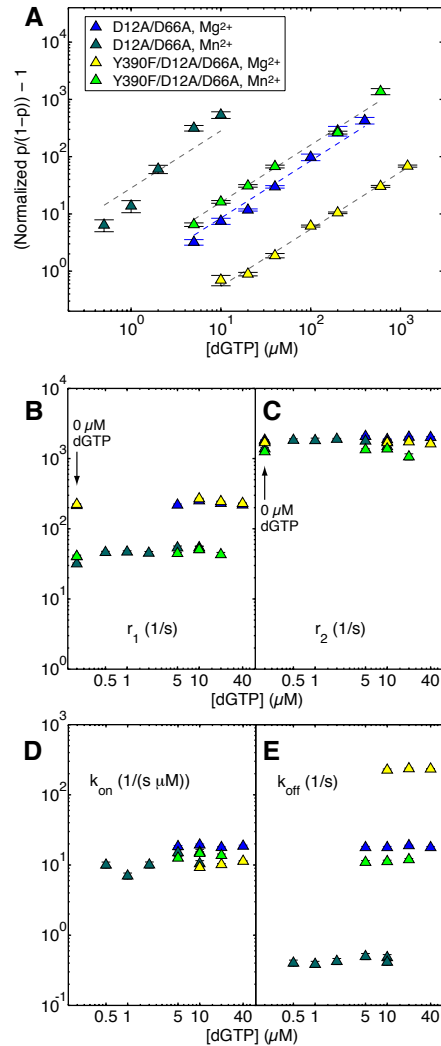
Figure 2.7: Effects of Mn^{2+} on dNTP binding to mutant phi29 DNAP-DNA complexes. (**A**) The (normalized p/(1-p)) -1 is plotted on a log scale as a function of dGTP concentration, for complexes formed between DNA1 and the D12A/D66A mutant in the presence of Mg^{2+} (blue triangles) or Mn^{2+} (dark green triangles), and between DNA1 and the Y390F/D12A/D66A mutant, in the presence of Mg^{2+} (yellow triangles) or Mn^{2+} (light green triangles). Complexes were captured at 180 mV. Plots of (**B**) r_1 vs. [dGTP] and (**C**) r_2 vs. [dGTP] for complexes formed between the D12A/D66A mutant or the Y390F/D12A/D66A mutant with DNA1, captured at 180 mV in the presence of Mg^{2+} or Mn^{2+}. Because there is no zero value on the log scale plot, in (B and C), the values of r_1 and r_2 for binary complexes of the two mutants, captured in the presence of Mg^{2+} or Mn^{2+} are placed on the plot at the position for 0.2 μM dGTP, and are indicated by an arrow and label (0 μM dGTP). Plots of (**D**) k_{on} vs. [dGTP] and (**E**) k_{off} vs. [dGTP] for complexes formed between the D12A/D66A mutant or the Y390F/D12A/D66A mutant with DNA1, captured at 180 mV in the presence of Mg^{2+} or Mn^{2+}. Plot symbols in (B-D) are the same as in panel A. Errors bars indicate the standard error.

dissociation rate of Mn^{2+}-dNTPs was slower than the rate for Mg^{2+}-dNTPs. A decrease in the dissociation rate that brings the dwell time in the dNTP-bound state ($1/k_{off}$) into a regime that is sufficient for chemistry could improve the function of the mutant enzyme in synthesis. (Note that it is not necessary that the dwell time in the dNTP-bound state be increased to the level of the wild type for this to obtain, only that it be increased sufficiently to increase the probability of chemistry).

We compared the effects of the Y390F mutation on the translocation step and on dNTP binding in the presence of Mn^{2+} or Mg^{2+}. To avoid bulk phase exonucleolytic degradation, we performed these experiments using the Y390F/D12A/D66A and D12A/D66A enzymes. In binary complexes, the identity of the divalent metal cation affects the equilibrium across the translocation step; for complexes formed in Mn^{2+}, the probability of the post-translocation state is smaller than it is for complexes formed in Mg^{2+}. When binary complexes formed between D12A/D66A and DNA1 are captured at 180 mV, p/(1-p) in Mg^{2+} = 0.117 ± 0.002; p/(1-p) in Mn^{2+} = 0.023 ± 0.0004. For complexes formed with the Y390F/D12A/D66A enzyme and captured at 180 mV, p/(1-p) in Mg^{2+} = 0.132 ± 0.001; p/(1-p) in Mn^{2+} = 0.032 ± 0.001. Plots of normalized p/(1-p) for dGTP titration experiments show that complexes formed with both the D12A/D66A and Y390F/D12A/D66A mutants have a greater affinity for dNTP when they are captured in Mn^{2+} than when they are captured in Mg^{2+} (Fig. 2.7 A). Nonetheless, when the two enzymes are

58

compared in Mn^{2+} the Y390F/D12A/D66A mutant retains a significantly diminished affinity for dNTP relative to the D12A/D66A mutant (Fig. 2.7 A).

To determine the kinetic mechanisms by which Mn^{2+} alters the translocation and dNTP binding, we compared the translocation fluctuation rates and the dNTP binding rates for complexes formed between the D12A/D66A and Y390F/D12A/D66A enzymes in Mg^{2+} or Mn^{2+}. The shift in the translocation equilibrium toward the pre-translocation state caused by Mn^{2+} in binary complexes is primarily due to a significant decrease in the forward translocation rate relative to complexes formed in Mg^{2+} (Fig. 2.7 B and 2.7 C; Table 2.1). At 180 mV, for the D12A/D66A mutant, r_1 in Mg^{2+} = 214.36 ± 8.56; r_1 in Mn^{2+} = 40.05 ± 0.88. For the Y390F/D12A/D66A mutant, r_1 in Mg^{2+} = 223.08 ± 4.6; in Mn^{2+}, r_1 = 40.43 ± 1.74. It is possible that this decrease in the forward translocation rate caused by Mn^{2+} may contribute to the generally decreased level of processive synthesis supported by Mn^{2+} relative to Mg^{2+} for both wild type and mutant enzymes [14]. We will examine the effects of divalent metals on the dynamics of the phi29 DNAP translocation step in detail in a separate study.

As we have shown for complexes captured in Mg^{2+} (Fig. 2.6 A and 2.6 B; [66] when complexes are captured in Mn^{2+}, neither r_1 nor r_2 are affected by [dNTP] (Fig. 2.7 B and 2.7 C). The effect of Mn^{2+} on the dNTP association rate (relative to Mg^{2+})

differed between the two enzymes (Table 2.2, Fig. 2.7 D). For the D12A/D66A

mutant k_{on} was slower in Mn^{2+} (11.45 ± 0.86 $\mu M^{-1}s^{-1}$) than in Mg^{2+} (19.87 ± 0.27 $\mu M^{-1}s^{-1}$); for the Y390F/D12A/D66A mutant, k_{on} increased very slightly in Mn^{2+} (k_{on} in Mn^{2+} = 13.64 ± 0.25 $\mu M^{-1}s^{-1}$; k_{on} in Mg^{2+} = 10.42 ± 0.18 $\mu M^{-1}s^{-1}$). For both enzymes, Mn^{2+} had a dramatic effect on the dNTP dissociation rates (Table 2.2, Fig. 2.7 E). For the D12A/D66A mutant, k_{off} in Mg^{2+} = 18.02 ± 0.30 s^{-1}; k_{off} in Mn^{2+} = 0.44 ± 0.02 s^{-1}. For the Y390F/D12A/D66A mutant, k_{off} in Mg^{2+} = 220.09 ± 2.63 s^{-1}; k_{off} in Mn^{2+} = 11.35 ± 0.25 s^{-1}. Thus, Mn^{2+} decreases the dNTP dissociation rate for the Y390F/D12A/D66A mutant to a value smaller than 30.41 ± 0.59 s^{-1}, the dNTP dissociation rate for the wild type enzyme in Mg^{2+} (Table 2.2, Fig. 2.6 D). These data are consistent with the proposal that the rescue of processive DNA synthesis by the Y390F mutant in the presence of Mn^{2+} could result from a decrease in the dNTP dissociation rate for the mutant, relative to this rate for complexes formed in the presence of Mg^{2+}.

The findings that both r_1 and k_{off} are significantly slower for complexes captured in Mn^{2+} than they are for complexes captured in Mg^{2+} suggests the possibility that Mn^{2+} may exert its effects on r_1 and k_{off} via a common mechanism: it may diminish the rate of fingers opening from the closed complex. The structural model proposed for the translocation step [6] predicts that decreasing the rate of fingers opening in the pre-translocation state would lead to a decrease in the rate of the forward translocation (r_1). In the dNTP-bound post-translocation state, decreasing the rate of fingers opening

could yield a decrease in the dissociation rate of dNTP (k_{off}).

Pyrophosphate binding to the pre-translocation state in mutant enzyme complexes. The strong (and opposing) effects of the Y226F and Y390F mutations on the kinetics of dNTP binding to post-translocation state complexes prompted us to examine whether these mutations affected the binding of pyrophosphate to phi29 DNAP-DNA complexes. As a product of phosphodiester bond formation, pyrophosphate is bound to pre-translocation state complexes immediately following the chemical step. We compared the effects of pyrophosphate on complexes formed between each of the six phi29 DNAP enzymes and DNA1 in pyrophosphate titration experiments (Fig. 2.8 A). To highlight the effect of pyrophosphate on the pre-translocation state probability, we plotted the normalized (1-p)/p as a function of pyrophosphate concentration, where p is the probability of the post-translocation state and consequently (1-p) is the probability of the pre-translocation state. The normalized (1-p)/p is determined solely by the binding affinities of pyrophosphate. This quantity is independent of the transitions between the two translocation states in the absence of pyrophosphate (and thus independent of any differences in the translocation rates among the six enzymes; Fig. 2.4); the effect of these transitions is eliminated when (1-p)/p is normalized by its value in the absence of pyrophosphate.

When pyrophosphate was titrated into the nanopore chamber in the presence of complexes formed between DNA1 and the wild type, Y226F, Y390F, D12/

D66A, Y226F/D12/D66A, or Y390F/D12/D66A enzymes, the equilibrium across the translocation step was shifted toward the pre-translocation state for the Y226F and Y226F/D12/D66A mutants, while the translocation equilibrium for the wild type, Y390F, D12/D66A, and Y390F/D12/D66A enzymes was unaffected (Fig. 2.8 A). Thus, the introduction of the Y226F mutation specifically causes a concentration-dependent shift in the equilibrium across the phi29 DNAP translocation step, toward the pre-translocation state. The enzymes bearing the Y226F mutation displayed a ~ 2.5-fold increase in pre-translocation state probability at the highest concentration of pyrophosphate that could be tested without the risk of precipitation (Fig. 2.8 A).

For each of the six enzymes, complexes captured in the presence of pyrophosphate resided atop the nanopore and fluctuated between the upper and lower amplitude states for tens of seconds, during which we rarely observed pyrophosphorolytic (or exonucleolytic) cleavage (which can be readily discerned by a change in ionic current amplitude [24, 65]. This is consistent with our prior finding that pyrophosphorolysis is extremely slow for complexes formed with DNA substrates bearing 3'-H termini [24].

To determine the dynamic mechanism by which pyrophosphate shifts the translocational equilibrium in complexes with enzymes bearing the Y226F mutation, we examined the effects of pyrophosphate on the average dwell time in the pre-translocation and post-translocation states. We used a half-
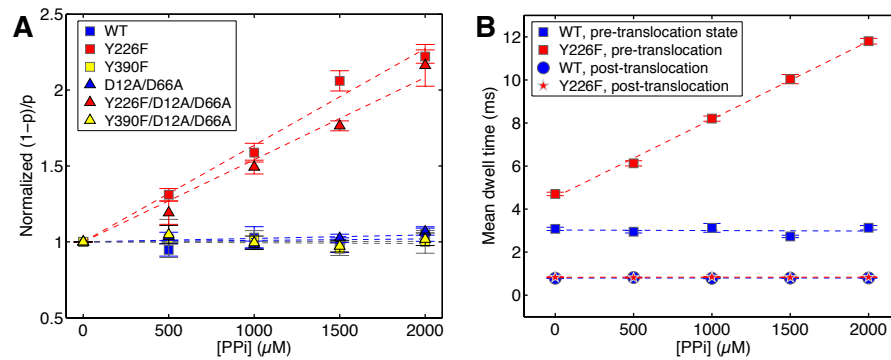
Figure 2.8: Pyrophosphate binding to mutant phi29 DNAP-DNA complexes. (**A**) The normalized (1-p)/p (where (1-p) is the probability of pre-translocation state occupancy, and the normalized (1-p)/p is the value of (1-p)/p in the presence of a given concentration of pyrophosphate, divided by the value for (1-p)/p for the same phi29 DNAP-DNA complex at 0 mM pyrophosphate) is plotted as a function of pyrophosphate concentration for complexes formed between wild type, Y226F, Y390F, D12A/D66A, Y226F/D12A/D66A or Y390F/D12A/D66A phi29 DNAP and DNA1. Plot symbols for each of the enzymes are given in the legend for Fig. 4. In (**B**), the mean dwell time in the pre-translocation and post-translocation states for complexes formed between wild type phi29 DNAP (pre-translocation, blue squares; post-translocation, blue circles) and the Y226F mutant (pre-translocation, red squares; post-translocation, red stars) are plotted as a function of pyrophosphate concentration. Complexes were captured at 160 mV. Errors bars indicate the standard error.

amplitude threshold method [23] to extract dwell time samples from ionic current traces, for complexes formed between each of the six enzymes and DNA1. Pyrophosphate causes a concentration-dependent, linear increase in the average pre-translocation state dwell time for complexes formed with the Y226F and Y226F/D12A/D66A mutants, while average dwell time in the post-translocation state for these mutants is unaffected (Fig. 2.8 B; Table 2.3). For complexes formed with the wild type, Y390F, D12A/D66A, or Y390F/D12A/D66A enzymes, neither the average dwell time in the pre-translocation state nor the average dwell time in the post-translocation state was affected by the

Table 2.3: Average dwell times in the pre-translocation and post-translocation states.

phi29 DNAP[a]	[PPi] (mM)	Pre-translocation state mean dwell time (ms)	Post-translocation state mean dwell time (ms)
Wild type	0	3.08 ± 0.07	0.79 ± 0.01
	0.5	2.95 ± 0.05	0.82 ± 0.01
	1.0	3.13 ± 0.16	0.78 ± 0.01
	1.5	2.73 ± 0.04	0.79 ± 0.01
	2.0	3.13 ± 0.09	0.8 ± 0.01
Y226F	0	4.7 ± 0.07	0.84 ± 0.02
	0.5	6.13 ± 0.10	0.82 ± 0.01
	1.0	8.2 ± 0.11	0.85 ± 0.01
	1.5	10.04 ± 0.16	0.83 ± 0.01
	2.0	11.8 ± 0.12	0.84 ± 0.01
Y390F	0	3.58 ± 0.04	0.97 ± 0.01
	2.0	3.3 ± 0.04	0.94 ± 0.01
D12A/D66A	0	3.89 ± 0.07	0.87 ± 0.01
	2.0	3.65 ± 0.06	0.83 ± 0.01
Y226F/ D12A/ D66A	0	5.92 ± 0.09	1.13 ± 0.02
	2.0	14.96 ± 0.36	1.2 ± 0.03
Y390F/ D12A/ D66A	0	3.58 ± 0.02	0.95 ± 0.01
	2	3.52 ± 0.04	0.9 ± 0.01

[a] Complexes were formed between the indicated phi29 DNAP enzymes and DNA1, and captured atop the nanopore at 160 mV applied potential.

Error bars indicate the standard errors.

presence of pyrophosphate (Fig. 2.8 B; Table 2.3). Thus, the Y226F mutation specifically increases the affinity of the pre-translocation state for pyrophosphate. For the wild type, D12A/D66A, Y390F, and Y390F/D12A/ D66A enzymes, the affinity for pyrophosphate is below the limit of detection of our assay.

The Y226F mutant has a slower forward translocation rate than the wild type enzyme (Fig. 2.4; Table 2.1), and its forward translocation is further diminished in the presence of pyrophosphate (Fig. 2.8 B). The immediate product of the chemical step in DNA synthesis is the fingers-closed, pre-translocation state complex with pyrophosphate bound, and the probability of fingers opening and the consequent pre-translocation to post-translocation transition may be lower when pyrophosphate is bound. Nonetheless, the Y226F mutant is not impaired in the rate or processivity of DNA synthesis when assayed in the bulk phase (Fig. 2.3). This suggests that the effects of the Y226F mutation on the forward translocation rate, revealed under opposing force, and on the affinity for pyrophosphate in the pre-translocation state, do not reduce the forward translocation rate enough to impede the rate of DNA synthesis in the absence of an opposing force.

2.5 Conclusion

The Y390F and Y226F mutations in phi29 DNAP both disrupt the H-bond between the hydroxyl groups of these two conserved residues, an interaction

that was proposed to stabilize the orientation of Y390 in the fingers-closed, post-translocation ternary complex, and in the fingers-closed, pre-translocation state complex [6]. Nonetheless, biochemical studies revealed that the Y390F and Y226F mutations yield very different properties in phi29 DNAP [11, 95, 14, 86]. In this study we have shown that the Y390F and Y226F mutations differ in their effects on the dynamics of the translocation step, and on the dynamics of substrate binding in the both the pre-translocation and post-translocation states. Somewhat surprisingly, we found that neither the Y226F nor the Y390F mutation caused an increase in the forward translocation rate; this rate is decreased by the Y226F mutation and largely unaffected by the Y390F mutation. This indicates that the disruption of the H-bond between the two residues does not destabilize the fingers-closed pre-translocation state, implying that the H-bond makes a relatively small contribution to the free energy of this state.

While the most dramatic consequences of introducing the Y390F or Y226F mutations are their opposing effects on dNTP binding rates and affinity, it is difficult to clearly attribute these effects to the destabilization of the fingers-closed, dNTP-bound, post-translocation state complex. The effects on the dNTP binding rates elicited by the Y390F mutation, particularly the significantly increased dNTP dissociation rate, could instead be due to a distortion of dNTP binding within the closed complex, rather than destabilization of the closed complex itself, since the closed pre-translocation

state does not appear to be destabilized by this mutation. Distortion of dNTP binding may arise as a consequence of the perturbation of the water-mediated interactions of Y390 with the template strand in the DNA duplex.

The increased rate of dNTP dissociation observed with the Y390F mutation is not observed when the H-bond is disrupted by the Y226F mutation; by contrast this mutation leads to a dramatic decrease in the dNTP dissociation rate. The Y226F mutation also yields a higher affinity for pyrophosphate in the pre-translocation state, which may reflect a slower dissociation rate for this ligand. This suggests that the higher affinity for both dNTP and for pyrophosphate caused by the Y226F mutation may arise from common mechanism. As with the Y390F mutation, it is possible that the effects of the Y226F mutation on ligand binding arise from an indirect effect on the water-mediated interactions of Y390 with the template strand of the DNA duplex. In this scenario, the loss of the H-bond between Y390 and Y226 results in a change in the distance or angle of the bonds formed between Y390 and the template strand, leading to tighter binding of both dNTP in the post-translocation state, and pyrophosphate in the pre-translocation state. Indeed, it cannot be ruled out that such a structural perturbation might also contribute to the effects of the Y226F mutation on the translocation rates.

It is difficult to unequivocally assign the effects of introducing the Y390F or Y226F mutations on the translocation kinetics, or on substrate binding kinetics in the pre-translocation and post-translocation states, directly to

disruption of the H-bonding potential between them in fingers-closed phi29 DNAP-DNA complexes. These findings do not refute the proposed structural mechanism which links the spatial displacement of translocation to the fingers opening transition [6], but they suggest that the H-bond that forms between Y390 and Y226 in the fingers-closed state is not directly essential to the mechanism.

Chapter 3

Kinetic Mechanisms Governing Stable Ribonucleotide Incorporation in Individual DNA Polymerase Complexes

3.1 Abstract

Ribonucleoside triphosphates (rNTPs) are frequently incorporated during DNA synthesis by replicative DNA polymerases (DNAPs), and once incorporated are not efficiently edited by the DNAP exonucleolytic function. We examined the kinetic mechanisms that govern selection of complementary deoxyribonucleoside triphosphates (dNTPs) over complementary rNTPs, and that govern the probability of a complementary ribonucleotide at the primer terminus escaping exonucleolytic editing and becoming stably incorporated. We studied the quantitative responses of individual phi29 DNAP complexes to ribonucleotides using a kinetic framework, based on our prior work, in which transfer of the primer strand from the polymerase to exonuclease site occurs prior to translocation, and translocation precedes dNTP binding. We determined transition rates among the pre-translocation and post-translocation states, the polymerase and exonuclease sites, and for dNTP or

rNTP binding, with single-nucleotide spatial precision and sub-millisecond temporal resolution from ionic current time traces recorded when individual DNAP complexes are held atop a nanopore in an electric field. The predominant response to the presence of a ribonucleotide in phi29 DNAP complexes before and after covalent incorporation is significant destabilization, relative to the presence of a deoxyribonucleotide. This destabilization is manifested in the post-translocation state prior to incorporation as a substantially higher rNTP dissociation rate, and manifested in the pre-translocation state after incorporation as rate increases for both primer strand transfer to the exonuclease site and the forward translocation, with the probability of editing not directly increased. In the post-translocation state the primer terminal 2′-OH group also destabilizes dNTP binding.

3.2 Introduction

While RNA has the same coding potential as DNA, the genomes of contemporary cellular organisms consist of DNA, which is more chemically stable. Nonetheless, a mounting body of evidence has revealed that ribonucleoside triphosphates (rNTPs), which are more abundant in cells than deoxyribonucleoside triphosphates (dNTPs) [94, 76, 39], are frequently incorporated during DNA synthesis catalyzed by replicative DNA polymerases (DNAPs) both *in vitro* and *in vivo* [76, 104, 44, 83, 72, 107]. For example, yeast Pol-ε, the B-family DNAP that catalyzes leading strand nuclear replication in eukaryotes, incorporates one ribonucleoside monophosphate (rNMP) residue

per 1250 deoxyribonucleoside monophosphates (dNMPs) in the presence of physiological concentrations of dNTPs and rNTPs *in vitro* [76]. Under the same conditions, yeast Pol-δ, the B-family DNAP responsible for lagging strand nuclear replication in eukaryotes, incorporates one rNMP per 5000 dNMPs [76].

When an rNMP residue is incorporated during replication, the 2′-OH group renders the linkage more labile than DNA linkages, and can lead to replicative stress and genome instability [77, 21, 90, 17]. The rNMP-containing product can be repaired after synthesis, primarily via RNase H2-mediated pathways [90, 33, 84, 110, 70, 96]. If left unrepaired, DNAPs may encounter rNMPs in the template strand during subsequent rounds of replication, and these encounters can lead to termination of DNA synthesis [76, 104]. In addition to these potentially deleterious effects, incorporated rNMPs have also been implicated as strand-marking signals in the mismatch repair pathway [72, 110, 28, 43].

DNAPs favor the initial selection of complementary dNTPs over complementary rNTPs, due to polymerase active site structural features that discriminate sterically against rNTPs [51, 108, 15, 19, 16], and due to conformational checkpoints that selectively promote optimal alignment of complementary dNTP and misalignment of complementary rNTP [41, 101, 8]. Levels of selectivity of dNTP over rNTP range from ~10^3 to >10^6, depending upon the DNAP and the identity of the base pair examined [101]. Nonetheless, rNTPs can escape these initial discrimination mechanisms and become

incorporated at significant levels during replication. For example, the incorporation of one rNMP per 5000 dNMPs by yeast Pol-δ occurs despite rNTP discrimination levels for this enzyme ranging from 10^4 to $>10^6$ (depending upon the base pair) [76]. It is therefore essential to understand the mechanisms that govern the probability of replicative DNAPs continuing processive synthesis once an rNMP is incorporated.

In addition to an active site for polymerization, many DNAPs have a separate active site for 3′-5′ exonucleolytic editing of newly incorporated replication errors. The exonuclease and polymerase sites of DNAPs are located in separate protein domains (or in some cases, in separate subunits), typically separated by ~ 30 Angstroms [6, 32, 37, 58, 98], and primer strand transfer from the polymerase to exonuclease site requires that ~ 3 base pairs of the primer-template duplex be melted [87]. In contrast to noncomplementary dNMPs, when complementary rNMPs are incorporated they are not efficiently edited by replicative DNAPs, including yeast Pol-ε, human and yeast Pol-δ, and the B-family replicative DNAP from bacteriophage phi29 [15, 106, 22].

During replication DNAPs translocate along their DNA substrates in single nucleotide increments. We have developed a single-molecule approach using a nanoscale pore to quantify the rates of translocation, primer strand transfer between the polymerase and exonuclease sites, and dNTP binding, with single-nucleotide spatial precision and sub-millisecond temporal resolution [24, 66, 67, 68] (see also Supporting Information) When individual DNAP complexes

are captured atop an α-hemolysin (α-HL) nanopore in an electric field, forward and reverse fluctuations across the translocation step are observed as discrete transitions between two amplitudes in time traces of ionic current. Rates of the translocation fluctuations, rates of primer strand transfer in both directions between the polymerase and exonuclease sites, and rates of dNTP binding can be quantified mathematically from dwell time samples extracted from the ionic current time traces [66, 67, 25]. Using the phi29 DNAP as a model system, we have shown that transfer of the primer strand from the polymerase to exonuclease site occurs prior to translocation [68, 26]; the pre-translocation state is therefore the branchpoint between the DNA synthesis and editing pathways (Fig. 3.1 A). The forward translocation step precedes dNTP binding; the translocation is rectified but not driven by dNTP binding, and dNTP has no influence on the rates across the translocation step (Fig. 3.1 A) [67]. If complementary rNTP escapes initial discrimination and is covalently incorporated, the rNMP residue at the primer terminus can affect the probability of primer strand transfer between the polymerase and exonuclease sites, of the translocation step in the polymerase site, and of subsequent dNTP incorporation after the translocation. Together, the transition rates for these processes determine the net probability of stably incorporating the rNMP.

In the B family of DNAPs, the core structure, catalytic mechanisms, and functional properties that contribute to replication fidelity are highly conserved
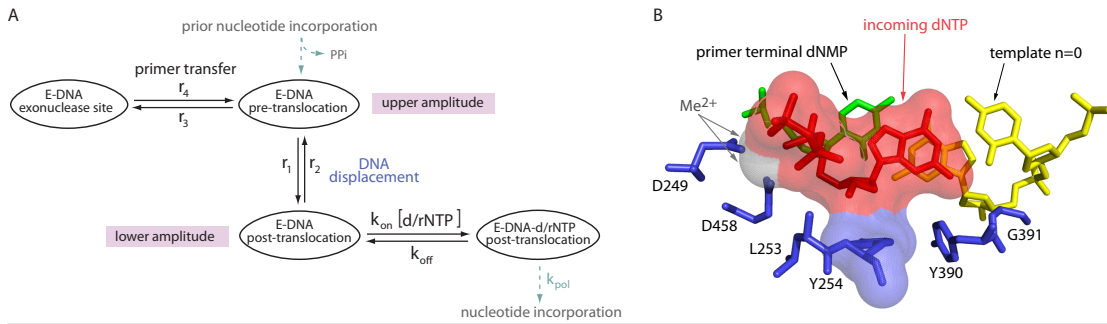
Figure 3.1: Kinetic and structural determinants of stable dNTP or rNTP incorporation. (**A**) Model for the kinetic relationships among the steps of translocation, primer strand transfer between the polymerase and exonuclease active sites, and nucleotide triphosphate binding. The kinetic model is fully described by six transition rates: the rates of translocation (r_1, r_2), the rates of primer stand transfer between the pre-translocation state polymerase site and the exonuclease site (r_3, r_4), and the rates of dNTP binding to post-translocation state complexes (k_{on}[dNTP] and k_{off}). The mathematical framework based on the model allows us to determine these rates from experimental measurements [66, 67, 68]. (**B**) View of the polymerase active site in the phi29 DNAP-DNA-dNTP, post-translocation state ternary complex from the crystal structure model in PDBID: 2PYJ. The structure is from ref. (6), and was determined using the D12A/D66A mutant of phi29 DNAP. Protein residues are blue, the template strand (the n=0 and n= -1 template residues) are yellow, the primer terminal residue is green, and the incoming dNTP is shown in red. The two active site Me^{2+} ions are rendered as grey spheres. The solvent accessible surfaces of the incoming dNTP and residue Y254, in red and blue, respectively, are shown to highlight the stacking of the deoxyribose sugar of the incoming dNTP on the phenyl ring of residue Y254. Surfaces were rendered in PyMol using a solvent radius of 1.4 Å.

[6, 93, 100, 48, 53, 54]. phi29 DNAP serves as an excellent model system for leading strand DNA synthesis catalyzed in more complex replisomes; it catalyzes highly processive DNA synthesis [12, 13, 85], without the need for accessory proteins such as sliding clamps or helicases, because it remains tightly associated with its DNA substrate and promotes downstream strand displacement during replication [6, 73]. The phi29 DNAP exonuclease active site is located ~ 30 Å from the polymerase active site [6, 58], and primer strand transfer between the polymerase and exonuclease sites is an intramolecular

process [29]. In accord with the inefficient editing of rNMPs by replicative DNAPs Pol-ε and Pol-δ, the probability of exonucleolytic editing of a newly incorporated rNMP by phi29 DNAP is similar to the basal level observed for complementary dNMP [15]. For both phi29 DNAP and the B-family DNAP from bacteriophage T4, exonucleolytic cleavage of an rNMP residue at the 3′ terminus of a single-stranded substrate is not significantly impaired [15, 71]. Thus, it is possible that the inefficient editing observed after complementary rNMP incorporation could occur because the rNMP:dNMP pair is not recognized as incorrect; in this scenario the probability of primer strand transfer from the polymerase to exonuclease site for the complementary rNMP:dNMP pair would not be greater than it is for a complementary dNMP:dNMP pair. But this has not been directly tested, and the contribution of the translocation rates and the dNTP binding rates after translocation to the probability of editing has not been examined.

phi29 DNAP discriminates against complementary rNTP incorporation by > 2 million-fold, in steady-state kinetic assays conducted in Mn^{2+} [15]. In the phi29 DNAP ternary complex crystal structure the dNTP is snugly bound, with its deoxyribose moiety stacked upon the phenyl ring of the conserved Y254, leaving no space to accommodate a 2′-OH group on the sugar (Fig. 3.1 B) [6], leading to the hypothesis that discrimination is due at least in part to steric exclusion of the ribose 2′-OH group. Consistent with this hypothesis, while the Y254F mutation had negligible effects on discrimination between dNTP and

rNTP incorporation, introduction of the Y254V mutation decreased the discrimination against rNTP incorporation relative to dNTP incorporation by ~ 3 orders of magnitude [15].

Because primer strand transfer from the polymerase site to the exonuclease site precedes translocation [68], there is a kinetic competition between primer transfer and the forward translocation. This kinetic competition could be perturbed by a newly incorporated rNMP. Based upon a structural model for the pre-translocation state in which the terminal base pair of the duplex occupies the same configuration that the nascent base pair between the templating base and the incoming dNTP occupies in the post-translocation state ternary complex [6], we hypothesized that the presence of a $2'$-OH group at the primer terminus would destabilize the pre-translocation state, due to a structural mechanism of steric exclusion directly analogous to the mechanism that yields discrimination against rNTP binding in the post-translocation state. Specifically, in the structural model for the pre-translocation state, the sugar moiety of the primer terminal residue is predicted to stack on the phenyl ring of Y254, leaving no space to accommodate a primer-terminal $2'$-OH group. Destabilization of the pre-translocation state following rNTP incorporation could lead to an increase in the rates of both pathways out the pre-translocation state: i) the transition from the pre-translocation state, polymerase site to the exonuclease site (r_3; Fig. 2.1 A), and ii) the transition from the pre-translocation state to the post-translocation state (r_1; Fig. 2.1 A).

76

The rates of both of these transitions can directly affect the probability that an incorporated rNMP will be exonucleolytically edited. Furthermore, following the forward translocation, the presence of a 2′-OH group on the primer strand could diminish the probability of binding and incorporating the next dNTP. Because dNTP binds to phi29 DNAP complexes only after the forward translocation and has no influence on the forward or reverse translocation rates (r_1 and r_2, respectively; Fig. 2.1 A) [67], a decrease in dNTP binding affinity or in the rate of progression to chemistry (k_{pol}) caused by a primer terminal rNMP could in turn indirectly affect the probability of transfer to the exonuclease site by increasing the probability of fluctuation back to the pre-translocation state.

In this study we first examined the kinetic binding mechanisms that govern selection of complementary dNTPs over complementary rNTPs by phi29 DNAP, and by which the Y254V mutation attenuates this discrimination. We then determined the consequences of covalent incorporation of an rNMP residue at the primer terminus on the rates of translocation, the rates of primer strand transfer between the polymerase and exonuclease sites, and the dNTP binding rates. We tested the hypothesis that steric occlusion of a primer terminal 2′-OH group destabilizes the pre-translocation state, which is the branchpoint between the synthesis and editing pathways. The transition rates determined in the study provide insight into the mechanisms that govern

the net probability of continuing processive synthesis beyond an rNMP residue, thus stably incorporating it into the nascent strand.

3.3 Experimental Methods

Enzymes. Wild type phi29 DNAP was obtained from Enzymatics (Beverly, MA). The D12A/D66A mutant was obtained from XPol Biotech (Madrid, Spain). Construction, expression and purification of the Y254V mutant has been described [15, 69]. The D12A/D66A-6His mutant was constructed by adding an oligonucleotide coding for 6 histidines at the carboxyl end of the enzyme to the plasmid containing the D12A/D66A mutation [89]. To the resulting plasmid, the Y254V mutation was added to obtain the mutant D12A/D66A/Y254V-6His. The histidine tagged proteins were expressed and purified essentially as described [69], with an additional step of a nickel-agarose column. The D12A/D66A and D12A/D66A-6His enzymes were compared in nanopore experiments and the dynamics of complexes formed with the two enzymes were found to be indistinguishable.

DNA substrates. DNA1-H_H, DNA1-H_OH, DNA1-OH_OH, and a version of the DNA1 substrate shortened by one residue at the 3$'$ end were synthesized at Stanford Protein and Nucleic Acid Facility and purified by denaturing PAGE. DNA1-OH_H was generated from the shortened substrate by incorporating 3'-deoxycytidine-5'-triphosphate (Trilink Biotechnologies) using the D355A/E357A exonuclease-deficient mutant of the Klenow fragment (New England Biolabs) in the presence of Mn^{2+}. Reactions were monitored by

denaturing PAGE and routinely yielded > 95% extended product. The product was purified on Bio-Rad P6 gel filtration spin columns. DNA hairpins were annealed by heating at 90°C for 4 min followed by snap cooling in ice water.

Nanopore methods. Nanopore experiments were conducted as described [41, 24, 65, 67, 5, 2]. Briefly, a single α-HL nanopore is inserted in a ~ 25 μm-diameter lipid bilayer that separates two chambers (*cis* and *trans*) containing buffer solution (10 mM K-Hepes, pH 8.0, 0.3 M KCl, and 1 mM EDTA). $MgCl_2$ and DTT were added to the nanopore *cis* chamber to final concentrations of 11 mM and 1 mM, respectively. DNA and phi29 DNAP were added to the *cis* chamber to final concentrations of 1 and 0.75 μM, respectively. Ionic current was measured with an integrating patch clamp amplifier (Axopatch 200B, Molecular Devices) in voltage clamp mode. Data were sampled using an analog-to-digital converter (Digidata 1440A, Molecular Devices) at 100 kHz in whole-cell configuration and filtered at 5 kHz using a low pass Bessel filter.

Analysis of ionic current time traces. Each recorded time trace of ionic current amplitude from a captured phi29 DNAP complex is analyzed in the following series of steps:

i) Determine the centers and relative fractions of the 2 amplitude clusters: The collection of amplitude samples from each time trace is fitted to a model of 2 Gaussian modes. The fraction of the upper or lower amplitude mode gives the equilibrium probability of the complex residing at the upper or lower

amplitude, respectively. The centers of the upper and lower amplitude modes give approximately the underlying noiseless values of the 2 amplitude levels.

ii) Extract dwell time samples of each amplitude level from a time trace: We model the measured time trace as a true underlying noiseless time trace plus Gaussian noise. The true underlying noiseless time trace jumps between the 2 noiseless amplitude levels obtained in i). Each time the measured amplitude cross the middle point between the 2 noiseless amplitude levels, a transition is tentatively detected. To weed out spurious transitions caused by noise, we set a cut-off threshold $t_C = 100$ μs. A transition from one amplitude to the other is accepted only if the dwell time at the target amplitude level is above the threshold t_C. Transitions with dwell time below the threshold t_C are rejected. All accepted dwell time samples are shifted by t_C; these observed dwell time samples represent the dwell time beyond the threshold t_C.

iii) Determine the intermediate parameters describing the dwell time distribution of each amplitude level: If an amplitude level corresponds to a single kinetic state, the dwell time is exponentially distributed. If an amplitude level corresponds to 2 kinetic states (see the model diagram in Fig. 3.1 A), the dwell time distribution is a combination of 2 exponential modes (see Supporting Information). The observed dwell time samples of each amplitude level are fitted to their theoretical distribution using the maximum-likelihood estimation (MLE) to determine the intermediate parameters. The 2 fittings are then corrected to account for the side effects of cut-off threshold $t_C = 100$ μs.

iv) The intermediate parameters obtained in iii) are used to calculate kinetic transition rates (see Supporting Information). At each experimental condition, we typically have a set of n = 20 ~ 30 time traces, which yields a set of n estimated values for each parameter. The final estimation and the associated standard error for each kinetic rate are calculated based on this set of estimated values.

3.4 Results

In the nanopore experiments, a single α-HL nanopore is inserted into a lipid bilayer that separates two chambers (termed *cis* and *trans*) containing buffer solution (Fig. 3.2 A). A patch clamp amplifier applies voltage across the
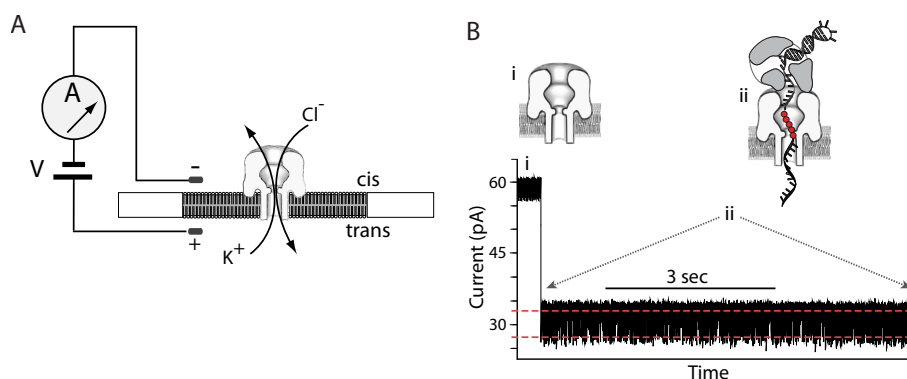
Figure 3.2: Capturing and measuring individual phi29 DNAP complexes on the α-HL nanopore. In the nanopore device (**A**), a single α-HL nanopore is inserted in a ~25 µm-diameter lipid bilayer separating two chambers (cis and trans) that contain buffer solution. A patch clamp amplifier applies voltage across the bilayer and measures ionic current, which is carried through the nanopore by K $^+$ and Cl$^-$ ions. (**B**) Representative current trace for an individual binary complex of phi29 DNAP and a DNA substrate (DNA1-H_H, in Figure 3b, i), formed in the bulk phase in the cis chamber and captured at 180 mV. Cartoons above the current trace illustrate the sequence of events, which is described in the text. In the cartoons, a reporter group of five consecutive abasic (1′, 2′-H) residues is shown as red circles. In the plot of ionic current vs. time, the two red dashed lines indicate the two amplitude levels between which the current fluctuates (detailed views of current traces are shown in Figure 3).

81

A   DNA1 sequence:

n=0

+12 +8

5´-ACTATCATTATCTACATCCATTXXXXXACTACTCCGCATGCAGGTAGCC ^{T} T
                              CGTACGTCCATCGG $_T$ ^{T}

B   DNA1-2'_3' primer termini:

i   DNA1-H_H
300 ms

ii   DNA1-OH_H

iii   DNA1-H_OH

iv   DNA1-OH_OH
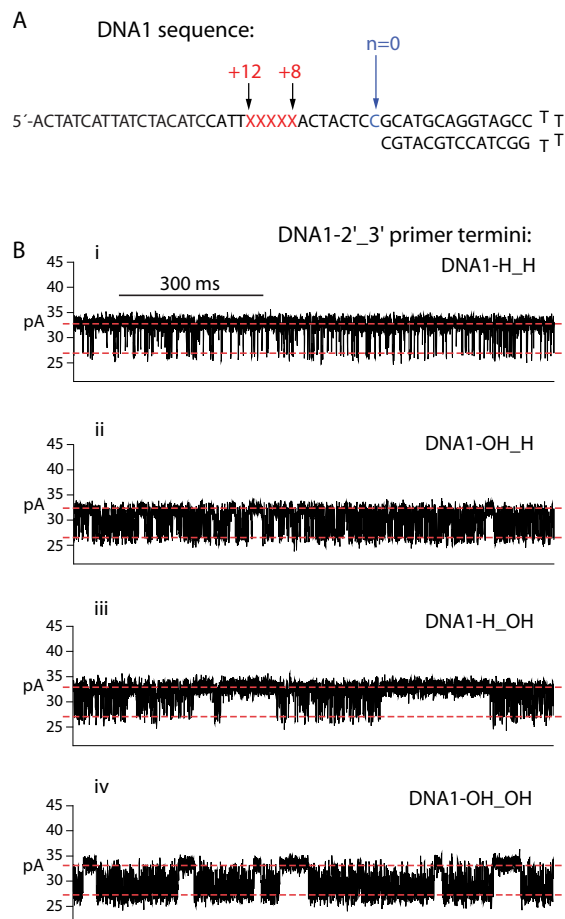
Figure 3.3: Influence of the DNA substrate primer-terminal sugar manifested in ionic current traces for captured phi29 DNAP complexes. (**A**) DNA sequence of the hairpin primer-template substrates used in the study. Each substrate features a 14-base pair duplex region and a single- stranded template region of 35 nucleotides. The template strands contain a reporter group of five consecutive abasic (1´,2´-H) residues spanning positions +8 to +12 (indicated as red Xs in the sequence). (**B**) Ionic current traces for individual binary complexes formed between the D12A/D66A mutant of phi29 DNAP with (**i**) DNA1-H_H, which bears a 2′-H, 3′-H primer terminus; (**ii**) DNA1-OH_H, which bears a 2′-OH, 3′-H primer terminus; (**iii**) DNA1-H_OH, which bears a 2′-H, 3′-OH (natural deoxyribose) primer terminus or (**iv**) DNA1-OH_OH, which bears a 2′-OH, 3′-OH (natural ribose) primer terminus. Complexes were captured at 180 mV.

bilayer, and measures the ionic current that flows through the nanopore, which is carried by K^+ and Cl^- ions in the buffer. Figure 3.2 B shows a typical ionic current trace that results when an individual binary complex formed

between phi29 DNAP and a DNA substrate (Fig. 3.3 A and B, i; DNA1-H_H) is captured atop a nanopore at 180 mV applied potential. The ionic current through the open pore (Figure 2B, i) drops rapidly when a complex is captured (Fig. 3.2 B, ii). The enzyme is too large to enter the nanopore, and therefore the phi29 DNAP-DNA complex perches atop the pore. The DNA template strand of the captured complex is suspended through the nanopore lumen, which is just wide enough to accommodate a single strand of DNA. The DNA displacement associated with the translocation is detected by the use of a reporter group comprising five consutive abasic ($1'$, $2'$-H) residues in the template strand (shown as solid red circles in Figure 2B and as red Xs in Fig. 3.3 A). When a complex is captured atop the pore, movement of the DNA relative to the enzyme results in a displacement of the template strand reporter group in the nanopore lumen, which, in turn, is manifested as a change in the measured ionic amplitude.

DNAP-DNA complexes are captured from the bulk phase in the *cis* chamber, one complex at a time; tens to hundreds of individual complexes are examined, sequentially, in each experiment. Each captured phi29 DNAP-DNA complex resides atop the nanopore for several seconds (Fig. 3.2 B, ii), during which the measured ionic current fluctuates on the millisecond time scale between two amplitude levels (Fig. 3.2 B, ii; Figure 3.3 B). These fluctuations are due to movement of the DNA substrate relative to the enzyme; the distance of this displacement is ~ one nucleotide [24]. The observed

fluctuations between the two amplitude levels correspond to fluctuations across the translocation step [24]. In the complexes formed with DNA1-H_H, at the upper amplitude, the primer-template junction of the DNA substrate is bound in the polymerase active site, in the pre-translocation state [24]. At 180 mV, the pre-translocation state amplitude is centered at ~32 pA. At the lower amplitude, the primer-template junction of the DNA substrate resides in the polymerase active site, in the post-translocation state [24]. The post-translocation state amplitude is centered at ~26 pA at 180 mV. The amplitude fluctuations continue until the complex dissociates or is ejected, after which another individual complex can be captured.

For binary complexes formed between DNA1-H_H and either the wild type phi29 DNAP (Fig. 3.2 B) or the exonuclease-deficient mutant D12A/D66A (Fig. 3.3 B, i), the pre-translocation state at the upper amplitude and the post-translocation state at the lower amplitude are discrete kinetic states [66]. The transition in each direction across the translocation is a single kinetic step; complexes fluctuate between the two states, with two transition rates: r_1, the forward transition from the upper amplitude, pre-translocation state to the lower amplitude post-translocation state, and r_2, the reverse transition from the post-translocation state to the pre-translocation state (Fig. 3.1 A) [66]. For complexes formed with DNA1-H_H, the translocation rates r_1 and r_2 for the wild type enzyme are almost indistinguishable from those of the D12A/D66A mutant [25].

The D12A/D66A mutant lacks two of the ligands for the catalytic Me^{2+} ions in the exonuclease active site and thus has negligible exonucleolytic activity [9, 42], permitting us to conduct experiments under conditions in which DNA substrates bearing 3′-OH termini would be degraded if the wild type enzyme was used [65, 68]. When binary complexes are formed between the D12A/D66A mutant and the natural deoxyribose-terminated substrate, DNA1-H_OH, the lower amplitude still corresponds to one kinetic state, the post-translocation state [68]. But in contrast to complexes formed with DNA1-H_H, in complexes formed with DNA1-H_OH there is an additional kinetic state at the upper amplitude; the presence of this state can be observed in the time traces as pauses of tens to hundreds of milliseconds that punctuate intervals of rapid fluctuation between the two amplitudes (Fig. 3.3 B, iii) The periods of rapid fluctuation are due to transitions between the pre-translocation and post-translocation states (with rate r_1 for the forward translocation and r_2 for the reverse translocation; Fig. 3.1 A), and the pauses in the upper amplitude arise when the primer strand is transferred from the polymerase active site to the exonuclease site (r_3; Fig. 3.1 A). When the primer strand is transferred from the exonuclease active site back to the polymerase active site, (r_4; Fig. 3.1 A), the rapid fluctuations between the pre-translocation and post-translocation states resume [68].

Discrimination against complementary rNTP binding in the post-translocation state. Wild type phi29 DNAP strongly discriminates against complementary

rNTP incorporation. In primer extension assays conducted in Mn^{2+}, the steady-state kinetic parameter V_{max}/K_m was > 2 million-fold lower for rNTP incorporation than for dNTP incorporation [15]. This selective incorporation of dNTPs over rNTPs could be due to a lower binding affinity for rNTPs relative to dNTPs or due to a slower rate of progression to phosphodiester bond formation (k_{pol}) with rNTP substrates, or caused by perturbation of both of these processes. A lower binding affinity may be caused by a lower association rate constant (k_{on}[rNTP]; Fig. 3.1 A), a faster dissociation rate (k_{off}; Fig. 3.1 A), or both. The snug fit of the dNTP bound in the post-translocation state complex observed in the phi29 DNAP ternary complex crystal structure (Fig. 3.1 B) [6] indicates that an rNTP residue would not be well accommodated in the polymerase active site; in particular, stacking of the deoxyribose moiety of the incoming dNTP on the phenyl ring of Y254 leaves little space for a 2′-OH group on the sugar. This suggests that the discrimination against complementary rNTP incorporation is at least in part due to ground state binding discrimination.

The 3′-H group at the primer terminus of DNA1-H_H (Fig. 3.3 B) significantly inhibits the onset of phi29 DNAP-catalyzed exonucleolytic digestion of the primer strand [24, 68]. This inhibition of exonucleolytic digestion affords protection for DNA in the bulk phase, permitting experiments with wild type phi29 DNAP under conditions in which substrates bearing 2′-H, 3′-OH termini are rapidly degraded. In addition, complexes formed with DNA1-H_H support

the formation of phi29 DNAP-DNA-dNTP ternary complexes but not the chemical step of phosphodiester bond formation, allowing the direct comparison of dNTP or rNTP binding affinities and binding rates among complexes [24, 67, 25]. Complementary dNTP binds with high affinity to complexes in the post-translocation state but has negligible affinity for complexes in the pre-translocation state; the dNTP association and dissociation rates are independent of the applied voltage [24, 67]. The transition from the pre-translocation state to the post-translocation state and dNTP binding are sequential (Fig. 3.1 A); dNTP can bind to phi29 DNAP-DNA complexes only after the transition from the pre-translocation state to the post-translocation state, and the transition back to the pre-translocation state cannot occur before the dissociation of dNTP [67].

Like complementary dNTP, complementary rNTP stabilizes the post-translocation state in a concentration-dependent manner (Fig. 3.4 A and Fig. S1, c and d), while having no effect on the dwell time of complexes in the pre-translocation state (Fig. S2). We compared the binding affinity of phi29 DNAP for complementary dNTP or complementary rNTP in titration experiments, using complexes formed with DNA1-H_H, captured at 180 mV in 10 mM MgCl$_2$ (Fig. 3.4 A). The data are displayed by plotting the normalized p/(1-p), where p is the probability of post-translocation state occupancy, and the normalized p/(1-p) is defined as the value of p/(1-p) in the presence of a given concentration of dNTP or rNTP, divided by the value of p/(1-p) for the same

phi29 DNAP-DNA complex in the absence of dNTP or rNTP [24] (see also Supporting Information) The normalized p/(1-p) is independent of the transitions between the two translocation states in the absence of dNTP or rNTP; the effect of these transitions is eliminated when p/(1-p) is normalized by its value measured in the absence of dNTP or rNTP. Based on the model shown in Fig. 3.1 A, we derive that (normalized p/(1-p) - 1) = [d/rNTP]/K_d, where [d/rNTP] refers to the concentration of dNTP or rNTP (see Supporting Information). This theoretical expression predicts that the log-log plot of (normalized p/(1-p) - 1) vs [d/rNTP] is a straight line with slope = 1 and vertical intercept = -log(K_d) at [dNTP] or [rNTP] =1 µM. K_d is determined from the vertical intercept obtained in fitting observed data points to the theoretical expression. Therefore, the log-log plot of (normalized p/(1-p) - 1) vs [d/rNTP] allows direct comparison of the post-translocation state binding affinities among the complexes: a higher vertical position in the log-log plot corresponds to a smaller value of K_d (Figure 4A).

The discrimination by phi29 DNAP against rNTP incorporation is due at least in part to lower rNTP binding affinity; in complexes formed between the wild type enzyme and DNA1-H_H, concentration-dependent stabilization of the post-translocation state by complementary rNTP is substantially diminished relative to complementary dNTP (Fig. 3.4A). The K_d for rNTP binding to wild type complexes is ~ 1300-fold greater than the K_d for dNTP binding (Table 4.1). The D12A/D66A enzyme displays a level of binding discrimination

88

against rNTP similar to the wild type enzyme; the K_d for rNTP binding to the D12A/D66A enzyme is ~ 700-fold larger than the K_d for dNTP binding to this enzyme (Table 4.1). As we have previously shown, the D12A/D66A mutations themselves yield a modest (< 2-fold) increase in complementary dNTP binding affinity, when compared to the wild type enzyme [24] (Fig. 3.4, Table 3/1). The D12A/D66A mutations cause an increase in binding affinity of similar magnitude for complementary rNTP, relative to the affinity of the wild type enzyme for rNTP (Fig. 3.4, Table 3.1).

Introduction of the Y254V mutation partially relieves the discrimination by phi29 DNAP against rNTP incorporation; the steady-state kinetic parameter V_{max}/K_m for rNTP incorporation measured in Mn^{2+} was ~ 3 orders of magnitude higher for the Y254V mutant than for the wild type enzyme [15]. This partial alleviation of the discrimination against rNTP incorporation by the Y254V mutant is caused in part by an increase in rNTP binding affinity; the K_d for rNTP binding to the post-translocation state of complexes formed with the Y254V mutant was ~ 20-fold smaller than the K_d for complexes formed with the wild type enzyme (Fig. 3.4; Table 3.1). Similarly, the K_d for rNTP binding for the D12A/D66A/Y254V mutant was ~ 13-fold lower than the K_d for complexes with the D12A/D66A enzyme. The presence of the D12A/D66A mutations in the D12A/D66A/Y254V mutant causes a modest increase in binding affinity (decrease in K_d value) for both dNTP and rNTP, relative to the

binding affinities for each of these ligands in complexes formed with the

Y254V mutant (Fig. 3.4; Table 3.1).

Table 3.1 legend:

[a] The dNTP or rNTP association rate constant. [b] The dNTP or rNTP dissociation rate. [c] K_d values are determined from the vertical intercepts of the fitting lines to the log-log plot of (Normalized p/(1–p) –1) vs. [dGTP] or [rGTP], where p is equilibrium probability of the lower amplitude level (see the model diagram in Figure 1a). The plots of (Normalized p/(1–p) –1) vs. [dGTP] or [rGTP] are shown in Figure 5 and Figure 7. [d] Not determined. Rates were determined using dwell time samples extracted from ionic current traces (see Methods section) and a three-state kinetic model (consisting of transitions r_1, r_2, k_{on} and k_{off} in the model diagram in Figure 1a). Experiments were conducted at 180 mV; k_{on}[dNTP] and k_{off} are independent of the applied voltage [67]. All values are reported with the standard error.

Table 3.1. Complementary dNTP or rNTP binding rates for wild type phi29 DNAP and mutants.

enzyme	DNA	ligand	$k_{\text{on}}(\text{s}^{-1}\mu\text{M}^{-1})$ [a]	$k_{\text{off}}(\text{s}^{-1})$ [b]	$k_{\text{off}}(\text{s}^{-1})$ / $k_{\text{on}}(\text{s}^{-1}\mu\text{M}^{-1})$	$K_{\text{d}}(\mu\text{M})$ [c]
wt	DNA1-H_H	dGTP	21.4 ± 0.6	31.9 ± 0.5	1.49 ± 0.03	1.31 ± 0.04
Y254V	DNA1-H_H	dGTP	19.7 ± 0.9	140 ± 4	7.3 ± 0.3	6.59 ± 0.2
D12A/ D66A	DNA1-H_H	dGTP	16.7 ± 0.4	17.7 ± 0.3	1.02 ± 0.02	0.99 ± 0.04
D12A/ D66A/ Y254V	DNA1-H_H	dGTP	32.4 ± 1.0	126 ± 2.3	3.9 ± 0.2	4.76 ± 0.17
wt	DNA1-H_H	rGTP	> 3.2	> 5400	ND[d]	1702 ± 239
Y254V	DNA1-H_H	rGTP	8.9 ± 1.8	777 ± 16	91 ± 24	83.6 ± 3.6
D12A/ D66A	DNA1-H_H	rGTP	> 7.1	> 4900	ND[d]	695 ± 32
D12A/ D66A/ Y254V	DNA1-H_H	rGTP	12.6 ± 1.9	546 ± 32	45.2 ± 4.2	53.1 ± 1.6
wt	DNA1-OH_H	dGTP	8.9 ± 1.8	526 ± 28	60.8 ± 9.1	59.4 ± 7.3
Y254V	DNA1-OH_H	dGTP	12.7 ± 0.3	1126 ± 39	85.7 ± 1.8	86.4 ± 6.1
D12A/ D66A	DNA1-OH_H	dGTP	9.9 ± 0.6	318 ± 26	32 ± 1.0	27.2 ± 1.7
D12A/ D66A/ Y254V	DNA1-OH_H	dGTP	14.8 ± 0.2	897 ± 19	58.8 ± 1.3	63.4 ± 2.6
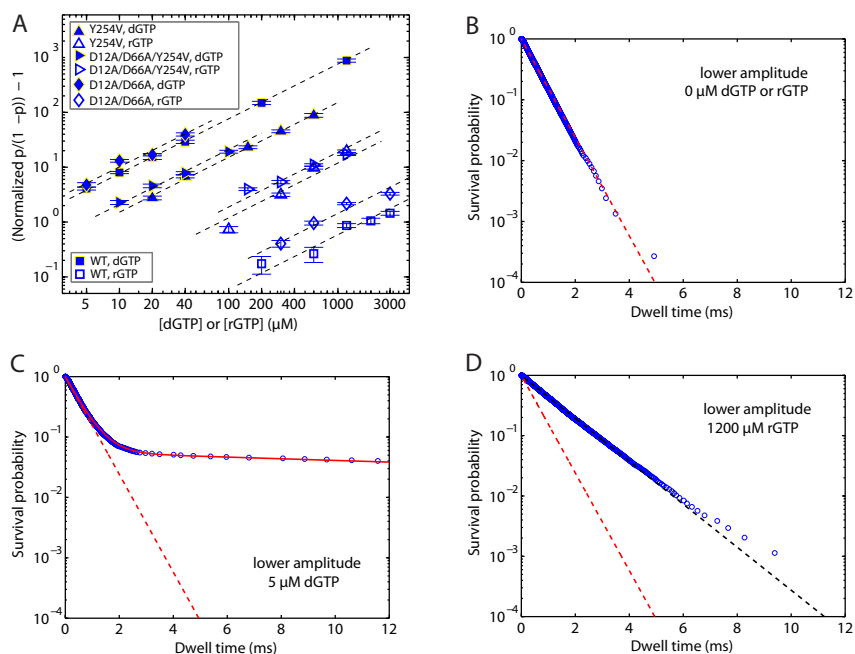
Figure 3.4: Complementary dNTP or rNTP binding to wild type phi29 DNAP and mutants. (**A**) The (normalized p/(1-p) − 1) is shown in a log-log plot as a function of the concentration of dGTP (filled symbols) or of rGTP (open symbols), for complexes formed between DNA1-H_H and wild type phi29 DNAP or mutants. Enzyme identities are indicated by symbol shapes: the wild type phi29 DNAP (squares), the Y254V mutant (upright triangles), the D12A/D66A mutant (diamonds), or the D12A/D66A/Y254V mutant (sideways triangles). The normalized p/(1-p) is defined as the value of p/(1-p) in the presence of a given concentration of dNTP or rNTP, divided by the value of p/(1-p) for the same phi29 DNAP-DNA complex at 0 μM dNTP or rNTP [24] where p is the probability of post-translocation state occupancy. Complexes were captured at 180 mV. Error bars show the standard errors. Each data point was determined from 15–30 ionic current time traces for individual captured complexes; each time trace had a duration of 5–10 s. (**B-D**) Plots of log(survival probability) vs. dwell time for the lower amplitude level for complexes formed between wild type phi29 DNAP and DNA1-H_H, captured in the presence of (**B**) 0 μM dGTP or rGTP, (**C**) 5 μM dGTP, or (**D**) 1200 μM rGTP. In panels b-d the dashed red line represents an exponential distribution with rate r_2, which is the constant slope of log(survival probability) at 0 μM dGTP or rGTP, and which is obtained by fitting to the data in Figure 4b (see text). The solid red fitting line in panel c shows the fit of the data to a model of two exponential modes. The dashed black fitting line in panel d shows the fit of the data to an exponential distribution. The dwell time samples were extracted from data files collected when complexes were captured at 180 mV; each file yields ~8,000-80,000 dwell time samples for each amplitude level. In the plots, while 1 out of every 20 points is shown, the curves are fit to the full set of dwell time samples.

Y254 is located in the polymerase active site and is directly involved in structural transitions critical to both the translocation and dNTP binding [6, 25, 14, 86]. While the Y254V mutation can relieve the steric clash with rNTP substrates, it does so with the concomitant loss of the stacking interaction with the sugar moiety of the bound nucleotide; the loss of this interaction would be predicted to significantly perturb dNTP binding. When the effects of the Y254V mutant were initially examined in Mn^{2+}, the steady-state kinetic parameter V_{max}/K_m for complementary dNTP incorporation was diminished < 2-fold by the presence of the mutation [15]. When the effects of introducing the Y254V mutation on dNTP incorporation are examined in Mg^{2+}, a more prominent impairment of dNTP incorporation is observed. In the context of a wild type exonuclease site, the Y254V mutant requires an ~ 8-fold higher concentration of dNTPs than is required by the wild type enzyme to shift the equilibrium between the polymerization and exonucleolytic reactions and achieve a similar primer extension efficiency (Fig. S3a). When measured in an exonuclease-deficient background, the D12A/D66A/Y254V enzyme requires an ~ 8-fold higher concentration of dNTPs to support a similar level of primer extension as does the D12A/D66A enzyme (Fig. S3b). In accord with this requirement for a higher concentration of dNTPs, the K_d value for dNTP binding was increased ~ 5-fold by the introduction of the Y254V mutation, in both the wild type or D12A/D66A backgrounds (Fig. 3.4; Table 3.1). Thus, the Y254V mutation diminishes binding discrimination against

rNTP relative to dNTP in two distinct ways: it increases the binding affinity for complementary rNTP while simultaneously decreasing the binding affinity for complementary dNTP.

Kinetic mechanism of discrimination against rNTP binding. The predicted steric occlusion of the 2'-OH group in post-translocation state ternary complexes formed with complementary rNTP [6] could yield a decrease in k_{on}, an increase in k_{off}, or a change in both rates, relative to complementary dNTP. To understand the kinetic mechanism that yields the low rNTP binding affinity of phi29 DNAP, we sought to compare k_{on} and k_{off} for dNTP and rNTP, calculated from dwell time samples of each amplitude level extracted from measured time traces of ionic current. The calculation of k_{on} and k_{off} is based on the mathematical formulation of a three-state model for translocation and nucleotide triphosphate binding, in which r_1, r_2, k_{on}, and k_{off} are determined simultaneously from the data [67]. We have shown that dNTP has no effect on the translocation rates, and that the dNTP association and dissociation rates are independent of the applied voltage (force) [67]. This three-state model has proven robust for determining dNTP binding rates [67, 25]. However, the rNTP binding rates could not be determined for complexes formed between DNA1-H_H and the wild type enzyme or the D12A/D66A enzyme using this method. To understand why we could not determine the rNTP binding rates, we examined the effects of complementary rNTP (rGTP) on the observed dwell time samples. We consider the escape problem of exiting from the lower

amplitude level. Let S(t) be the probability of dwell time $\geq t$ (survival probability). Plots of S(t) vs. t at various dGTP or rGTP concentrations are shown in Fig. 4, B-D and in Fig. S1, a-c, for complexes formed between DNA1-H_H and the wild type enzyme. In these plots, the survival probability S(t) is shown in the logarithmic scale. Based on the model in Fig. 3.1 A, the slope of log(S(t)) vs. t has the expression

$$-\frac{d\log\big(S(t)\big)}{dt} = r_2 \cdot \frac{S_{\text{Post-unbound}}(t)}{S(t)}$$

where $S_{\text{Post-unbound}}(t)$ is the probability of surviving in the lower amplitude level beyond time t and being in the unbound state at time t. The quantity Q(t) = $S_{\text{Post-unbound}}(t)/S(t)$ is the fraction of the survival probability in the unbound state. At t = 0, Q(0) = 1. In the presence of dGTP (or rGTP), as t increases, the relative fractions of unbound and the bound states reach an equilibrium and Q converges to Q(∞), which is smaller than 1. Correspondingly, the slope of log(S(t)) decreases from r_2 at t = 0 to r_2Q(∞) at large time. For complexes formed between DNA1-H_H and the wild type enzyme, plots of log(S(t)) are shown for binary complexes (0 μM dGTP or rGTP; Fig. 3.4 B), and for complexes captured in the presence of 5 μM dGTP (Fig. 3.4 C), or in the presence of 1200 μM rGTP (Fig. 3.4 D). The dashed red line in each of the three plots has slope r_2, which is the constant slope of log(S(t)) at 0 μM dGTP or rGTP, and which is obtained by fitting to the data in Fig. 3.4 B. Plots of log(S(t)) at additional concentrations of dGTP or rGTP are shown in Figure

96

S1. Fig. 3.4 C shows clearly the decrease of slope from r_2 to a much smaller value in the plot of log(S(t)) at 5 µM dGTP; the solid red line shows the fitting to a model of two exponential modes. In contrast, in the presence of rGTP, the transition in the slope of log(S(t)) is not observed, as shown for 1200 µM rGTP (Fig. 3.4 D). Instead, the observed slope is a constant smaller than r_2; the dashed black line shows a single exponential fitting to the observed dwell time samples.

When $k_{off} \gg r_2$, Q relaxes to equilibrium much faster than the transition rate r_2 and the transition in the slope occurs in a very short time, too short to be resolved in our current experiments. Nevertheless, we can estimate how large k_{off} has to be given that the transition in the slope is absent in Fig. 3.4 D. In extracting dwell time samples, to weed out spurious transitions caused by noise, we set a lower threshold of 100 µs. Transitions with dwell time < 100 µs at the target amplitude level are rejected; all accepted dwell time samples are shifted by 100 µs and these observed dwell time samples represent the dwell time beyond the threshold. Thus, a more precise meaning of S(t) plotted in Fig. 3.4 B-D is S(t) = conditional probability of dwell time $\geq t + 100$ µs given that dwell time ≥ 100 µs. In Fig. 3.4 D, the observed slope stays at a constant level and is already smaller than r_2. This observation indicates that the transition in slope from r_2 to a smaller value has already occurred within the 100 μs that was cut off. On the other hand, the rate of Q relaxing to equilibrium is bounded by ($r_2 + k_{on}[rGTP] + k_{off}$), which is the sum of all rates

involved in the escape problem of exiting from the lower amplitude level (see

Supplemental Information). Combining these 2 results, we obtain

$$(r_2 + k_{on}[rGTP] + k_{off}) > -\log(0.33)/(100 \text{ μs})$$

where the factor -log(0.33) reflects the assertion that at least two-thirds (67%)

of the transition in slope has occurred in 100 μs. For complexes formed

between DNA1-H_H and the wild type enzyme, $r_2 = 1868$ s^{-1} at 180 mV, and

K_d (for rGTP binding) = 1702 μM (Table 1). Fig. 3.4 D shows that at 1200 μM

rGTP the magnitude of transition in slope is substantial (the difference

between the observed constant slope and r_2 is substantial) but the transition

is too fast to be observed after the 100 μs cut-off at the lower end.

Substituting these values into the inequality above, we arrive at $k_{off} > 5400$ s^{-1}.

Accordingly, $k_{on} = k_{off}/K_d > 3.18$ s^{-1}(μM)$^{-1}$. These are lower bounds for k_{on} and

k_{off}. It is reasonable to expect that k_{on} for rNTP does not exceed k_{on} for dNTP.

For complexes formed between wild type phi29 DNAP and DNA1-H_H, $k_{on} =$

21.38 ± 0.62 s^{-1}μM^{-1} for dGTP (Table 3.1). In the case that k_{on} for rGTP

binding is the same as k_{on} for dGTP binding, and that the increase in K_d value

for rGTP relative to dGTP is entirely attributable to the increase in k_{off}, the

dissociation rate for rGTP would be $k_{off} = k_{on} \times K_d$ (for rGTP) = 36346 s^{-1},

which can be viewed as an upper bound for the rate of dissociation of rGTP

from wild type phi29 DNAP complexes.

The plots of log(S(t)) vs. t for rNTP binding to complexes formed between the

D12A/D66A enzyme and DNA1-H_H are very similar to the plots for rNTP

binding to wild type complexes (Fig. S2, c and d); the transition in slope from r_2 to a smaller value has already occurred within 100 μs, the cut-off threshold used in extracting dwell time samples. An analysis similar to the one above yields that for rNTP binding to D12A/D66A complexes, $k_{off} > 4900$ s^{-1}. The affinity of rNTP binding to D12A/D66A complexes is $K_d = 695$ μM (Table 3.1). Accordingly, for rNTP binding to D12A/D66A complexes, $k_{on} = k_{off} /K_d > 7.05$ s^{-1}μM^{-1}. For complexes formed between the D12A/D66A enzyme and DNA1-H_H, $k_{on} = 16.70 \pm 0.4$ s^{-1}μM^{-1} for dGTP (Table 3.1). In the case that k_{on} for rGTP binding is the same as k_{on} for dGTP binding and the increase in K_d value for rGTP relative to dGTP is solely attributable to the increase in k_{off}, we calculate $k_{off} = k_{on} \times K_d$ (for rGTP) = 11607 s^{-1}, which can be viewed as an upper bound for the rGTP dissociation rate for the D12A/D66A mutant.

The large increase in k_{off} suggests that the majority of encounters with rNTP do not progress to phosphodiester bond formation. While we do not have k_{pol} values for the incorporation of either dNTP or rNTP by phi29 DNAP, values determined for other B-family DNAPs may provide a plausible range: for dNTP incorporation by T4 DNAP, $k_{pol} = 400 \pm 4$ s^{-1} [18] and for dNTP incorporation by human Pol-ε $k_{pol} = 248 \pm 6$ s^{-1} [111]. For the B-family DNAP from bacteriophage RB69, which discriminates against rNTP incorporation by 64,000-fold, $k_{pol} = 200 \pm 13$ s^{-1} for dNTP incorporation and $k_{pol} = 0.74 \pm 0.2$ s^{-1} for rNTP incorporation [108], a 270-fold decrease in k_{pol} for rNTP incorporation. It is highly probable that k_{pol} for rNTP incorporation by phi29 DNAP is also

significantly slower than k_{pol} for dNTP incorporation; a decrease in k_{pol} would conspire with the increased rNTP dissociation rate to yield the substantial discrimination exhibited by phi29 DNAP against rNTP incorporation.

Kinetic mechanisms by which the Y254V mutation diminishes discrimination against rNTP binding. In contrast to complexes formed with the wild type or D12A/D66A enzymes, for rNTP binding to complexes formed with the Y254V or the D12A/D66A/Y254V mutants, the transition in the slope of log(S(t)) can be resolved in the observed dwell time samples. As a result, the values of k_{on} and k_{off} for rGTP binding to these mutants were determined from the data (Table 3.1). Introduction of the Y254V mutation into either the wild type or the D12A/D66A background diminishes the discrimination against rNTP binding by causing both an increase in affinity for rNTP and a decrease in affinity for dNTP (Fig. 3.4 A; Table 3.1). The 20-fold increase in the binding affinity of the Y254V mutant for rNTP relative to the wild type enzyme is primarily due to a decrease in the rNTP dissociation rate. While the plausible range for the rate of rNTP dissociation from the wild type enzyme is 5400 to ~ 36346 s^{-1}, the rate of rNTP dissociation from the Y254V mutant is k_{off} = 776.8 ± 16.4 s^{-1}. Similarly, while the plausible range for the rate of rNTP dissociation from the D12A/D66A enzyme is 4900 to ~ 11607 s^{-1}, the rate of rNTP dissociation from the D12A/D66A/Y254V mutant is k_{off} = 546.0 ± 31.8 s^{-1}, indicating that the 13-fold increase in the binding affinity of the D12A/D66A/Y254V mutant for rNTP

relative to the D12A/D66A enzyme is also due to a decrease in the rNTP dissociation rate.

In both the wild type and the D12A/D66A backgrounds, the ~ 5-fold decrease in dNTP binding affinity caused by the Y254V mutation is largely due to an increase in the dNTP dissociation rate (Table 3.1). While there is negligible difference in k_{on} for dNTP binding attributable to the Y254V mutation, the dNTP dissociation rate of $k_{off} = 140.2 \pm 3.9$ s^{-1} for the Y254V mutant is a ~4.4-fold increase over $k_{off} = 31.88 \pm 0.49$ s^{-1} for the wild type enzyme. There is a modest (< 2-fold) increase in k_{on} and an ~ 7-fold increase in k_{off} for dNTP binding to the D12A/D66A/Y254V enzyme relative to the D12A/D66A enzyme. The Y254V mutation may also cause a decrease in k_{pol} for dNTP, which together with the increase in k_{off} of the Y254V mutant for dNTP could yield the requirement for a higher concentration of dNTPs to achieve efficient primer extension (Fig. S3).

Kinetic consequences of introducing a primer terminal 2′-OH group in the context of a 3′-H terminated DNA substrate. Immediately following covalent incorporation of an rNTP residue, the product occupies the polymerase site in the pre-translocation state. This state is the branchpoint between the DNA synthesis and editing pathways, and we sought to understand the consequences of introducing a ribose sugar at the primer terminus on the fate of phi29 DNAP binary complexes at this critical juncture. We first examined the effects of introducing a 2′-OH group at the primer terminus by comparing

complexes formed with DNA1-OH_H to complexes formed with DNA1-H_H. Inspection of ionic current traces for complexes formed between DNA1-OH_H and the wild type enzyme or the D12A/D66A enzyme (Fig. 3.3 B, ii) reveals that, in contrast to complexes formed with DNA1-H_H (Fig. 3.3 B, i), intervals of rapid fluctuation between the two amplitudes are punctuated by pauses at the upper amplitude. The pauses are reminiscent of those observed in the time traces for complexes formed between the D12A/D66A enzyme and DNA1-H_OH (Fig. 3.3 B, iii) but appear to be of much shorter duration.

For binary complexes formed between DNA1-OH_H and the wild type enzyme or the D12A/D66A enzyme, survival probability vs. dwell time plots for the lower amplitude fit well to a single exponential distribution (Fig. S4d, i and ii), indicating that, like the complexes formed with DNA1-H_H (Figure S4c, i and ii), the lower amplitude corresponds to one kinetic state, the post-translocation state. In contrast, survival probability vs. dwell time plots of the upper amplitude for complexes formed between DNA1-OH_H and the wild type enzyme or the D12A/D66A enzyme are inconsistent with a single exponential distribution, but are well explained by a model of two exponential modes (Fig. S4b, i and ii). Thus, the pauses at the upper amplitude for the complexes formed with DNA1-OH_H (Fig. 3.3 B, ii; Fig. S4b, i and ii), which are not apparent when complexes are formed between DNA1-H_H and either wild type or D12A/D66A phi29 DNAP (Fig. 3.3 B, i; Fig. S4a, i and ii), correspond to an additional kinetic state. We have shown that this additional

102

kinetic state in the upper amplitude is one in which the primer strand occupies the exonuclease site [68].

The substantial increase in k_{off} for complementary rNTP relative to complementary dNTP (Fig. 3.4; Table 3.1) indicates that in the ground state, the rNTP-bound complex is significantly destabilized relative to the dNTP-bound complex. Based upon a model for the structure of the pre-translocation state in which the terminal base pair of the duplex occupies the same configuration as the nascent base pair does in the post-translocation state ternary complex [6], we therefore hypothesized that an analogous destabilization of the pre-translocation state would be caused by the presence of a $2'$-OH group at the primer terminus. This hypothesis predicts that the rates of the transitions out of the pre-translocation state, the forward translocation (r_1) and the primer strand transfer to the exonuclease site (r_3), would increase upon the introduction of a $2'$-OH group at the primer terminus. As a first test of this prediction, compared the dynamic transitions of binary complexes formed with DNA1-H_H or DNA1-OH_H (Fig. 3.5). We used dwell time samples extracted from ionic current traces and a two-state kinetic model [66] to determine the translocation rates for complexes formed with DNA1-H_H; we used a three-state model for translocation and primer strand transfer between the polymerase and exonuclease sites [68] to determine the translocation rates and primer strand transfer rates for complexes formed with DNA1-OH_H.
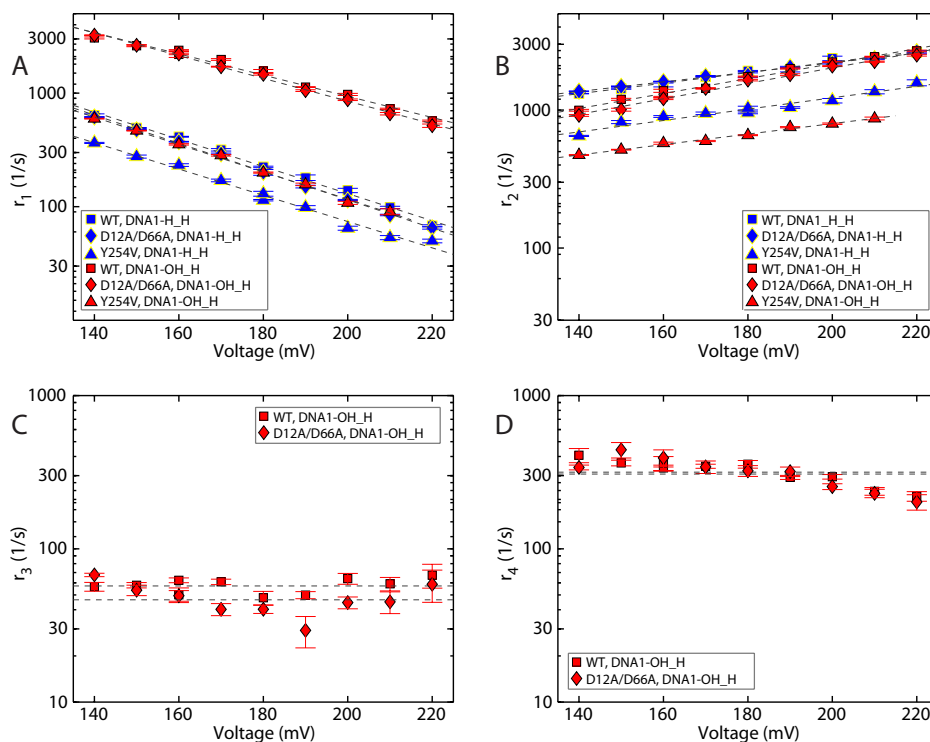
Figure 3.5: Effects of a 2′-OH, 3′-H primer terminated substrate on the rates of translocation and primer strand transfer between the polymerase and exonuclease sites. Plots of (**A**) $\log(r_1)$ vs voltage, (**B**) $\log(r_2)$ vs voltage, (**C**) $\log(r_3)$ vs voltage, and (**D**) $\log(r_4)$ vs voltage for complexes formed between DNA1-H_H (blue symbols) or DNA1-OH_H (red symbols) and wild type (squares), D12A/D66A (diamonds), or Y254V (upright triangles) phi29 DNAP. Rates were determined from dwell time samples extracted from ionic current traces and a three-state kinetic model [40] consisting of transitions r_1, r_2, r_3 and r_4 in the model diagram in Figure 1a. Errors bars indicate the standard error.

The translocation rates and primer strand transfer rates are plotted as function of applied voltage (Fig. 3.5). The voltage applies a force along the direction of the translocation, which impedes the rate of the forward translocation (r_1) and increases the rate of the reverse translocation (r_2). The slope of $\log(r_1)$ vs. voltage is negative and proportional to the distance between the pre-translocation state and the transition state in the translocation step; the slope of $\log(r_2)$ vs. voltage is positive and proportional

to the distance between the transition state and the post-translocation state [66]. The rates of primer transfer from the polymerase to the exonuclease site (r_3) and from the exonuclease site to the polymerase site (r_4) do not vary with voltage; this pathway is not associated with a spatial displacement along the direction of the applied force [68].

The forward (r_1) and reverse (r_2) translocation rates in complexes formed by the wild type enzyme or the D12A/D66A enzyme with DNA1-H_H are almost indistinguishable [25], and this is also the case for complexes formed with DNA1-OH_H (Fig. 3.5 A and B). Consistent with the prediction that the presence of a $2'$-OH group at the primer terminus would destabilize the pre-translocation state, r_1 was significantly faster in complexes formed with DNA1-OH_H and the wild type enzyme or the D12A/D66A enzyme than in complexes formed with DNA1-H_H (Fig. 3.5 A). For example, comparing the values at 180 mV for both the wild type and D12A/D66A enzymes, the presence of the primer terminal $2'$-OH caused an ~ 7-fold increase in r_1. The primer terminal $2'$-OH has modest effects on the magnitude of r_2 (Fig. 3.5 B).

In complexes formed with DNA1-OH_H, the rates of primer strand transfer from the polymerase to exonuclease site (r_3; Fig. 3.5 C) for the wild type and D12A/D66A enzymes are very similar. This is also the case for the rates of primer strand transfer from the exonuclease to polymerase site (r_4; Figure 5D). As we have shown for complexes formed with DNA1-H_OH [40], in complexes formed with DNA1-OH_H neither r_3 nor r_4 displays a systematic

105

trend with the applied voltage (Fig. 3.5 C and D) suggesting that r_3 and r_4 are not affected by the applied force. We therefore treat the data points at individual voltages as independent samples, and calculate the mean and standard error of each rate, for each enzyme. In complexes formed between wild type phi29 DNAP and DNA1-OH_H, $r_3 = 57.58 \pm 2.05$ s^{-1} and $r_4 = 321.6 \pm 18.77$ s^{-1}; in complexes formed between the D12A/D66A mutant and DNA1-OH_H, $r_3 = 47.65 \pm 3.75$ s^{-1} and $r_4 = 316.02 \pm 25.6$ s^{-1}.

In contrast to the D12A/D66A mutant, the exonuclease site in the wild type enzyme is catalytically active. Nonetheless, in the complexes formed between the wild type enzyme and DNA1-OH_H, the rate of primer strand transfer from the exonuclease to polymerase site (r_4) measured when complexes are held atop the pore is for the case when the primer strand is returned to polymerase site uncleaved. The return of the uncleaved primer strand to the polymerase site is indicated by the subsequent resumption of fluctuations between the amplitudes that are characteristic of the pre-translocation and post-translocation states for the intact DNA substrate. Cleavage of the primer strand is rare in the complexes formed between the wild type enzyme and DNA1-OH_H, and if it occurs, can be readily detected in the ionic current traces. If the primer terminal residue is cleaved, the position of the reporter group in the nanopore lumen is shifted by the distance of 1 nucleotide, and this yields a new set of two amplitude clusters for the pre-translocation and post-translocation states for the cleaved DNA substrate that is shifted from

the set for the substrate before the cleavage [68]. The transition rates r_3 and r_4 are determined from time traces prior to cleavage of the primer terminal residue, and complexes formed between DNA1-OH_H and the wild type or the D12A/D66A enzymes have nearly identical values for r_4.

In addition to the increase in r_1 predicted to occur upon introduction of a primer terminal 2'-OH group, the rate of primer strand transfer to the exonuclease site (r_3) is also predicted to increase. Because we cannot reliably detect an upper amplitude dwell time sub-population corresponding to the state in which the primer strand occupies the exonuclease site for complexes formed with DNA1-H_H (Fig. S4 a, i and ii), we cannot directly compare values of r_3 and r_4 in complexes with DNA1-OH_H to values in complexes with DNA1-H_H. The lack of a separate sub-population in the upper amplitude dwell time for complexes formed with DNA1-H_H is consistent with the case where r_3/r_4 is very small (either r_3 is small or r_4 is large, or both).

Effects of the Y254V mutation in complexes formed with 3'-H terminated DNA substrates. While Y254 functions to sterically occlude rNTP binding to post-translocation state complexes, it also participates in dynamic structures in the polymerase active site directly involved in the translocation [6]. In the post-translocation state binary complex, the side chains of Y254 and Y390 are stacked; this interaction is presumably disrupted by the Y254V mutation. In the post-translocation state ternary complex, the side chains of Y254 and Y390 are both rotated relative to their positions in the binary complex, which

disrupts the stacking interaction between them and allows the deoxyribose sugar of the dNTP to stack on Y254. Thus in the pre-translocation state binary complex, where the terminal primer strand residue is predicted to occupy the configuration that is occupied by incoming dNTP in the post-translocation state ternary complex, the Y254V mutation would alter the stacking interaction between the Y254 phenyl ring and the primer terminal sugar moiety. Since binary complex interactions in both the pre-translocation state and post-translocation state are perturbed by the Y254V mutation, and because of the intimate involvement of Y254 in the active site movements predicted to accompany the translocation displacement [6], it is reasonable to expect that the Y254V mutation would affect the translocation rates. Indeed, introduction of the Y254V mutation in complexes formed with DNA1-H_H causes an ∼ 1.75 to 2-fold decrease in r_1 (Fig. 3.5 A) and an ∼ 1.75 to 2-fold decrease in r_2 (Fig. 3.5 B) relative to the wild type enzyme. The effects of the Y254V mutant on r_1 and on r_2 are of similar magnitude, yielding little change in the equilibrium across the translocation step in binary complexes.

If the terminal base pair of the duplex occupies the same configuration in the pre-translocation state that the nascent base pair occupies in the post-translocation state ternary complex[6] the destabilization of the pre-translocation state by a primer terminal 2′-OH group would be predicted to be attenuated by the Y254V mutation, via a structural mechanism analogous to the mechanism by which the Y254V mutation increases rNTP binding affinity

in the post-translocation state (Fig. 3.4). Consistent with this prediction, when complexes are formed between the Y254V enzyme and DNA1-OH_H, the Y254V mutation almost fully alleviates the $\sim$ 7-fold increase in r_1 caused by the introduction of the 2′-OH into complexes formed with the wild type or D12A/D66A enzymes (Fig. 3.5 A); r_1 for complexes of Y254V with DNA1-OH_H is very similar to r_1 for wild type or D12A/D66A complexes formed with DNA1-H_H. In the complexes formed with DNA1-OH_H, the Y254V mutant yields an $\sim$ 2-fold decrease in r_2 and a small decrease in the slope of $\log(r_2)$ vs. voltage, relative to the wild type or D12A/D66A complexes formed with DNA1-OH_H. These effects of the Y254V mutation on r_2 in complexes formed with DNA1-OH_H are similar to the effects of introducing the mutation in DNA1-H_H complexes.

For complexes formed between the Y254V mutant and DNA1-OH_H, primer strand transition rates (r_3 and r_4) could not be reliably determined. This is reminiscent of complexes formed between phi29 DNAP with DNA1-H_H, and it is in contrast to complexes of the wild type or D12A/D66A enzymes with DNA1-H_OH, where a second dwell time cluster at the upper amplitude could be well resolved, and r_3 and r_4 could be determined (Fig. 3.5 C and D). Inspection of a plot of survival probability vs. dwell time for the upper amplitude of complexes formed between the Y254V mutant and DNA1-OH_H shows that while it deviates from a single exponential distribution (Fig. S4b, iii), the second exponential mode that emerged upon the introduction of the

2′-OH in wild type or D12A/D66A complexes (Fig. S4b, i and ii) is significantly attenuated in the complexes formed with the Y254V mutant, which precludes a reliable estimate of r_3 and r_4 from data.

Kinetic consequences of introducing a 2′-OH group in the context of a 3′-OH terminated DNA substrate. In order to examine the effects of introducing a primer terminal 2′-OH group on the fate of complexes at the branchpoint between synthesis and editing in a context where we can simultaneously detect and quantify the translocation rates (r_1 and r_2 in Fig. 3.1 A) and the primer strand transfer rates between the polymerase and exonuclease sites (r_3 and r_4 in Fig. 3.1A), we next compared complexes using DNA substrates with natural deoxyribose (DNA1-H_OH) or ribose (DNA1-OH_OH) primer strand termini, formed with the exonuclease-deficient D12A/D66A mutant. In complexes formed with DNA1-H_H, the translocation rates (r_1 and r_2) for the D12A/D66A enzyme are nearly indistinguishable from those of the wild type enzyme ([39]; Fig. 3.5 A and 3.5 B), whereas the primer strand transfer rates r_3 (Fig. 3.5 C) and r_4 cannot be determined from the data. For complexes formed with DNA1-OH_H, the D12A/D66A enzyme and the wild type enzyme yield nearly identical results in the translocation rates, r_1 (Fig. 3.5 A) and r_2 (Fig. 3.5 B), as well as in the primer strand transfer rates r_3 (Fig. 3.5 C) and r_4 (Fig. 3.5 D). These data suggest that the wild type and D12A/D66A enzymes have very similar responses to the introduction of a 2′-OH group at the primer strand terminus. Thus, it is reasonable to expect that the effects on the kinetic

rates caused by introducing a primer terminal ribose in D12A/D66A complexes will reflect the effects of introducing a primer terminal ribose in wild type complexes.

The translocation rates and primer strand transfer rates for complexes formed with DNA1-H_OH (Fig. 3.3 B, iii) or DNA1-OH_OH (Fig. 3.3 B, iv) are plotted as a function of voltage in Figure 6. Consistent with the prediction that a ribose residue at the primer terminus would destabilize the pre-translocation state, in complexes formed between the DNA1-OH_OH and the D12A/D66A enzyme, r_1 was ~ 4.8-fold faster than in complexes with DNA1-H_OH (Fig. 3.6 A). This is in accord with the increase in r_1 caused by the $2'$-OH group in the context of a $3'$-H terminated primer strand (Fig. 3.5 A). Across the measured voltage range, the ribose-terminated primer strand causes a ~3.2-fold (lower voltage) to ~1.8-fold (higher voltage) decrease in r_2 (Fig. 3.6 B).

Unlike the case of complexes formed with DNA1-H_H, where rates r_3 and r_4 cannot be reliably estimated, for complexes formed with a natural $3'$-OH terminated primer strand we can directly assess the consequences of introducing a $2'$-OH group on the rates of primer transfer between the polymerase and exonuclease sites (Fig. 3.6 C and D). Since neither the rate of primer transfer from the polymerase site to the exonuclease site (r_3) nor the rate of primer transfer from the exonuclease site to the polymerase site (r_4) displays a systematic trend with the applied voltage (Fig. 3.6 C and D), we treat the data points at individual voltages as independent samples, and
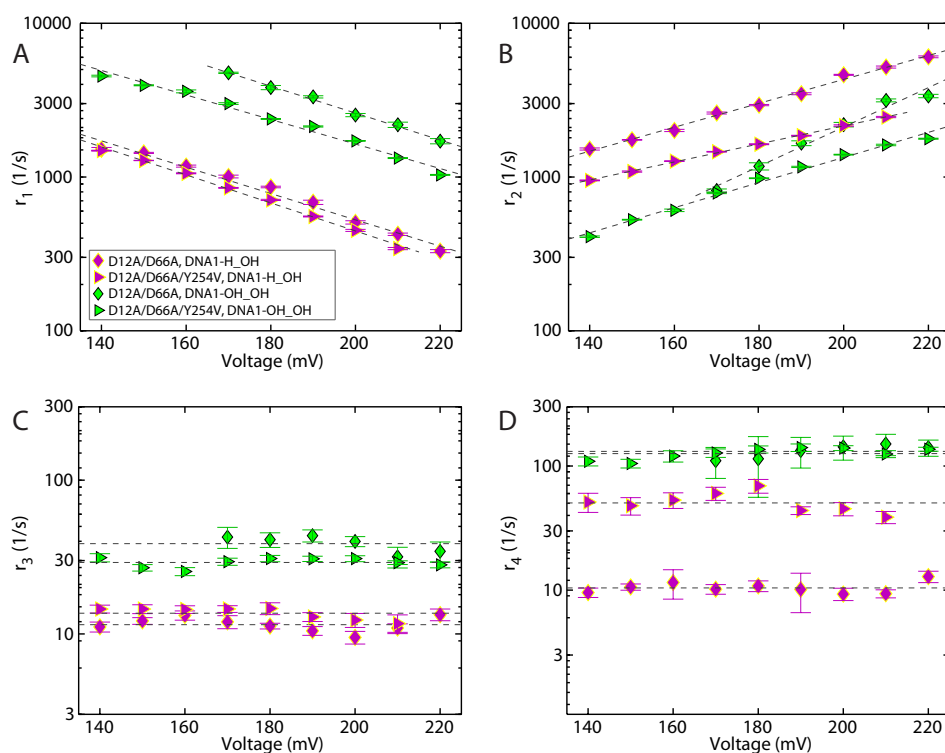
Figure 3.6: Effects of a 2′-OH, 3′-OH primer terminated substrate on the rates of translocation and primer strand transfer between the polymerase and exonuclease sites. Plots of (**A**) $\log(r_1)$ vs voltage, (**B**) $\log(r_2)$ vs voltage, (**C**) $\log(r_3)$ vs voltage, and (**D**) $\log(r_4)$ vs voltage for complexes formed between DNA1-H_OH (magenta symbols) or DNA1-OH_OH (green symbols) and D12A/D66A (diamonds), or D12A/D66A/Y254V (sideways triangles) phi29 DNAP. Rates were determined from dwell time samples extracted from ionic current traces and a three-state kinetic model [40]consisting of transitions r_1, r_2, r_3 and r_4 in the model diagram in Figure 1a. Errors bars indicate the standard error.

calculate the mean and standard error for each of these two rates, for each of the two DNA substrates. This yields r_3 = 11.54 ± 0.42 s^{-1} for D12A/D66A complexes formed with DNA1-OH_H, the deoxyribose-terminated DNA substrate, and r_3 = 38.86 ± 1.94 s^{-1} for D12A/D66A complexes formed with DNA1-OH_OH, the ribose-terminated DNA substrate DNA1-OH_OH. This is a 3.3-fold increase in the rate of primer transfer from the pre-translocation state polymerase site to the exonuclease site, caused by the presence of the 2′-OH

group at the primer terminus. Thus, the rates of both pathways out of the pre-translocation state (r_1 and r_3) are increased after the covalent incorporation of an rNTP residue.

The rate of primer strand transfer from the exonuclease to polymerase site was increased > 10-fold by the presence of a primer terminal 2′-OH group; r_4 = 10.48 ± 0.38 s^{-1} for D12A/D66A complexes formed with DNA1-H_OH, the deoxyribose-terminated substrate, and r_4 = 131.95 ± 6.64 s^{-1} for D12A/D66A complexes formed with DNA1-OH_OH, the ribose-terminated substrate. This increase in r_4 for the ribose-terminated primer strand suggests that the 2′-OH group destabilizes primer strand binding in the exonuclease site. While we cannot assess whether this increase in r_4 is of sufficient magnitude to diminish the probability of exonucleolytic cleavage, in experiments that compared exonucleolytic cleavage of single stranded oligonucleotides bearing 3′-terminal deoxyribose or ribose residues conducted with the B family T4 DNAP, as well as with the A family DNAPs, KF and T7 DNAP, cleavage was not impaired for the ribose-terminated substrates [71]. Similarly, the efficiency of exonucleolytic cleavage by wild type phi29 DNAP was not impaired by the presence of a ribose residue at the 3′ terminus of a single stranded oligonucleotide substrate [15]. These data suggest that the ribose 2′-OH group does not significantly alter the rate of exonucleolytic chemistry. Further, a complementary rNMP residue at the 3′ terminus of the primer strand did not impair cleavage efficiency by wild type phi29 DNAP relative to a

113

complementary dNMP residue [15], suggesting that if the increase in r_4 (measured in the nanopore experiments using the exonuclease-deficient D12A/D66A phi29 DNAP) also occurs when the wild type enzyme encounters a primer-template substrate with a ribose-terminated primer strand, the 2′-OH group does not render r_4 large enough to compete with the rate of exonucleolytic chemistry.

Effects of the Y254V mutation in complexes formed with 3′-OH terminated DNA substrates. In complexes formed between the D12A/D66A/Y254V enzyme and DNA1-H_OH, the DNA substrate with the deoxyribose terminated primer strand, the Y254V mutation causes a small decrease in r_1 (~ 1.2-fold), relative to complexes with the D12A/D66A enzyme (Fig. 3.6 A). This decrease in r_1 is more modest than the decrease caused by the Y254V mutation in complexes with DNA1-H_H (Fig. 3.5 A). We have previously shown that the 3′-OH is a determinant in the energy landscape of the translocation step; r_1 is faster when the primer terminus bears a 3′-OH group than when it bears a 3′-H group [68]. In the 3′-OH context, the effect of the Y254V mutation on r_1 is not prominently manifested, suggesting that the Y254V mutation does not significantly perturb the effect of the 3′-OH group in increasing r_1.

In the context of complexes formed with DNA1-H_OH, the magnitude of r_2 as well as the dependence of r_2 on opposing force is altered by the Y254V mutation in D12A/D66A/Y254V complexes, relative to D12A/D66A complexes

114

(Figure 6b). Across the measured voltage range, the Y254V mutation causes a ~1.6-fold (lower voltage) to ~2-fold (higher voltage) decrease in r_2. The decrease in r_2 is similar to the ~ 1.75 to 2-fold decrease in r_2 observed for the Y254 mutation in the context of complexes formed with DNA1-H_H and wild type or D12A/D66A enzymes (Fig. 3.5 B).

The effects of the Y254V mutation on the translocation step are intriguing, since structural predictions can be made that the mutation disrupts the stacking interaction between Y254 and Y390 in the post-translocation state binary complex, and disrupts the stacking interaction between Y254 and the sugar moiety of the primer terminal residue in the pre-translocation state binary complex [6]. Therefore, the Y254V mutation might be expected to destabilize both the pre-translocation and post-translocation states in binary complexes, leading to increases in r_1 and r_2. Instead, both r_1 and r_2 are decreased when the Y254V mutation is introduced (Fig. 3.5 A and B; Fig. 3.6 A and B). The loss of the stacking interactions in both of the translocation states makes it implausible the decreases in both the forward and reverse translocation rates occur because the Y254V mutation lowers the free energy of the pre-translocation or post-translocation state energy wells. It is likely that the Y254V mutation raises the free energy barrier of the transition state by disrupting coordination along the pathway of the translocation step.

The presence of the Y254V mutation has little effect on r_3 in complexes with the deoxyribose-terminated substrate, DNA1-H_OH (Fig. 3.6 C); for D12A/

115

D66A/Y254V complexes with DNA1-H_OH $r_3 = 13.68 \pm 0.42$ s^{-1}, compared to $r_3 = 11.54 \pm 0.42$ s^{-1} for D12A/D66A complexes with DNA1-H_OH. In contrast, the Y254V mutation causes an $\sim$ 5-fold increase in r_4 in complexes formed with the deoxyribose-terminated substrate (Fig. 3.6 D); in complexes of the D12A/D66A/Y254V enzyme with DNA1-H_OH, $r_4 = 51.16 \pm 3.45$ s^{-1} compared to $r_4 = 10.48 \pm 0.38$ s^{-1} for the D12A/D66A enzyme with DNA1-H_OH. We cannot assign a mechanistic cause for this increase in the rate of primer strand transfer from the exonuclease to polymerase site, caused by a mutation located in the polymerase active site. The Y254V mutation may alter the energy landscape across the primer strand transfer pathway, but because the rates across primer strand transfer pathway are insensitive to the force that is applied in the nanopore experimental setup, we cannot directly probe details of the energy landscape.

Consistent with the prediction that the Y254V mutation would attenuate the destabilization of the pre-translocation state by the 2′-OH group, in complexes formed between the D12A/D66A/Y254V mutant and the ribose-terminated substrate DNA1-OH_OH, the rates of both pathways out of the pre-translocation state are modestly decreased relative to the D12A/D66A enzyme with DNA1-OH_OH (Fig. 3.6 A). The rate of the pre-translocation to post-translocation state transition, r_1, is decreased $\sim$ 1.6-fold, and the rate of primer strand transfer from the polymerase to exonuclease site, r_3, is decreased $\sim$ 1.3-fold (for the D12A/D66A enzyme, $r_3 = 38.86 \pm 1.94$ s^{-1}; for

116

the D12A/D66A/Y254V enzyme, r_3 = 29.22 ± 0.68 s^{-1}). In contrast to the case

with the 3′-H terminated DNA substrates (Fig. 3.5 A), in the context of the

natural deoxyribose or ribose-terminated DNA substrates, the Y254V

mutation only partially alleviates the increase in r_1 caused by introducing the

2′-OH group (Fig. 3.6 A).

Finally, while introduction of the Y254V mutation caused an increase in r_4 for

complexes formed with DNA1-H_OH, in complexes formed with the ribose-

terminated DNA substrate, DNA1-OH_OH, it has negligible effect on r_4; in

complexes of D12A/D66A/Y254V, r_4 = 126.54 ± 4.37 s^{-1}, and in complexes

with D12A/D66A, r_4 = 131.95 ± 6.64 s^{-1}.

The direct effect of rNTP incorporation on the fate of complexes at the

branchpoint between the synthesis and editing pathways. The increase in the

rate of primer strand transfer from the polymerase to exonuclease site (r_3)

elicited by the introduction of a 2′-OH group at the primer terminus (Fig. 3.6 C)

indicates that a complementary base pair between a primer terminal rNMP

and template dNMP (the product of rNTP incorporation) is recognized as

incorrect by phi29 DNAP. The ~ 3.3-fold increase in r_3 in complexes formed

between D12A/D66A and DNA1-OH_OH relative to DNA1-H_OH is likely

modest compared to the increase in r_3 expected for a mismatched

(noncomplementary) pair at the primer terminus. Moreover, the increase in r_3

for a newly incorporated rNMP residue is accompanied by an increase in the

rate of the forward translocation (r_1). The direct effect of a newly incorporated

117

rNMP residue on the fate of the pre-translocation state complexes can be expressed as the ratio r_1/r_3. For complexes formed with the D12A/D66A enzyme and captured at 180 mV, $r_1/r_3 = 76.56 \pm 4.28$ with the deoxyribose-terminated DNA1-H_OH, and $r_1/r_3 = 93.14 \pm 13.10$ with the ribose-terminated DNA1-OH_OH. Thus for phi29 DNAP complexes formed with both of these DNA substrates, at the branchpoint between the synthesis and editing pathways the forward transition is more probable, even given that captured complexes are held atop the pore by a force that diminishes the rate of the forward translocation. Further, in spite of the increased rate of primer strand transfer to the exonuclease site elicited by a newly incorporated rNMP residue, the probability of the forward pathway is slightly increased for a complex bearing a ribose-terminated primer strand, relative to a complex bearing a deoxyribose-terminated primer strand, due to the concomitant increase in the rate of the transition from the pre-translocation to post-translocation state. These findings provide a kinetic basis for the absence of efficient exonucleolytic editing by B-family replicative DNAPs, which likely contributes to the abundant rNMP incorporation *in vivo* that can be observed prominently in cells that are comprised in RNase H-dependent repair processes [83, 77, 106].

Effects of a primer terminal 2′-OH group on dNTP binding after the translocation. Following the covalent incorporation of complementary rNTP and the forward translocation, complexes attain a state in which they are

competent to bind the next incoming dNTP [67]. It is reasonable to predict that in the post-translocation state, the presence of a 2′-OH group on the primer terminus could perturb dNTP binding (k_{on} or k_{off}), progression to phosphodiester bond formation (k_{pol},) or both of these processes. We compared complementary dNTP binding for complexes formed between phi29 DNAP and DNA1-H_H or DNA1-OH_H. For each of the four phi29 DNAP enzymes examined (wild type, D12A/D66A, Y254V, and D12A/D66A/Y254V) introduction of a 2′-OH group at the primer terminus caused a decrease in dNTP binding affinity (Fig. 3.7; Table 3.1). For the wild type and D12A/D66A enzymes, the K_d values for dNTP in complexes formed with DNA1-OH_H were increased ~ 45-fold and ~ 27.5-fold, respectively, relative to complexes formed with DNA1-H_H. The diminished affinity for the wild type
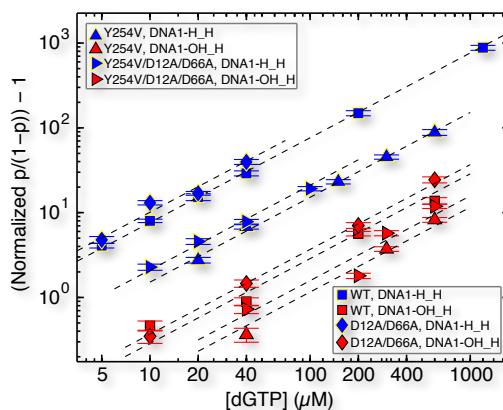
Figure 3.7: Effects of a primer terminal 2′-OH group on complementary dNTP binding to wild type phi29 DNAP and mutants. The (normalized p/(1-p)-1) is shown in a log-log plot as a function of the concentration of dGTP, for complexes formed between DNA1-H_H (blue symbols) or DNA1-OH_H (red symbols) with the wild type (squares), or Y254V (upright triangles), D12A/D66A (diamonds), or D12A/D66A/Y254V (sideways triangles) phi29 DNAP. Error bars show the standard errors.

enzyme is caused by a ~ 2.4-fold decrease in k_{on} and a larger (~ 16.5-fold)

increase in k_{off}. For the D12A/D66A enzyme, the decrease in affinity is due to

a ~ 1.7-fold decrease in k_{on} and an ~ 18-fold increase in k_{off}. The reduction in

dNTP binding caused by the primer terminal 2'-OH group is not attenuated by

the Y254V mutation (Fig. 3.7; Table 3.1); rather, in both the wild type and

D12A/D66A backgrounds, this impairment in dNTP binding appears to be

roughly additive with the impairment in dNTP binding caused by the Y254V

mutation itself (Fig. 3.4; Table 3.1).

The spatial displacement of translocation separates the state in which dNTP

binds from the state in which primer strand transfer from the polymerase to

the exonuclease site occurs (Fig. 3.1 A); dNTP binds to complexes only after

the forward translocation and has no influence on the forward or reverse rates

of the translocation, which are inherent to the binary complex [67]. Therefore a

decrease in k_{on}, an increase in k_{off}, or a decrease in k_{pol} caused by a primer

terminal 2'-OH group will reduce the net probability of the forward synthetic

pathway vs. the editing pathway after rNTP incorporation by increasing the

probability of fluctuating back to the pre-translocation state before the binding

of dNTP proceeds to covalent incorporation and carries the complex to the

pre-translocation state of the next cycle. Each time a complex fluctuates back

to the pre-translocation state, there is a new chance for the primer strand to

be transferred to the exonuclease site, with probability governed by the ratio

of r_1/r_3.

Once dNTP is covalently added onto the 2′-OH-terminated primer strand, the resulting substrate bears a complementary rNMP:dNMP pair at the -2 position of the duplex. Noncomplementary dNMPs that escape exonucleolytic editing in the -1 position of the primer terminus can continue to exert a strong influence on the probability of editing when they occupy the -2 to -4 positions of the duplex, presumably by increasing the rate of primer transfer to the exonuclease site [55, 82 and references therein]. In a future study we will examine the influence of complementary rNMP:dNMP pairs in these duplex positions on the net probability of the forward synthetic pathway vs. the editing pathway, by determining their effects on the rates of primer strand transfer between the polymerase and exonuclease site, the translocation rates, and the rates of dNTP binding.

3.5 Conclusion

The predominant response to the presence of a complementary ribonucleotide in phi29 DNAP complexes both before and after covalent incorporation is significant destabilization, relative to the presence of a complementary deoxyribonucleotide. This response is exhibited prior to incorporation by the substantially higher dissociation rate for rNTP from post-translocation state complexes (Table 3.1). Destabilization is also exhibited in the pre-translocation state after rNMP incorporation, as a faster rate of primer strand transfer to the exonuclease site and a faster rate of the forward translocation. When the primer terminus bears a complementary

dNMP:dNMP pair, the rate of the forward translocation is much greater than the rate of primer strand transfer to the exonuclease site (Fig. 3.6). While a complementary rNMP:dNMP pair at the primer terminus yields a modest ~ 3-fold increase in the rate of transfer to the exonuclease site, the concomitant increase in the forward translocation rate ensures that the probability of exonucleolytic editing is not directly increased by the presence of the incorrect sugar moiety. The net probability of the forward synthetic pathway vs. the editing pathway, however, can still be increased indirectly if the presence of the incorrect sugar moiety reduces the probabilities of subsequent forward reactions in the synthetic pathway. Indeed, an example of this is observed when, after incorporation and the forward translocation, the presence of a $2'$-OH group at the primer terminus also destabilizes the dNTP-bound state.

The findings in this study are in accord with a common structural cause for the destabilization of the rNTP-bound post-translocation state and of the pre-translocation state after rNMP incorporation. They provide strong functional evidence that the polymerase active site configuration surrounding the primer terminal base pair after covalent incorporation, but prior to translocation, resembles the structure of the post-translocation state ternary complex [28]. The destabilization that is a consequence of the presence of a ribonucleotide rather than a deoxyribonucleotide in both the pre-translocation and post-translocation states is at least in part due to a steric clash of the $2'$-OH group

122

with the phenyl ring of the highly conserved residue Y254; in both translocation states the destabilization is partially alleviated by the Y254V mutation. It is reasonable to predict that other B-family DNAPs, with their conserved polymerase active site structures, have similar responses to complementary ribonucleotides, and to further predict that the destabilization of both states may extend to replicative DNAPs from other families. This study thus reveals kinetic mechanisms that contribute to the initial discrimination against rNTPs over dNTPs, that underlie the inefficient editing of newly incorporated rNMPs by replicative DNAPs, and that contribute to the probability that replicative DNAPs will continue synthesis beyond a newly incorporated rNMP and stably incorporate it into the nascent strand.

Supporting Information Available: Supporting Information Text (Mathematical derivations); Figures S1-S4. This material is available free of charge via the Internet at http://pubs.acs.org.

Chapter 4

MODULATION OF DNA POLYMERASE NONCOVALENT KINETIC TRANSITIONS BY DIVALENT CATIONS

4.1 Abstract

Replicative DNA polymerases (DNAPs) require divalent metal cations for phosphodiester bond formation in the polymerase site and for hydrolytic editing in the exonuclease site. Me^{2+} ions are intimate architectural components of each active site, where they are coordinated by a conserved set of amino acids and functional groups of the reaction substrates. Therefore Me^{2+} ions can influence the noncovalent transitions that occur during each nucleotide addition cycle. Using a nanopore, transitions in individual phi29 DNAP complexes are resolved with single-nucleotide spatial precision and sub-millisecond temporal resolution. We studied Mg^{2+} and Mn^{2+}, which support catalysis, and Ca^{2+}, which supports deoxynucleoside triphosphate (dNTP) binding but not catalysis. We examined their effects on translocation, dNTP binding, and primer strand transfer between the polymerase and exonuclease sites. All three metals cause a concentration-dependent shift in the translocation equilibrium, predominantly by decreasing the forward translocation rate. Me^{2+} also promotes an increase in the backward

translocation rate that is dependent upon the primer terminal 3′-OH group. Me^{2+} modulates the translocation rates but not their response to force, suggesting that Me^{2+} does not affect the distance to the transition state of translocation. Absent Me^{2+}, the primer strand transfer pathway between the polymerase and exonuclease sites displays additional kinetic states not observed at > 1 mM Me^{2+}. Complementary dNTP binding is affected by Me^{2+} identity, with Ca^{2+} affording the highest affinity, followed by Mn^{2+}, and then Mg^{2+}. Both Ca^{2+} and Mn^{2+} substantially decrease the dNTP dissociation rate relative to Mg^{2+}, while Ca^{2+} also increases the dNTP association rate.

4.2 Introduction

DNA polymerases (DNAPs) are responsible for accurate replication of DNA genomes. Replicative DNAPs achieve this by catalyzing template-directed polymerization of deoxyribonucleoside triphosphates (dNTPs) with extremely high fidelity (with error rates of $\sim 10^{-6}$ to $\sim 10^{-8}$) [63, 54]. Phosphodiester bond formation, the chemical transformation during polymerization, is catalyzed in the polymerase active site, in the 5′ to 3′ direction. Many replicative DNAPs also catalyze a second nucleotidyl transfer reaction, the exonucleolytic cleavage of the primer strand in the 3′ to 5′ direction. This editing reaction allows for removal of incorrect nucleotides inserted during polymerization, contributing $\sim$1-2 orders of magnitude to replication fidelity [63]. Exonucleolytic editing occurs in a separate active site that is typically $\sim$ 30-40 Angstroms from the polymerase site [6, 32, 37, 58, 98, 47, 48], and it requires that $\sim$3 base pairs

125

of the nascent primer-template duplex be melted to allow the primer strand to be transferred from the polymerase site to the exonuclease site.

An essential role for two divalent metal cations (Me^{2+} ions) in the mechanism of numerous enzyme-catalyzed nucleotidyl transfer reactions has been characterized [109], in which one Me^{2+} ion (termed metal A) serves primarily to activate the nucleophile, while the other (termed metal B) mitigates the negative charge that builds in the transition state [109, 92]. For replicative DNAPs, this two Me^{2+} ion mechanism applies to both phosphodiester bond formation in the polymerase site, and to hydrolysis of the primer terminal dNMP residue in the exonuclease site [10, 91]. In accord with their roles in the chemical transformations, the Me^{2+} ions are intimate architectural components of each of the active sites. They are coordinated by a highly conserved set of acidic amino acid side chains in each site, as well as by specific functional groups of the reaction substrates. Therefore, in addition to their essential role in the chemical reactions, Me^{2+} ions may also influence the reversible noncovalent transitions that govern the fate of DNAP-DNA complexes after each covalent nucleotide addition. These transitions include i) the primer strand transfer between the polymerase and exonuclease sites, ii) the translocation fluctuations, in which the DNA substrate moves between the pre-translocation and the post-translocation states in the DNAP polymerase site, a spatial displacement of the distance of a single nucleotide, and iii) dNTP binding in the polymerase site. In contrast to the roles of the

126

Me^{2+} ions in the chemical steps of dNTP polymerization and exonucleolysis, much less is known about the effects of Me^{2+} ions on these noncovalent transitions in DNAP complexes.

The conserved architecture of the DNAP domain that contains the polymerase active site resembles a partially closed hand, comprising palm, thumb, and fingers subdomains [6, 32, 48, 57, 40]. The palm subdomain contains residues that participate in catalysis of phosphodiester bond formation, including the acidic residues involved in Me^{2+} coordination. The fingers subdomain contains residues essential for binding incoming nucleotide substrates. When dNTP binds to the polymerase active site in the post-translocation state, elements of the fingers subdomain move relative to their position in complexes lacking dNTP, closing in toward the active site to achieve a tight steric fit for the nascent base pair. In the fingers-closed dNTP-bound complex, the 3′-OH group of the DNA primer strand and a non-bridging oxygen of the α phosphate are ligands for one of the Me^{2+} ions (metal A) and non-bridging oxygens of all three phosphate groups are ligands for the metal B ion [32, 40, 79, 4].

A hypothesis for the structural basis of translocation in the B family of replicative DNAPs has been proposed [6]. The B family includes Pol-ε and Pol-δ, the enzymes that catalyze leading and lagging strand genomic replication, respectively, in eukaryotes [80, 75]. Core structure, catalytic mechanisms, and functional properties are highly conserved in this DNAP
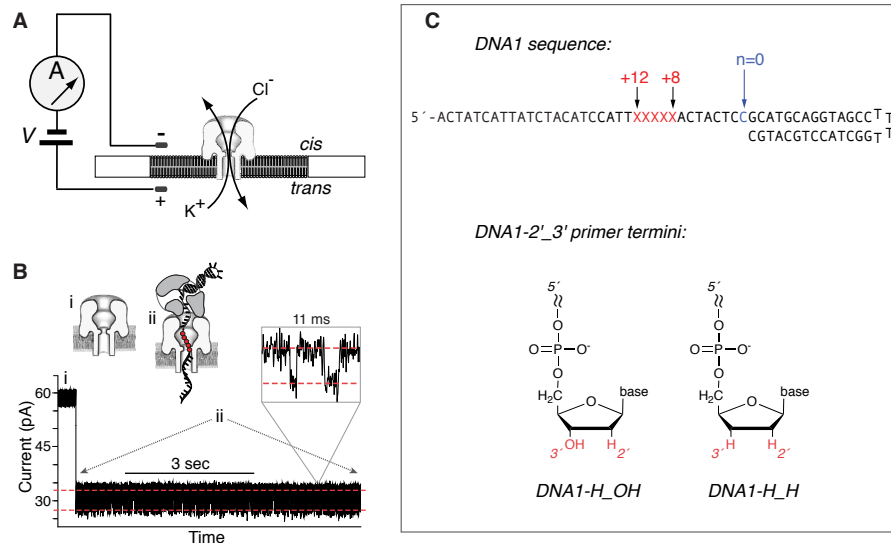
127

A

A
V
Cl⁻
cis
trans
K⁺

B

i
ii
11 ms
Current (pA)
60
45
30
3 sec
Time

C

DNA1 sequence:

+12 +8   n=0

5´-ACTATCATTATCTACATCCATTXXXXXACTACTCCGCATGCAGGTAGCC^T T
                          CGTACGTCCATCGG_T ^T

DNA1-2´_3´ primer termini:

5´
O
O=P—O⁻
O
H₂C   O   base
3´OH H 2´

DNA1-H_OH

5´
O
O=P—O⁻
O
H₂C   O   base
3´H   H 2´

DNA1-H_H

Figure 4.1: Capture of individual DNAP complexes. **A:** A single α-HL nanopore is inserted in a ~25 µm-diameter lipid bilayer separating two chambers (*cis* and *trans*) containing buffer solution. A patch clamp amplifier applies voltage across the bilayer and measures ionic current carried through the pore by K^+ and Cl^- ions. **B:** Ionic current time trace for a binary complex formed between wild type phi29 DNAP and a DNA substrate (DNA1-H_H, in panel C) captured at 180 mV applied potential in buffer containing 10 mM K-Hepes, pH 8.0, 0.3 M KCl, 1 mM EDTA, 1 mM DTT, and 11 mM $MgCl_2$. phi29 DNAP and DNA1-H_H were added to the nanopore *cis* chamber to final concentrations of 0.75 and 1 µM, respectively. Cartoons above the current trace illustrate the sequence of events: **i**) the ionic current through the open nanopore; **ii**) the current drops rapidly when a complex is captured. The enzyme is too large to enter the nanopore, and therefore the complex, with the enzyme bound at the primer–template junction of the DNA substrate, perches atop the pore. The DNA template strand of the complex is suspended through the nanopore lumen, which is just wide enough to accommodate a single strand of DNA. Individual complexes reside atop the nanopore for tens of seconds, during which the measured ionic current fluctuates on the millisecond time scale between two amplitude levels (inset). Transition between the two amplitudes corresponds to movement of the DNA substrate relative to the enzyme and the nanopore by the distance of a single nucleotide [24, 65, 66]. This spatial displacement of the DNA substrate that occurs during the translocation is detected in the time traces of ionic current by the use of a reporter group of five consecutive abasic (1′-H, 2′-H) residues in the template strand (red circles in the cartoon). The fluctuations between the two amplitudes continue until complexes dissociate or are ejected, after which another complex can be captured. **C:** DNA hairpin substrates comprise a 14-base pair duplex region and a single-stranded template region of 35 nucleotides. The reporter group of five abasic (1′, 2′-H) residues spans positions +8 to +12 (red Xs in the sequence). DNA1-H_OH bears a 2′-H, 3′-OH primer terminus and DNA1-H_H bears a 2′-H, 3′-H primer terminus.

128

family [53, 54, 6, 93, 100]. Based upon the comparison of crystal structures of the DNAP from bacteriophage phi29 in the fingers-open, post-translocation state DNAP-DNA binary complex or in the fingers-closed, post-translocation state DNAP-DNA-dNTP ternary complex [6], it was proposed that the post-translocation state ternary complex can serve as a model for the structure of a fingers-closed, pre-translocation state complex. In this view, the nascent base pair between the templating base at $n = 0$ and the incoming complementary dNTP in the post-translocation state ternary complex occupies the site that would be occupied by the terminal base pair of the primer-template duplex in the pre-translocation state complex. The structure of the binary complex in the fingers-open, post-translocation state indicates that the pre-translocation state in the fingers-open conformation is sterically precluded, and hence fingers opening was proposed to accompany the forward translocation [6]. Little is known about whether Me^{2+} ions that interact with complexes in the polymerase active site in the pre-translocation or post-translocation states exert any influence on the translocation fluctuations. For the X-family mammalian repair DNAP, pol β, there is evidence that Me^{2+} can stabilize the pre-translocation state in DNAP-DNA binary complexes [59].

We have developed a single-molecule approach using a nanoscale pore (Fig. 4.1A and B) that permits quantification of the rates of translocation fluctuations, rates of the primer strand transfer between the polymerase and exonuclease sites, and rates of dNTP binding, in individual DNAP-DNA

complexes [24, 66, 67, 68]. We used the B-family replicative phi29 DNAP, which serves as an excellent model system for leading strand DNA synthesis catalyzed in more complex B family replisomes. It catalyzes highly processive DNA synthesis while remaining tightly associated with its DNA substrate and promoting downstream strand displacement during replication [6, 73, 12, 13, 85], obviating the need for accessory proteins such as sliding clamps or helicases. Primer strand transfer between the phi29 DNAP polymerase and exonuclease sites is an intramolecular process [29].

When a complex formed between phi29 DNAP and a primer template DNA substrate (Fig. 4.1 C) is captured atop an α-hemolysin (α-HL) nanopore in an applied electric field (Fig. 4.1 A and B), it undergoes iterative fluctuations across the translocation step. The forward and reverse fluctuations across the translocation step are recorded as discrete transitions between two amplitudes in time traces of ionic current (Fig. 4.1 B, ii, inset; Fig. 4.2 A). Each individual captured complex resides atop the nanopore and undergoes the amplitude fluctuations for several seconds, and in each experiment, hundreds of DNAP-DNA complexes can be captured and examined one-by-one, in series. Dwell time samples of each amplitude level are extracted from the ionic current time traces and used to quantify transitions with sub-millisecond temporal resolution and single nucleotide spatial precision.

Complexes formed between phi29 DNAP and DNA1-H_H captured in the presence of 10 mM Mg^{2+} fluctuate rapidly between the two ionic current
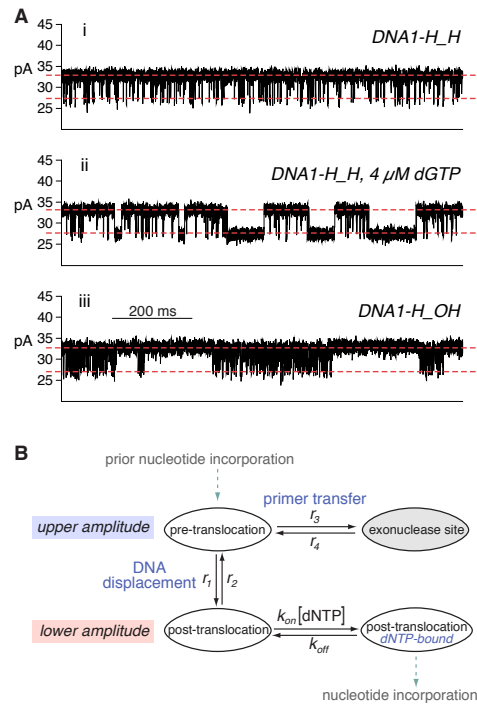
Figure 4.2: Kinetic transitions resolved in individual phi29 DNAP complexes captured atop a nanoscale pore. A: Ionic current time traces for complexes captured at 180 mV in the presence 10 mM Mg^{2+}. Complexes were formed between the D12A/D66A mutant of phi29 DNAP and **i**) DNA1-H_OH; or **ii**) and **iii**) DNA1-H_H. In iii) 4 μM dGTP was present in the *cis* chamber. The two red dashed lines indicate the two amplitude levels between which the measured current fluctuates; at 180 mV, the upper amplitude is centered at ~32 pA, and the lower amplitude is centered at ~26 pA at 180 mV. **B:** Four-state model for the reversible, noncovalent transitions that govern the fate of replicative DNAP complexes following each covalent nucleotide addition. The model shows the kinetic relationships among the steps of translocation, primer strand transfer between the polymerase and exonuclease active sites, and dNTP binding. Transfer of the primer strand from the polymerase to the exonuclease site precedes the forward translocation (26), and the forward translocation precedes dNTP binding; the translocation is rectified but not driven by dNTP binding, and dNTP has no influence on the rates across the translocation step (25). The kinetic model comprises six transition rates: the rates of translocation (r_1, r_2), the rates of primer stand transfer between the pre-translocation state polymerase site and the exonuclease site (r_3, r_4), and the rates of dNTP binding to post-translocation state complexes (k_{on}[dNTP] and k_{off}). The mathematical framework based on the model allows these rates to be determined from experimental measurements [66, 67, 68]. Each of the three unshaded ellipses corresponds to a state (pre-translocation state, post-translocation state, or dNTP-bound post-translocation state) in which the primer strand, in duplex with the template strand, resides in the polymerase site; the ellipse that is shaded grey corresponds the state in which the primer terminus resides in the exonuclease site.

amplitudes (Figs. 4.1 B and 4.2 A, i). Addition of dNTP complementary to the templating base at n=0 (dGTP) stabilizes the lower amplitude post-translocation state (Fig. 4.2 A, ii) (23, 25). In contrast to complexes formed with DNA1-H_H (Fig. 4.2A, i), complexes formed with DNA1-H_OH display an additional kinetic state at the upper amplitude. The presence of this state can be directly observed in the time traces as pauses of tens to hundreds of milliseconds that punctuate intervals of rapid fluctuation between the two amplitudes (Fig. 4.2 A, iii). We showed that the periods of rapid fluctuation are due to transitions between the pre-translocation and post-translocation states and that the pauses in the upper amplitude arise when the primer strand is transferred from the polymerase active site to the exonuclease site [68]. Thus, the upper amplitude comprises the pre-translocation state in the polymerase site and the state in which the primer strand occupies the exonuclease site; the lower amplitude comprises the post-translocation state in the polymerase site [24, 66, 67, 68].

When an individual DNAP-DNA complex is held atop the nanopore, the applied voltage exerts a force on the complex along the direction of the translocation. This applied force shifts the translocation equilibrium toward the upper amplitude, pre-translocation state [24, 66]; it impedes the rate of the forward translocation and increases the rate of the reverse translocation [24, 66], but it does not affect the rates of dNTP binding [67] or the rates of primer strand transfer between the polymerase and exonuclease sites [68]. Based on

translocation fluctuations measured under varying opposing force loads and in the presence of varying nucleotide concentrations, we established the kinetic relationship of the translocation to the step of nucleotide binding and to the step of primer strand transfer between the polymerase and exonuclease sites (Fig. 4.2 B). Specifically, we showed that the forward translocation precedes dNTP binding; the translocation fluctuation is rectified but not driven by dNTP binding, and dNTP has no influence on the rates of the translocation fluctuations (Fig. 4.2 B) [67]. Further, we established that transfer of the primer strand from the polymerase to the exonuclease site occurs prior to the forward translocation fluctuation; the pre-translocation state is therefore the branchpoint between the DNA synthesis and editing pathways (Fig. 4.2 B) [68]. We developed mathematical methods to determine the forward (r_1) and reverse (r_2) rates of translocation fluctuations, the forward (r_3) and reverse (r_4) rates of primer strand transfer between the polymerase and exonuclease sites, and the rates of dNTP or rNTP association (k_{on}) and dissociation (k_{off}) in individual DNAP complexes (Fig. 4.2 B) [66, 67, 68, 25].

In prior studies, we applied these capabilities to examine the kinetic mechanisms by which mutations of highly conserved DNAP residues, alterations in DNA substrate sequence or structure, or alterations in nucleotide substrate structure exert their influence on DNAP function [24, 25, 26, 66, 68]. The majority of these studies were conducted using complexes captured in the presence of a high concentration of Mg^{2+} (10 mM). In the

133

current study, we have examined the effects of Me^{2+} ions on noncovalent transitions that are essential to DNAP function. We determined the concentration-dependent effects of Mg^{2+}, Mn^{2+}, and Ca^{2+} on the translocation fluctuations, on primer strand transfer between the polymerase and exonuclease sites, and on dNTP binding, in individual phi29 DNAP-DNA complexes. Mg^{2+} and Mn^{2+} are metals that support catalysis in both the polymerase and exonuclease active sites, while Ca^{2+} supports nucleotide binding but not catalysis in either the polymerase or the exonuclease active sites [34].

4.3 Experimental Methods

DNA and Enzymes. DNA substrates were synthesized at Stanford Protein and Nucleic Acid Facility and purified by denaturing PAGE. DNA hairpins were annealed by heating at 90 °C for 4 min followed by snap cooling in ice water. Wild type phi29 DNAP was obtained from Enzymatics (Beverly, MA). The D12A/D66A mutant was obtained from XPol Biotech (Madrid, Spain).

Nanopore Methods. Nanopore experiments were conducted as described [23, 25, 35-38] single α-HL nanopore is inserted in a ~25 µm-diameter lipid bilayer that separates two chambers (*cis* and *trans*) containing buffer solution (10 mM K-Hepes, pH 8.0, 0.3 M KCl, and 1 mM DTT). DNA and phi29 DNAP were added to the *cis* chamber to final concentrations of 1 and 0.75 µM, respectively. $MgCl_2$, $CaCl_2$, $MnCl_2$ or EDTA were added to the *cis* well as

indicated in the text and figure legends. Ionic current was measured with an integrating patch clamp amplifier (Axopatch 200B, Molecular Devices) in voltage clamp mode. Data were sampled using an analog-to-digital converter (Digidata 1440A, Molecular Devices) at 100 kHz in whole- cell configuration and filtered at 5 kHz using a low pass Bessel filter.

Detection of the translocation in captured DNAP-DNA complexes. The displacement of the DNA relative to the enzyme that occurs during translocation is detected with single nucleotide spatial precision and sub-millisecond temporal resolution by the use of a reporter group in the template strand of the DNA substrate, and has been described in detail [24, 25, 26 65, 66, 67, 68, 20]. Briefly, the reporter comprises five consecutive abasic (1′, 2′-H) residues (shown as red Xs in the sequence in Fig. 4.1 B and as solid red circles in the cartoon in Fig. 4.2 B). When the reporter resides within the pore lumen, the abasic residues allow more ions to flow through the channel than a strand composed of normal DNA residues alone. The extent to which the abasic reporter augments the amplitude of captured complexes depends upon its position relative to the limiting aperture of the nanopore lumen. Thus, displacement of the abasic reporter group in the nanopore lumen is manifested as a change in measured ionic current [24, 65].

Analysis of ionic current time traces. The extraction of the dwell time samples of the upper and lower amplitudes from recorded time traces of ionic current, and the calculation of kinetic transition rates, has been described in detail [68,

Estimating free [dNTP] in bulk phase. Due to the high affinity of phi29 DNAP-DNA complexes for dNTP in Mn^{2+} and Ca^{2+} it was necessary to measure complementary dNTP binding by adding very low concentrations of dGTP to the bulk phase in the *cis* well (Fig. 4.10). In bulk phase, dGTP may bind onto complexes to form ternary complexes. As a result, not all added dGTP molecules are free in bulk. When the added dGTP concentration is lower than the enzyme concentration, the concentration of free dGTP in bulk may be significantly different from that of added dGTP. Therefore, we need to calculate the free dGTP concentration.

Let

p = probability of lower amplitude (post-translocation state, with or without dNTP bound)

$[dGTP]_0$ = concentration of added dGTP in bulk

$[DNAP \cdot DNA]_0$ = concentration of added enzyme-DNA complexes in bulk

$[dGTP]_{Free}$ = concentration of free dGTP in bulk

$K_d^{(bulk)}$ = apparent binding affinity of dGTP onto enzyme-DNA complexes in bulk

K_d = binding affinity of dGTP onto an enzyme-DNA complex that is captured atop the pore and is in the post-translocation configuration

Note that $K_d^{(bulk)}$ may be different form K_d since in bulk, not all enzyme

136

molecules are necessarily at the primer-template junction, and for complexes with enzyme at the primer-template junction, not all of them are in the post-translocation state (the state competent to bind dNTP) at all time. So, in our model, we allow $K_d^{(bulk)}$ to be different from K_d.

In experiments, $[DNAP \cdot DNA]_0$ is fixed at 0.75μM and $[dGTP]_0$ is a control variable. As a function of $[dGTP]_0$ and $K_d^{(bulk)}$, the concentration of free dGTP in bulk has the expression:

$$[dGTP]_{Free}\left([dGTP]_0, K_d^{(bulk)}\right) = \frac{[dGTP]_0 - [DNAP \cdot DNA]_0 - K_d^{(bulk)}}{2}$$
$$+ \frac{\sqrt{\left([dGTP]_0 + [DNAP \cdot DNA]_0 + K_d^{(bulk)}\right)^2 - 4[dGTP]_0 \cdot [DNAP \cdot DNA]_0}}{2}$$

Previously [23] we established that the normalized $p/(1-p)$ is related to the free dGTP concentration by

$$\text{normalized } \frac{p}{1-p} = 1 + \frac{[dGTP]_{Free}}{K_d}$$

It follows that log(normalized $p/(1-p)$ − 1) is a linear function of log($[dGTP]_{Free}$) with slope 1.

$$\log\left(\text{normalized } \frac{p}{1-p} - 1\right) = \log\left([dGTP]_{Free}\left([dGTP]_0, K_d^{(bulk)}\right)\right) - \log(K_d)$$

In dGTP titration experiments, values of normalized $p/(1-p)$ are observed for a sequence of $[dGTP]_0$ values. For a given value of $K_d^{(bulk)}$, a sequence of $[dGTP]_{Free}$ values are calculated from the $[dGTP]_0$ values. Data points of

$$\log\left([dGTP]_{\text{Free}}\right) \text{ vs } \log\left(\text{normalized } \frac{p}{1-p} - 1\right)$$ are fitted to a straight line with

slope 1. The deviation of data points from the fitting line is used to measure

how well data points follow a straight line with slope 1. The value of $K_d^{(\text{bulk})}$ is

determined by minimizing this deviation.

Once the value of $K_d^{(\text{bulk})}$ is calculated, we calculate the free dGTP

concentration $[dGTP]_{\text{Free}}$ corresponding to each added dGTP concentration

$[dGTP]_0$.

4.4 Results

When phi29 DNAP complexes are captured atop the nanopore in the

absence of added Me^{2+}, the measured ionic current fluctuates between the

two amplitudes that are characteristic of the fluctuations across the

translocation step observed in the presence of Me^{2+} [68]. Neither the forward

or reverse translocation displacement requires Me^{2+}. However, in the absence

of Me^{2+}, the probability of occupancy of the lower amplitude, post-

translocation state (p) is higher than when complexes are captured in the

presence of Me^{2+} (Fig. 4.3). Addition of Me^{2+} to the bulk phase elicits a

decrease in p for complexes formed with both DNA1-H_OH, which bears a

natural $2'$-H, $3'$-OH primer terminus, or DNA1-H_H, which bears a $2'$-H, $3'$-H

primer terminus (Fig. 4.1 C). This is illustrated in amplitude histograms for

complexes formed between the D12A/D66A mutant of phi29 DNAP and

DNA1-H_OH (Fig. 3A) or DNA1-H_H (Fig. 4.3 B), captured in either the
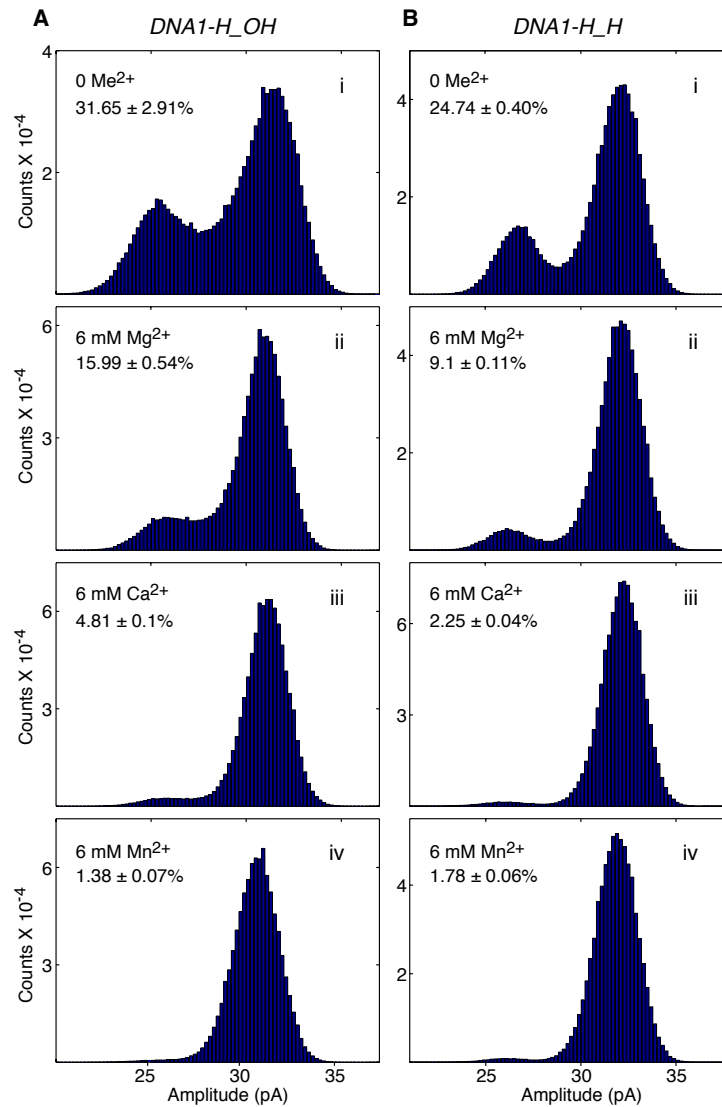
Figure 4.3: Me²⁺ decreases the probability of the post-translocation state. Amplitude histograms for complexes formed between the D12A/D66A mutant of phi29 DNAP and (**A**) DNA1-H_OH or (**B**) DNA1-H_H, captured in the presence of **i**) 0 Me²⁺, **ii**) 6 mM Mg²⁺, **iii**) 6 mM Ca²⁺, or **iv**) 6 mM Mn²⁺.

absence of Me²⁺ (Fig. 4.3 A, i, and 4.3 B, i) or in the presence of 6 mM Mg²⁺

(Fig. 4.3 A, ii, and 4.3 B, ii), 6 mM Ca²⁺ (Fig. 4.3 A, iii, and 4.3 B, iii), or 6 mM

Mn²⁺ (Fig. 4.3 A, iv, and 4.3 B, iv). The D12A/D66A mutant lacks two of the

ligands for the catalytic Me²⁺ ions in the exonuclease active site and thus has

negligible exonucleolytic activity [9, 42]. This permits experiments to be conducted with DNA1-H_OH in the presence of Mg^{2+} or Mn^{2+}, conditions under which the $2'$-H, $3'$-OH terminated primer strand would be degraded in the bulk phase by the wild-type enzyme [65, 68]. The extent to which p was decreased for the D12A/D66A complexes varied with the identity of the metal species; for both DNA1-H_OH and DNA1-H_H, Mn^{2+} caused the largest decrease in p, followed by Ca^{2+} and then Mg^{2+} (Fig. 4.3).

We examined the effects of Me^{2+} concentration on the probability of the post-translocation state in phi29 DNAP complexes by performing titration experiments spanning five orders of magnitude, from 0.2 μM through 10000 μM Me^{2+}. For DNA1-H_OH, we used the wild type enzyme in Ca^{2+} titration experiments, and the D12A/D66A enzyme in titration experiments with Mg^{2+}, Ca^{2+}, and Mn^{2+} (Fig. 4.4 A). Because the $2'$-H, $3'$-H primer terminus of DNA1-H_H inhibits the onset of phi29 DNAP-catalyzed exonucleolytic digestion of the primer strand relative to substrates bearing $2'$-H, $3'$-OH termini [65], it affords protection for DNA in experiments conducted with the wild type enzyme in the presence of Mg^{2+} or Mn^{2+}. This protection permitted us to examine the effects of all three metals (Mg^{2+}, Ca^{2+}, and Mn^{2+}) on complexes formed between DNA1-H_H with the wild type enzyme and those with D12A/D66A enzyme in the titration experiments (Fig. 4.4 B).

For complexes formed with DNA1-H_OH in the absence of Me^{2+}, but not those formed with DNA1-H_H in the absence of Me^{2+}, the identity of the
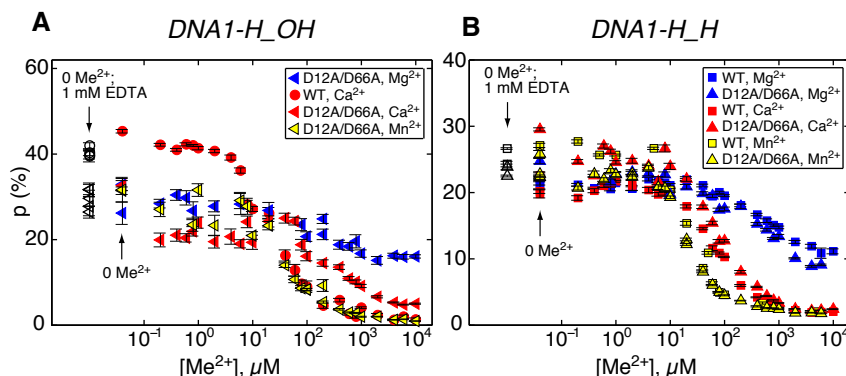
140

Figure 4.4: Concentration-dependent effects of Me^{2+} on the probability of the post-translocation state. **A:** Concentration-dependent effects on p of Mg^{2+} (blue symbols), Ca^{2+} (red symbols), or Mn^{2+} (yellow symbols) for complexes formed between DNA1-H_OH and the wild type enzyme (circles) or the D12A/D66A enzyme (sideways triangles). **B:** Concentration-dependent effects on p of Mg^{2+} (blue symbols), Ca^{2+} (red symbols), or Mn^{2+} (yellow symbols) for complexes formed between DNA1-H_H and the wild type enzyme (squares) or the D12A/D66A enzyme (triangles). For each plotted data point, dwell time samples at each experimental condition are extracted from 10-20 time traces of captured complexes. Each typical time trace comprises a segment of ~ 10 seconds, recorded at 100 kHz, and thus consisting of 1 million data points. Because there is no zero value on the log scale horizontal axis of the plots in (*A*) and (*B*), values for p determined in the absence of added Me^{2+} (from each individual titration experiment) are placed on the plots at the position for 0.04 μM Me^{2+}, and are indicated by an arrow and label (0 Me^{2+}). Values for equivalent complexes, captured in separate control experiments in the absence of Me^{2+} and in the presence of 1 mM EDTA, are placed on the plots at the position for 0.01 μM Me^{2+}. These values are indicated in (*A*) for complexes of DNA1-H_OH with the wild type enzyme (open black circles) or with the D12A/D66A enzyme (open black sideways triangles), and in (*B*) for complexes with DNA1-H_H with the wild type enzyme (open black squares) or the D12A/D66A enzyme (open black triangles), by an arrow and label (0 Me^{2+}; 1 mM EDTA). Within the range of experiment to experiment variation, the measured p values were unaffected by the presence or absence of EDTA. This indicates that Me^{2+} titration experiments, performed in the absence of EDTA, are not affected by residual or contaminating Me^{2+} in the system. Complexes in the Me^{2+} titration experiments were captured at 180 mV applied potential.

enzyme causes a difference in p; for wt phi29 DNAP at 0 Me^{2+}, $p \approx 45\%$ (Fig. 4.4 A) while for the D12A/D66A mutant at 0 Me^{2+} $p \approx 30\%$ (Fig. 4.3 A, i and 4.4 B). For all complexes examined, formed with either the wild type or D12A/

D66A enzyme, and either the DNA1-H_OH or DNA1-H_H substrate, Me^{2+} causes a concentration-dependent decrease in p. For most of the complexes, the decrease in p occurs primarily within the concentration range of ~ 10 μM to 1000 μM Me^{2+} (Fig. 4.4); the exception is the response of p to Ca^{2+} for the complexes of the wild type enzyme with DNA1-H_OH (Fig. 4.4 A). For this complex, the decrease in p begins at a lower concentration (between ~ 6 and 8 μM) than it does for D12A/D66A complexes formed with DNA1-H_OH (~ 10 μM). Further, for the wild type complexes formed with DNA1-H_OH, p drops precipitously in response to Ca^{2+}, reaching a lower level than for the D12A/ D66A enzyme at concentrations above $\sim 60\text{-}80$ μM. At $[Ca^{2+}] \geq \sim 1000$ μM, p for the wild type enzyme resolves to $\sim 1\text{-}2\%$ and to $\sim 4\text{-}5\%$ for the D12A/ D66A enzyme.

Me^{2+}-concentration dependent resolution of kinetic complexity in the primer strand transfer pathway. We sought to determine the kinetic mechanisms that contribute to the Me^{2+}-dependent decrease in probability of post-translocation state occupancy. While the net effect of Me^{2+} for each DNAP-DNA complex is a decrease in p relative to the same complex in the absence of Me^{2+}, any subset of the noncovalent transitions that govern the probability of the post-translocation state could be affected by Me^{2+} (Fig. 4.2 B), and the magnitude and Me^{2+}-concentration dependence of effects on each transition could vary depending upon the identity of the partners in the DNAP-DNA complex and upon the identity of the metal. We first examined the survival probability vs.

142

dwell time of upper amplitude dwell time samples and that of lower amplitude

dwell time samples, extracted from ionic current traces for complexes

captured in the Me^{2+} titration experiments (Figs. 4.5 and 4.6). The survival

probability as a function of time t is defined as the probability that the dwell

time is larger than t. The observed survival probability is calculated as the

fraction of dwell time samples larger than t.
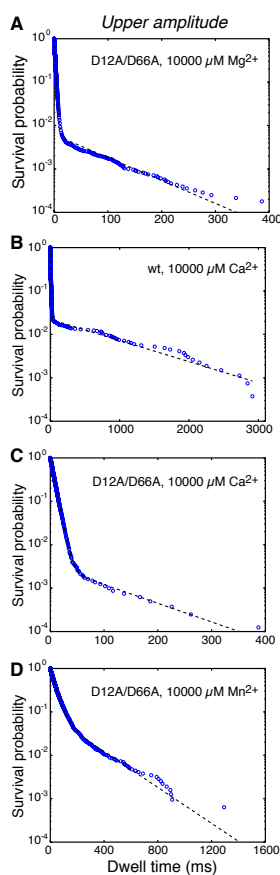
Figure 4.5: Survival probability of the upper amplitude at 10000 μM Me^{2+}. Survival probability vs. dwell time plots for dwell time samples extracted from from the upper amplitude of ionic current traces, for complexes formed between DNA1-H_OH and (**A**, **C-D**) the D12A/D66A enzyme or (**B**) the wild type enzyme. Complexes were captured in the presence of (A) 10000 μM Mg^{2+}, (B and C) 10000 μM Ca^{2+}, or (D) 10000 μM Mn^{2+}. Blue circles represent the data points, and the dashed black line represents MLE fitting to a model of two exponential modes.

In the presence of 10 mM Mg^{2+}, the upper amplitude survival probability for binary complexes formed between the D12A/D66A enzyme with DNA1-H_OH is well fit by a model of two exponential modes [68] (Fig. 4.5 A). We have shown that these two kinetic states in the upper amplitude correspond to i) stays of the complex in the pre-translocation state in the polymerase site, and ii) stays of the complex in the state in which the primer strand occupies the exonuclease site [68]. Upper amplitude survival probability for complexes captured in 10 mM Ca^{2+}, formed between DNA1-H_OH and either the wild type enzyme (Fig. 4.5 B) or the D12A/D66A enzyme (Fig. 4.5 C), are also well fit by a model of two exponential modes, as are complexes of the D12A/D66A enzyme with DNA1-H_OH captured in 10 mM Mn^{2+} (Fig. 4.5 D).

However, in the absence of Me^{2+} (Fig. 4.6 A, i), and across a wide range of $[Me^{2+}]$ up to almost 1000 μM the distribution of upper amplitude dwell time for complexes formed with the DNA1-H_OH substrate deviates from a distribution of two exponential modes. This is illustrated in Fig. 4.6 A, i-iv, for complexes captured in 0, 1, 10, and 100 μM Mg^{2+}. The upper amplitude survival probability contains more than two exponential modes, suggesting the possibility that at low $[Me^{2+}]$, residence in the exonuclease site may be described by multiple kinetic states, rather than a single kinetic state. As $[Mg^{2+}]$ is increased into the mM range (Fig. 4.6 A, v-vi, and Fig. 4.5 A) survival probability of the upper amplitude converges to a model of two exponential modes.

144
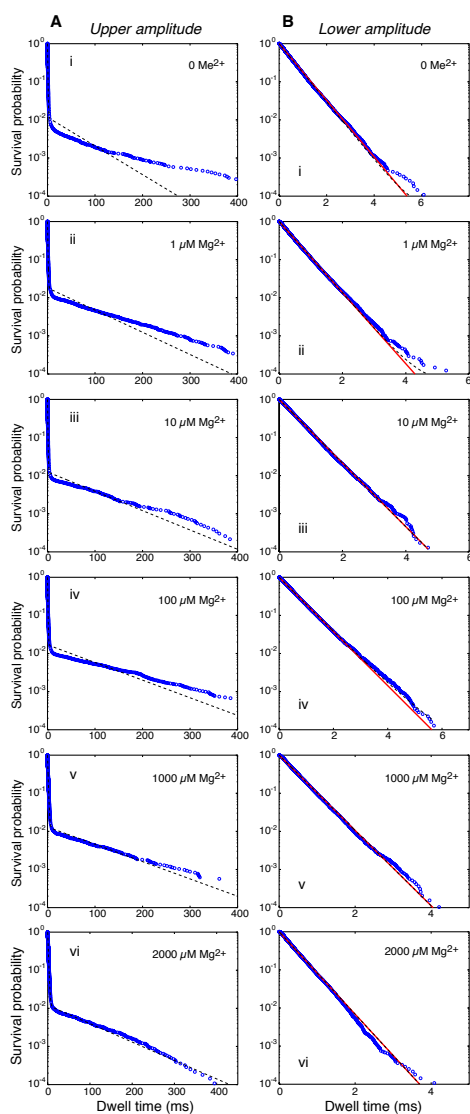
Figure 4.6: Effects of Me²⁺ on the survival probability of each of the two amplitude levels. Survival probability vs. dwell time plots based on dwell time samples extracted from from (**A**) the upper amplitude, or (**B**) the lower amplitude of ionic current traces, for complexes formed between D12A/D66A phi29 DNAP and DNA1-H_OH. Complexes were captured in the presence of **i**) 0 Me²⁺, **ii**) 1 μM Mg²⁺, **iii**) 10 μM Mg²⁺, **iv**) 100 μM Mg²⁺, **v**) 1000 μM Mg²⁺, **vi**) 2000 μM Mg²⁺. In (A) and (B), blue circles represent the data points; the dashed black line represents MLE (maximum likelihood estimate) fitting to a model of two exponential modes; in (B) the solid red line represents MLE fitting to a single exponential.

Qualitatively similar kinetic complexity (deviation from two exponential modes)

at Me²⁺ concentrations below ~1000 μM was observed for the upper

amplitude for all of the complexes examined in the Mg^{2+}, Ca^{2+}, or Mn^{2+} titration experiments. In addition, as we have previously reported, when complexes formed between either the wild type or D12A/D66A enzyme and the DNA1-H_H substrate are captured in the presence of high $[Me^{2+}]$, plots of survival probability vs. dwell time of the upper amplitude are fit by a single exponential decay function. This decay rate corresponds to the inverse of the forward translocation rate [66, 68]. Thus, for the DNA1-H_H complexes at high $[Me^{2+}]$, a population of dwell time samples that corresponds to stays of the primer strand in the exonuclease site cannot be resolved. This, combined with the kinetic complexity in the upper amplitude observed in the presence of 0 or sub-millimolar Me^{2+} for complexes formed with either DNA1-H_OH or DNA1-H_H, precludes our ability to reliably quantify the effects of Me^{2+} on the rates of primer strand transfer between the polymerase and exonuclease site (r_3 and r_4; Fig. 4.2 B).

Despite the kinetic complexity in the upper amplitude, the effects of Me^{2+} ions on the forward and reverse rates of translocation can be reliably determined. First, for each of the complexes examined, in the absence of Me^{2+}, and across the entire range of $[Me^{2+}]$, survival probability of the lower amplitude is well fit by a single exponential decay function (Fig. 4.6 B) [66, 68]; thus the reverse translocation rate (r_2) can be calculated from the data. Second, despite the complexity of the upper amplitude dwell time distribution observed at 0 Me^{2+} and at Me^{2+} concentrations below $\sim$1000 μM (Fig. 4.6 A, i-iv), r_1 can

also be determined across the [Me²⁺] titration series. This is illustrated using a hypothetical example for the kinetic complexity of the upper amplitude dwell time cluster (Fig. 4.7), in which the upper amplitude comprises four kinetic states: A, B, C and D, where state A is the pre-translocation state. The
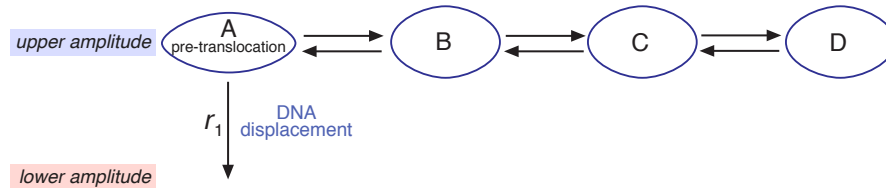
Figure 4.7: Hypothetical kinetic diagram governing the escape from the upper amplitude. At 0 Me²⁺ and low [Me²⁺], the upper amplitude dwell time shows more than two exponential modes (Fig. 6), which indicates that the primer strand transfer pathway between the polymerase and exonuclease sites contains multiple kinetic states, and is no longer described by a pair of transition rates (r_3 and r_4 as in Fig. 2B). Even in that case, the forward translocation rate r_1 can still be reliably calculated from measured dwell time samples (see Results section).

forward translocation rate (r_1) is the rate of escaping from state A to the lower amplitude, and can be accurately determined from the measured survival probability. Let S(t) be the upper amplitude survival probability as a function of dwell time t. Mathematically, no matter how kinetically complex the upper amplitude is, it holds that

$$r_1 = - \left. \frac{dS(t)}{dt} \right|_{t=0}$$

That is, the forward translocation rate r_1 is the initial decay rate of the survival probability.

The influence of Me²⁺ on phi29 DNAP translocation rates. Me²⁺ ions have notable effects on both the forward rate (r_1) and the reverse rate (r_2) of the

147

translocation fluctuations (Fig. 4.8). For all complexes examined, formed between either the wild type or D12A/D66A enzymes, with either the DNA1-H_OH or DNA1-H_H substrates, Me^{2+} causes a concentration-dependent decrease in r_1. The $[Me^{2+}]$-dependent decreases in r_1 (Fig. 4.8 A and C) largely mirror the $[Me^{2+}]$-dependent decreases in p (Fig. 4.4 C and D); the decrease in the forward translocation rate (r_1) makes the dominant contribution to the decrease in the probability of the post-translocation state
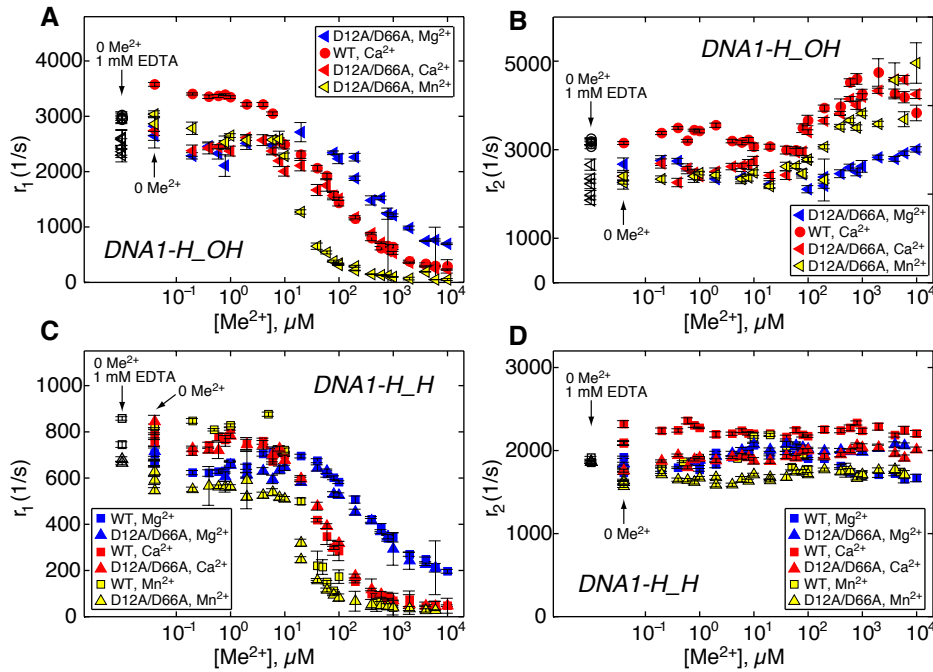
Figure 4.8: Concentration-dependent effects of Me^{2+} on phi29 DNAP translocation rates. Plots of (**A**, **C**) r_1 vs. $[Me^{2+}]$ and (**B**, **D**) r_2 vs. $[Me^{2+}]$ for complexes formed with the DNA1-H_OH substrate (**A**, **B**) or with the DNA1-H_H substrate (**C**, **D**). The shapes and colors of symbols for the complexes in (A) and (B) are displayed in the inset box and are described in the legend for Fig. 4*A*; the shapes and colors of symbols for the complexes in (C) and (D) are displayed in the inset box and are described in the legend for Fig. 4*B*. Within the range of experiment to experiment variation, the values for r_1 and r_2 for each of the DNAP-DNA complexes captured in the absence of added Me^{2+} (indicated by an arrow and label: 0 Me^{2+}) do not vary from those of the equivalent complexes, captured in control experiments in the absence of Me^{2+} and in the presence of 1 mM EDTA (indicated by an arrow and label: 0 Me^{2+}; 1 mM EDTA).

(p). As with the decrease in p, the effects of Me^{2+} on r_1 occur primarily within the concentration regime of ∼ 10 μM to ∼ 1000 μM Me^{2+} (Fig. 4.8 A and C). Also in accord with the decrease in p, the extent to which r_1 was decreased for the D12A/D66A complexes varied with the identity of the metal species; for both DNA1-H_OH and DNA1-H_H, Mn^{2+} elicited the greatest decrease in r_1, followed by Ca^{2+} and then Mg^{2+} (Fig. 4.8 A and C). In complexes formed with DNA1-H_OH, the higher value of p observed in the absence of Me^{2+} for the wild type enzyme relative to the D12A/D66A enzyme (Fig. 4.4) is also accounted for, at least in part, by a difference in r_1 for the two enzymes in the absence of Me^{2+} (Fig. 4.8 A). For wild type complexes formed with DNA1-H_OH at 0 Me^{2+}, $r_1 \approx 3400$/sec, while for D12A/D66A complexes formed with DNA1-H_OH at 0 Me^{2+} $r_1 \approx 2700$/sec. With DNA1-H_OH, the effects of Me^{2+} on complexes formed with both enzymes can only be compared for Ca^{2+}; we note that the Ca^{2+} concentration at which r_1 begins to decrease differs for the two enzymes, with the decrease in r_1 beginning at ∼ 0.2 μM Ca^{2+} for the wild type enzyme and at ∼ 1 μM for the D12A/D66A enzyme. At Ca^{2+} concentrations above ∼ 1 μM, the values of r_1 for the two enzymes converge (Fig. 4.8 A).

For complexes formed with either the wild type or D12A/D66A enzymes, Me^{2+} causes a concentration-dependent increase in r_2 that is dependent upon the presence of the primer terminal 3′-OH group: the increase is observed for complexes formed with DNA1-H-OH (Fig. 4.8 B) but not for complexes

formed with DNA1-H_H (Fig. 4.8 D). The increase in r_2 is much more modest than the decrease in r_1. The Me^{2+}-dependent increase in r_2 becomes detectable at ~100 μM Me^{2+}, an order of magnitude higher than the concentration required to observe the decrease in r_1. In complexes formed with the D12A/D66A enzyme, r_2 continues to increase as a function of either Mg^{2+} or Mn^{2+} up to at least 10000 μM, the highest concentration tested. In contrast, the ability of Ca^{2+} to cause an increase in r_2, in complexes formed with either the wild type or D12A/D66A enzymes, appeared to saturate above ~ 2000 μM (Fig. 4.8 B). Like the decrease in r_1, the extent to which r_2 was increased for the D12A/D66A complexes varied with the identity of the metal species, but the rank order of effectiveness among the three metal ions at eliciting the increase in r_2 differs from the effect on r_1; Ca^{2+} elicited the greatest increase in r_2, followed by Mn^{2+}, and then by Mg^{2+}, which had modest effects (Fig. 4.8 B).

The relationship among the effects of Me^{2+}, the identity of the primer terminus, and the translocation energy landscape. The influence of Me^{2+} ions on the magnitude of both r_1 and r_2 uncovered in the Me^{2+} titration experiments (Fig. 4.8) prompted us to examine the effect of Me^{2+} ions on the energy landscape across the translocation step (Fig. 4.9). The Me^{2+} titration experiments were conducted by capturing phi29 DNAP complexes at a single applied voltage (180 mV). When complexes are held atop the nanopore, the force applied by the voltage impedes the rate of the forward translocation (r_1)
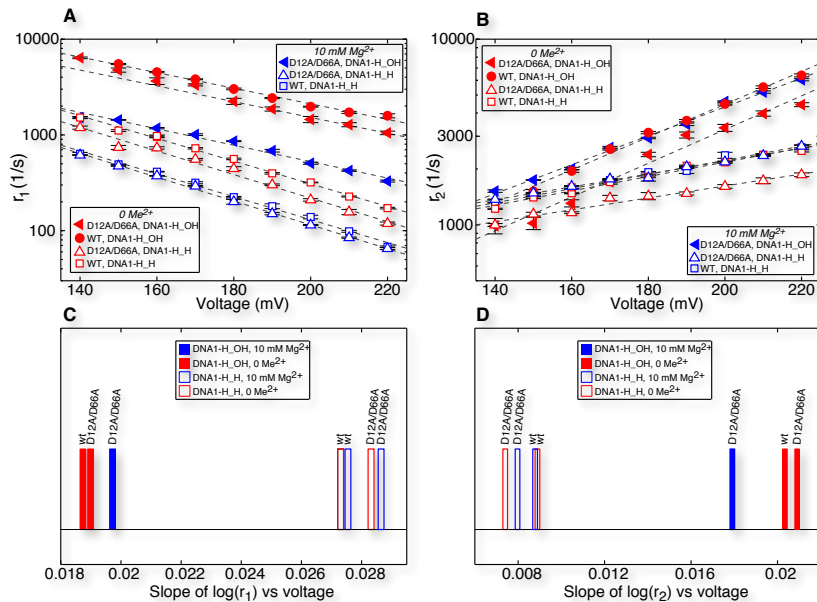
Figure 4.9: Effects of Me^{2+} on the response of the translocation rates to applied force. Plots of (**A**) log(r_1) vs. voltage and (**B**) log(r_2) vs. voltage for complexes formed between DNA1-H_OH (filled symbols) or DNA1-H_H (unfilled symbols) and either wild type phi29 DNAP or D12A/D66A phi29 DNAP. Complexes were captured in the presence of either 0 Me^{2+}, 1 mM EDTA (red symbols) or 11 mM Mg^{2+}, 1 mM EDTA (blue symbols). Rates were determined from dwell time samples extracted from ionic current traces and a three-state kinetic model (26)consisting of transitions r_1, r_2, r_3, and r_4 in the model diagram in Fig. 2*B*. Errors bars indicate the standard error. The fitting lines in (A) and (B) were generated by linear least squares fitting. The absolute values of the slopes of the fitting lines for each of the complexes are plotted in (**C**) for log(r_1) and in (**D**) for log(r_2).

and increases the rate of the reverse translocation (r_2). Plots of log(r_1) vs. voltage and log(r_2) vs. voltage both fit to straight lines, indicating that the force is applied along the direction of the translocation [24]. The Me^{2+} ions could influence both the rate at a given voltage and the dependence of the rates on voltage. The slope of log(r_1) vs. voltage is negative and proportional to the distance between the pre-translocation state and the transition state for the translocation step; the slope of log(r_2) vs voltage is positive and proportional to the distance between the transition state and the post-translocation state

151

(66).

We showed previously that the primer terminal 3′-OH group is a determinant in the energy landscape of the translocation [68]. When complexes are captured at 180 mV in the presence of 10 mM Mg^{2+}, r_1 is ~3.8 fold faster in complexes with DNA1-H_OH than it is in complexes of DNA1-H_H. The 3′-OH group causes a decrease in the slope of $\log(r_1)$ vs. voltage and an increase in the slope of $\log(r_2)$ vs. voltage, indicating that in 10 mM Mg^{2+}, along the coordinate of the translocation displacement, the transition state is closer to the pre-translocation state for complexes formed with DNA1-H_OH than it is for complexes formed with DNA1-H_H [68]. Since the increase in r_2 observed at high $[Me^{2+}]$ in the Me^{2+} titration experiments is dependent upon the presence of the primer terminal 3′-OH group (Fig. 4.8 B and D), it was of particular interest to compare the effects of Me^{2+} ions on the energy landscape between complexes formed with DNA1-H_OH and complexes formed DNA1-H_H. Is the influence of the 3′-OH group on the energy landscape of the translocation affected by the presence of Me^{2+}? Does the influence of the 3′-OH group in the translocation depend upon an interaction between the primer terminus and a Me^{2+} ion (or ions), or is it an inherent property of the interaction of the DNA substrate with the enzyme?

We examined the influence of Me^{2+} on the response of the translocation rates to applied force by comparing complexes captured at 0 Me^{2+} to those captured at 10 mM Mg^{2+}, across the range of voltages from 140 mV to 220

mV (Fig. 4.9). Experiments were conducted in the absence of Me^{2+} with both the wild type and the D12A/D66A enzymes, for complexes formed with both DNA1-H_OH and DNA1-H_H; in 10 mM Mg^{2+} with both enzymes for complexes formed with DNA1-H_H; and in 10 mM Mg^{2+} with the D12A/D66A enzyme for complexes formed with DNA1-H_OH. The higher values for r_1 in the absence of Me^{2+} relative to values for r_1 in 10 mM Me^{2+} that were observed in the Me^{2+} titration experiments conducted at 180 mV (Fig. 4.8 A and C) hold across the voltage range examined, for complexes formed with either DNA-H_OH or DNA1-H_H (Fig. 4.9 A). In addition, as previously observed for complexes captured in the presence of 10 mM Mg^{2+}, in the absence of Me^{2+}, r_1 is faster in complexes formed with DNA1-H_OH than it is for complexes formed with DNA1-H_H. For complexes formed with DNA1-H_OH, there is a modest decrease in the slope of $\log(r_1)$ vs. voltage relative to that for complexes formed with DNA1-H_H, in the absence of Me^{2+} or in the presence of 10 mM Mg^{2+} (Fig. 4.9 A). The 3′-OH group exerts an effect on the slope of $\log(r_2)$ vs. voltage that is more substantial in relative magnitude to its effect on the slope of $\log(r_1)$ vs. voltage. In complexes with DNA1-H_OH, the slope of $\log(r_2)$ vs. voltage is increased relative to that for complexes formed with DNA1-H_H, in the presence of 10 mM Mg^{2+} or the absence of Me^{2+} (Fig. 4.9 B).

We compared the absolute values of the slopes of the fitting lines in the $\log(rate)$ vs. voltage plots (Fig. 4.9 C and D). The slope values for $\log(r_1)$ vs.

voltage as well as those for $\log(r_2)$ vs. voltage show two discrete clusters. For r_1, a cluster for complexes formed with DNA1-H_OH is centered ~ 0.019, and a cluster for complexes formed with DNA1-H_H is centered at ~ 0.028 (Fig. 4.9 C). In the case of r_2, a cluster for complexes formed with DNA1-H_OH is centered ~ 0.02, and a cluster for complexes formed with DNA1-H_H is centered at ~ 0.008 (Fig. 4.9 D). For both $\log(r_1)$ vs. voltage and $\log(r_2)$ vs. voltage, the dependence of slope on force is a function of the primer terminus identity, regardless of the presence or absence of Me^{2+}.

The influence of Me^{2+} on complementary dNTP binding. Me^{2+} ions are essential for both dNTP binding and catalysis in the polymerase active site. The identity of the Me^{2+} ions bound in the polymerase active site can strongly influence dNTP binding affinity and binding rates [25], as well as the fidelity of dNTP selection and the ability to catalyze phosphodiester bond formation [109, 35, 56 and references therein]. Because the forward translocation fluctuation precedes dNTP binding, dNTP binds to complexes only when they are in the post-translocation state (Fig. 4.2 B) [67]. In the nanopore assay the two translocation states are fully resolved and the effects of dNTP on the equilibrium between the two translocation states can be quantified. Furthermore, since dNTP stabilizes the lower amplitude, post-translocation state (Fig. 4.2 A, ii) by extracting and analyzing dwell time samples in a stochastic model, the rates of dNTP binding to and dissociating from the post-translocation state can be quantified [67].

We compared the binding affinity for dGTP (complementary to dCMP at n=0 of DNA1 in Fig. 4.1 C) and dGTP binding rates, in complexes formed with D12A/D66A enzyme, captured in the presence of either 10 mM Mg^{2+}, Mn^{2+}, or Ca^{2+} (Fig. 4.10 and Table 4.1). For Ca^{2+}, we conducted dGTP titration experiments with both DNA1-H_OH and DNA1-H_H; for Mg^{2+} and Mn^{2+}, we used DNA1-H_H in dGTP titration experiments. The affinity for dNTP among different DNAP-DNA complexes can be directly compared by plotting the normalized $p/(1-p)$, where p is the probability of post-translocation state occupancy. The normalized $p/(1-p)$ is defined as the value of $p/(1-p)$ in the presence of a given concentration of dNTP, divided by the value of $p/(1-p)$ for the same complex captured in the absence of dNTP (in the presence of Me^{2+} of the same identity and concentration). The normalized $p/(1-p)$ is
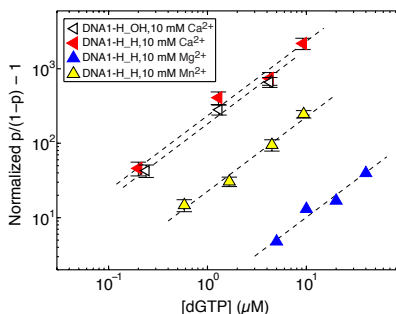
Figure 4.10: Complementary dNTP binding affinity compared as a function of Me^{2+} species and primer terminus identity. Complexes were formed between the D12A/D66A mutant of phi29 DNAP and DNA1-H-OH in 10 mM Ca^{2+} (black unfilled sideways triangles), or with DNA1-H_H in 10 mM Ca^{2+} (red filled triangles), 10 mM Mg^{2+} (blue filled triangles), or 10 mM Mn^{2+} (yellow filled triangles). The binding affinity is examined using the log-log plot of (normalized $p/(1 - p) - 1$) vs. dGTP concentration. We have shown that the vertical intercept of the log-log plot at [dGTP] = 1 μM is given by $-\log(K_d)$ [24]. The higher the log-log plot is vertically, the smaller the value of K_d and the stronger the binding affinity. Error bars show the standard errors.

155

independent of the transitions between the two translocation states in the absence of dNTP [24, 25], which vary among the complexes and conditions compared. In addition, we have shown that (normalized $p/(1-p) - 1$) = [dNTP]/K_d, where [dNTP] is the concentration of dNTP [24]. This theoretical expression predicts that the log–log plot of (normalized $p/(1 - p) - 1$) vs [dNTP] is a straight line with slope =1 and vertical intercept = $-\log(K_d)$ at [dNTP] = 1 μM. Thus K_d is determined from the vertical intercept at [dNTP] = 1 μM, obtained in fitting observed data points to the theoretical expression. A higher vertical position in the log–log plot of (normalized $p/(1 - p) - 1$) vs [dNTP] corresponds to a smaller value for K_d and a stronger binding affinity (Fig. 4.10; Table 4.1).

For D12A/D66A complexes formed with DNA1-H_H in Mg^{2+}, the dNTP binding affinity is $K_d \approx 1$ μM (in Mg^{2+}) (Table 4.1) [25]. Both Mn^{2+} and Ca^{2+} cause an enhancement in dNTP binding affinity compared to complexes captured in Mg^{2+}, with Mn^{2+} yielding a decrease in the value of K_d by ~ 25-fold, to $K_d \approx 0.04$ μM (in Mn^{2+}), and Ca^{2+} decreasing the value of K_d more significantly, by an additional order of magnitude to $K_d \approx 0.004$ μM (in Ca^{2+}). The stronger affinity afforded in Mn^{2+} relative to Mg^{2+} is primarily due to an ~ 18-fold decrease in the dNTP dissociation rate, in good accord with an earlier comparison between dNTP binding parameters for the D12A/D66A mutant captured in 10 mM Mg^{2+} or in 2 mM Mn^{2+} [25]. The more substantial enhancement in dNTP binding affinity for complexes captured in Ca^{2+} occurs

due to a combination of an ~ 9-fold increase in the dNTP association rate

constant and an ~ 23-fold decrease in the dNTP dissociation rate, relative to

the binding rates for complexes captured in Mg^{2+}. Since Ca^{2+} supports dNTP

Table 4.1: Complementary dNTP binding rates in complexes formed with the D12A/D66A mutant of phi29 DNAP.

DNA	Me^{2+}	$k_{on}(s^{-1}\mu M^{-1})$ [a]	$k_{off}(s^{-1})$ [b]	$k_{off}(s^{-1})$ / $k_{on}(s^{-1}\mu M^{-1})$	K_d (µM) [c]
DNA1-H_OH	Ca^{2+}	147.67 ± 14.75	0.767 ± 0.085	0.0052 ± 0.0006	0.00551 ± 0.0005
DNA1-H_H	Ca^{2+}	150.19 ± 13.85	0.624 ± 0.082	0.0043 ± 0.0007	0.00426 ± 0.0005
DNA1-H_H	Mg^{2+}	16.7 ± 0.4	17.7 ± 0.3	1.02 ± 0.02	0.988 ± 0.091
DNA1-H_H	Mn^{2+}	21.85 ± 2.81	0.978 ± 0.13	0.0448 ± 0.0067	0.0444 ± 0.0036

[a] The dNTP association rate constant.

[b] The dNTP dissociation rate.

[c] K_d values for the binding of dGTP to the post-translocation state are determined from the vertical intercepts of the fitting lines to the log-log plot of (Normalized p/(1–p) −1) vs. [dGTP], where p is equilibrium probability of the lower amplitude level (see the model diagram in Figure 1a). The plot of (Normalized p/(1–p) −1) vs. [dGTP] is shown in Figure 9. Rates were determined using dwell time samples extracted from ionic current traces (see Methods section) and a three-state kinetic model (consisting of transitions r_1, r_2, k_{on} and k_{off} in the model diagram in Figure 1a). Complexes were captured at 180 mV in the presence of 1 mM EDTA, with either 11 mM $MgCl_2$, or 11 mM $CaCl_2$, or 11 mM $MnCl_2$ as indicated; K_d, k_{on}[dNTP] and k_{off} are independent of the applied voltage [24, 67]. All values are reported with the standard error.

binding but does not support phosphodiester bond formation, this afforded an

opportunity to examine the effect of the primer terminal 3'-OH group, a ligand

for metal A [40], on dNTP binding by comparing complexes formed with DNA1-

H_OH or with DNA1-H_H. The presence or absence of the primer terminal 3'-

OH group does not significantly affect the ground state dNTP binding

parameters (Fig. 4.10; Table 4.1); k_{on}, $k_{off,}$ and K_d are each very similar for complexes formed with both DNA substrates.

4.5 Conclusion

In this study, we hypothesized that because the divalent cations (Me^{2+} ions) that are essential for catalytic function in both the polymerase and exonuclease active sites of replicative DNAPs and serve as intimate components of the architecture in each of the sites, they may exert significant influence on the noncovalent transitions that occur in DNAP-DNA complexes during each nucleotide addition cycle. To test this, we applied a nanopore-based, single molecule approach that features single-nucleotide spatial precision and sub-millisecond temporal resolution to determine the effects of Mg^{2+}, Ca^{2+} and Mn^{2+} on the kinetic transitions of translocation, dNTP binding, and primer strand transfer between the polymerase and exonuclease active sites (Fig. 4.2 B), in individual complexes formed with phi29 DNAP, a B-family replicative DNAP.

Me^{2+} ions alter the equilibrium across the translocation step, and the forward and reverse translocation rates. Fluctuations across the translocation step, during which the DNAP moves with respect to its DNA substrate by the distance of a single nucleotide, occur in the absence of Me^{2+}, indicating that the translocation fluctuations are an activity inherent to the DNAP-DNA complex. Nonetheless, Me^{2+} ions have a significant influence on the translocation. For all phi29 DNAP-DNA complexes examined, formed with

158

either the wild type or D12A/D66A enzyme, and either the DNA1-H_OH or DNA1-H_H substrate, Me^{2+} causes a concentration-dependent decrease in the probability of the post-translocation state (p) (Fig. 4.4). This shift in equilibrium is promoted by Mg^{2+} and Mn^{2+}, which support DNAP catalysis, and Ca^{2+}, which supports dNTP binding but not catalysis. It is caused primarily by a [Me^{2+}]-dependent decrease in the forward translocation rate (r_1) (Fig. 4.8 A and C). The Me^{2+}-dependent decrease in r_1 may be attributable to stabilization of the fingers-closed pre-translocation state, which would be predicted to slow the forward translocation [6]. The pre-translocation state that is sampled during the nanopore experiments corresponds to the product of the covalent addition of the primer terminal residue during DNA synthesis, after pyrophosphate has dissociated from the complex (presumably complexed with the metal B ion) [65, 68]. While we cannot assign whether the Me^{2+}-dependent decrease in r_1 is caused by binding of Me^{2+} at the metal A or metal B site, the Me^{2+}-dependent effects on the escape from the pre-translocation state suggest that while the environment of both metal binding sites necessarily changes significantly after catalysis, at least one of the sites retains affinity for Me^{2+}.

Each of the three metals examined also promoted a modest concentration-dependent increase in the backward transition rate from the post-translocation to pre-translocation state (r_2) that is dependent upon the presence of a 3′-OH group at the primer terminus of the DNA substrate (Fig.

4.8 B and D). The dependence of the increase in r_2 on the primer strand 3′-OH implicates this group as a ligand for the metal ion that elicits the increase, suggesting that it is the post-translocation state metal A ion and implying that the metal A ion binds to the phi29 DNAP-DNA complexes independent of the presence of dNTP. In complexes formed with a DNA substrate bearing a 3′-H primer terminus, where the increase in r_2 is not observed (Fig. 4.8 D), we cannot distinguish between the case in which metal A does not bind, or the case in which it binds but does not cause an increase in r_2.

In the nanopore experiments, the applied voltage exerts a force on the captured DNAP-DNA complex along the direction of the translocation. While Me^{2+} modulates the translocation rates at a given applied force (Fig. 4.8), it does not alter the dependence of the rates on force, for complexes formed with either DNA1-H_OH or with DNA1-H_H (Fig. 4.9). The primer terminal 3′-OH group is a determinant in the energy landscape of the translocation [68]; its presence alters the slopes of both $\log(r_1)$ vs. voltage and $\log(r_2)$ vs. voltage. When complexes captured at 0 Me^{2+} are compared to those captured at 10 mM Mg^{2+}, across a range of applied voltages from 140 mV to 220 mV (Fig. 4.9 A and B), the dependence of the slopes of both $\log(r_1)$ vs. voltage and $\log(r_2)$ vs. voltage are observed to be a function of the primer terminus identity, regardless of the presence or absence of Me^{2+} (Fig. 4.9 C and D). The smaller slope of $\log(r_1)$ vs. voltage and larger slope of $\log(r_2)$ vs. voltage for complexes formed with DNA1-H_OH relative to complexes formed with

DNA1-H_H indicates that the transition state for the translocation is closer to the pre-translocation state when the DNA substrate bears a primer terminal 3′-OH group than when the DNA substrate bears a primer terminal 3′-H group, independent of the presence or absence of Me^{2+}. Thus, Me^{2+} affects the translocation rates at each given voltage, but not the dependence of the rates on force, indicating that Me^{2+} does not perturb the translocation distances among the pre-translocation state well, the post-translocation state well, and the transition state, which are strongly influenced by the primer terminal 3′-OH group. Like the translocation fluctuations, the influence of the primer terminal 3′-OH on the energy landscape of the translocation is inherent to the DNAP-DNA complex.

Me^{2+} ion identity affects complementary dNTP binding equilibrium and rates. Complementary dNTP binding in the post-translocation state polymerase active site is strongly affected by Me^{2+} identity, with Ca^{2+} affording the highest affinity, followed by Mn^{2+}, and then Mg^{2+} (Fig. 4.10; Table 4.1). Both Ca^{2+} and Mn^{2+} substantially decrease the dNTP dissociation rate relative to Mg^{2+}; Ca^{2+} also yields an increase in the dNTP association rate constant. In contrast to Mn^{2+} and Mg^{2+}, Ca^{2+} does not support phosphodiester bond formation, yielding the opportunity to examine the effect of the primer terminal 3′-OH group, which is a ligand for metal A [40], on dNTP binding. Somewhat surprisingly, the presence or absence of the primer terminal 3′-OH group does not significantly affect the ground state dNTP binding parameters (Fig. 4.10;

Table 4.1), which are very similar for complexes formed with DNA1-H_OH or with DNA1-H_H. This indicates that the presence or absence of the primer terminal 3'-OH group is not a determinant for ground state dNTP binding in Ca^{2+}.

Because selective stabilization of the transition state over the ground state contributes to catalytic efficiency and specificity [38], it is possible that the enhanced ground state stabilization effected by both Mn^{2+} and Ca^{2+} contributes to the effects of these metals on phosphodiester bond formation. Since k_{off} reflects the free energy difference between the dNTP-bound post-translocation state and the dNTP-unbound post-translocation state, the smaller values of k_{off} for complexes formed in Mn^{2+} or Ca^{2+} relative to complexes formed in Mg^{2+} may obtain, at least in part, from a deeper free energy well for the dNTP-bound ground state. In addition to affecting the free energy difference between the dNTP-bound and unbound states, a deeper free energy well could affect the free energy difference between the dNTP-bound ground state and the transition state for catalysis. Among the three metals we examined, the largest stabilization of ground state dNTP binding is caused by Ca^{2+}, and this metal does not support catalytic function. Ground state stabilization by Mn^{2+}, while an order of magnitude more modest than Ca^{2+}, is almost 20-fold greater than that afforded by Mg^{2+}; while Mn^{2+} supports catalysis, it is at a cost to fidelity in nucleotide selection.

The relationship among Me^{2+} ions, dNTP binding, and conformational

transitions associated with dNTP binding has been studied for pol β [3], for the Klenow fragment of DNA polymerase I (an A-family repair DNAP) [7], and for the B-family replicative DNAP from bacteriophage RB69 [99, 64]. Not surprisingly, details vary among these enzymes, which have different functional roles and are from different DNAP families. Nonetheless, the studies are consistent with a general scenario in which metal B is sufficient to support dNTP binding and the associated fingers-closing conformational change, while both metals A and B are required for phosphodiester bond formation.

The effects of Me^{2+} ions on dNTP binding or on the translocation that we have measured in phi29 DNAP-DNA complexes may be related, at least in part, to Me^{2+}-dependent effects on the rates of fingers opening and closing. Based on earlier findings that both r_1 and k_{off} are slower for complexes captured in 2 mM Mn^{2+} than for complexes captured in 10 mM Mg^{2+}, we had hypothesized that Mn^{2+} may exert its effects on both r_1 and k_{off} by decreasing the rate of fingers opening from the closed complex [25]. This hypothesis is based upon the structural model proposed for translocation [6], which predicts that decreasing the rate of fingers opening in the pre-translocation state would lead to a decrease in the rate of the forward translocation (r_1). Similarly, since it is reasonable to assume that dNTP dissociation requires fingers opening, a decrease in the rate of fingers opening could yield a decrease in the dissociation rate of dNTP (k_{off}) from the dNTP-bound post-translocation state.

The findings in the current study that both r_1 and k_{off} are lower in Ca^{2+} (Fig. 4.8; Table 4.1) than they are in Mg^{2+} provides further evidence supporting the hypothesis.

Complex effects of Me^{2+} on the primer strand transfer pathway between the polymerase and exonuclease active sites. We previously established the correspondence between a second kinetic state in the upper amplitude and the primer strand transfer between the polymerase and exonuclease sites [68]. In the current study, we examined the effects of Me^{2+} ions on primer strand transfer by applying approaches developed in the prior study, which revealed that the primer strand pathway displays significant Me^{2+}-concentration dependent kinetic complexity. In the absence of Me^{2+}, or when the [Me^{2+}] is low, the pathway for primer strand transfer between the polymerase and exonuclease sites displays more than two kinetic states (Fig. 4.6 A). For complexes formed with DNA1-H_OH, this complexity resolves to two states (the pre-translocation state in the polymerase site, and the state in which the primer has been transferred to the exonuclease site) when the [Me^{2+}] is above ~ 1 mM (Figs. 4.5 and 4.6A) [68]. For complexes formed with DNA1-H_H captured in the presence of high [Me^{2+}], dwell time samples that correspond to stays of the primer strand in the exonuclease site cannot be resolved from the dwell time samples that decay at a single exponential rate that corresponds to the inverse of the forward translocation rate [66, 68]. While these observed effects of [Me^{2+}] are intriguing, we are not yet able to

164

speculate about the mechanisms governing the effects of Me^{2+} on the kinetics of the primer strand transfer pathway. Detailed dissection of the effects of $[Me^{2+}]$ on this pathway will require biochemical and mathematical models significantly more complex than the four-state model shown in Fig. 4.2 B, which we will pursue in future studies.

Finally, it is not yet known how generally applicable our findings regarding the effects of Me^{2+} ions on translocation, dNTP binding, or primer strand transfer with the phi29 DNAP will be to other DNAPs, in large part because most techniques that have been employed to examine DNAPs cannot resolve and quantify the translocation step and the kinetic transitions that it coordinates with both single nucleotide precision and sub-millisecond temporal resolution. However, given the structural and functional conservation of the polymerase and exonuclease active sites in the B-family replicative DNAPs, it is reasonable to expect that these effects will apply to other enzymes in this family. This includes the eukaryotic nuclear leading strand and lagging strand replicative DNAPs, Pol-ε and Pol-δ, respectively [80, 75]. In light of the finding that Me^{2+}, which stabilizes the pre-translocation state in phi29 DNAP-DNA complexes (Fig. 4.8 A and C), can also stabilize the pre-translocation state in DNAP-DNA complexes of the X-family mammalian repair DNAP, pol β [22], it is also possible that our findings will extend more broadly to other DNAP families.

Chapter 5

Conclusion

The advent of nanopore technology has opened the door for diverse applications to study single molecules with various objectives. Examples include protein and nucleic acid sequence and structure studies as well as biophysical studies of molecular motors [41, 24, 25, 26, 27, 68, 65, 5, 2, 30 and references therein]. We have developed an α-HL nanopore based technique to study the activity of the replicative DNA polymerase (DNAP) from the bacteriophage phi29 [20, 24, 25, 26, 27, 65, 66, 67, 68]. This enzyme is an excellent model for B-family replicative DNAPs because it is highly processive and has the ability to perform strand displacement without the need for accessory proteins like sliding clamps and helicases [6, 73, 12, 13, 85].

Possible limitations of the established nanopore assay. Like all tools developed for biochemical and biophysical analysis the nanopore is not without its constraints. Two important considerations are how the force applied to the DNAP at the interface with the α-HL channel may perturb measurements of the DNAP-DNA complex and how the buffer components of the nanopore experiments may change the activity of phi29 DNAP compared to a cellular environment. An important next step in addressing these specific questions would be to determine the rate of nucleotide incorporation, in single

turnover and processive assays, for individual complexes captured atop the nanopore for comparison to *in-vitro* ensemble experiments at a defined set of buffer conditions. An accurate description of the rates of synthesis for nanopore experiments in contrast to replicating complexes under different conditions could provide new mechanistic details.

Other technologies exist that have explored the behavior of phi29 DNAP replication on a DNA substrate under tension [73, 49, 74]. While the tension in these experiments applies very different energetic constraints to a replicating DNAP than the force applied to complexes captured with the nanopore it would be insightful to know the actual forces applied to the DNAP, across the range of voltages used in the nanopore studies presented here, for comparison to related kinetic values determined for complexes studied by alternative approaches. While these studies cannot resolve the translocation step with single nucleotide resolution they have provided some insight into the balance between the catalytic activities of the polymerase and exonuclease active sites [49, 74].

Phi29 DNAP is an excellent model enzyme. However, other replicative DNA polymerases, with the exception of the A-family DNAP Pol-I Klenow Fragment from *E. coli* [41], have yet to be captured and measured on the nanopore to study the fidelity of replicative DNAPs. Application of this technology to study other DNAPs, for example the viral DNAP from RB69 and the eukaryotic DNAPs responsible for leading and lagging strand replication, pol-ε and pol-δ

respectively, could parallel the mechanistic observations detailed in the studies conducted with phi29 DNAP. Application of existing nonopore technology to study related DNAPs could provide new data to support the findings of the studies conducted with phi29 DNAP. Such data would substantiate the utility of the this technique and corroborate the broader implications of the mechanistic descriptions that have been so elegantly established for phi29 DNAP.

Beyond corroborating the existing phi29 DNAP data, studies of other polymerases would create an opportunity to study more components of the eukaryotic DNA replisome, for example the processivity factor PCNA and the many diverse factors that associate with it. It would be mechanistically insightful to determine if PCNA affects the rate of translocation specifically or if it increases the activity of the polymerase in a more dynamic context. Nanopore technology presents a unique opportunity to make these measurements. That said, in unpublished nanopore experiments the fragment of yeast pol-ε that was used to determine one of its crystal structures [48] was drawn to the pore independent the presence of DNA. It is unclear if the residues removed, to promote high resolution structure determination, may have exposed an anionic surface that caused the protein to be drawn to the pore or if there is some other net anionic state of the protein that led to these observations. A focused effort to either modify the nanopore conditions or engineer pol-ε to make it amenable to standard nanopore experiments may

be required. Unfortunately, engineering enzymes in this way opens the door to speculation about the validity and broad implications of the data collected.

Alternative single molecule assays to further elucidate the mechanistic findings of nanopore data. It is important to discuss what orthogonal approaches could be used to support the findings of the nanopore experiments presented in the preceding chapters. A common mechanism was evoked for observations of the Y226 and Y390 mutant versions of phi29 DNAP [25], those exploring the mechanism for stable ribonucleotide incorporation [26] and those that examined the role of divalent cations on the non-covalent transitions associated with phi29 DNAP [27]; a change in the rates of the fingers opening and closing transitions. The nanopore cannot make this measurement directly and instead this process can only be inferred.

Ensemble and single molecule FRET studies of the Klenow fragment of *E. coli* pol-I, an A-family DNAP, conducted in the laboratory of Catherine Joyce labeled the fingers subdomain and the DNA substrate to measure the dynamics of the fingers-closing transition [52, 7]. Their approach required mutating all solvent exposed cysteine residues except one on the fingers subdomain, to which they conjugated a fluorophore. The techniques to tag a protein for single molecule labeling advanced considerably since the experiments conducted by Joyce et. al. [81 and references therein]. It is no longer essential to mutate all solvent exposed cysteine residues, instead a small

amino acid sequence can be added to the protein and a fluorophore can be enzymatically transferred with high efficiency and specificity. Using this labeling method to make dynamic measurements of the fingers subdomain could elucidate instances of perturbation to the fingers sub-domain movement evoked in the nanopore studies.

A hurdle to testing observations from nanopore experiments with FRET experiments is a clear distinction of the translocation state. In the pre-translocation state the fingers are closed. In the post translocation state the fingers must be open to accommodate nucleotide binding. Thus, across the translocation step the fingers are proposed to open. In concert with nucleotide binding the fingers close to determine complementarity of the nascent base pair and assemble the active site for chemical transformation. If the only measure in such FRET experiments is the rate of fingers opening and closing then the translocation transition may kinetically complicate the accuracy of fingers open and closing transitions. In the post translocation state there is an experimental variable to attenuate the fingers opening and closing transition, dNTP concentration. Presumably the rates of fingers opening and closing across the translocation step should not be perturbed by the rates of fingers opening and closing upon dNTP binding and release. Testing the rates of fingers opening and closing in the presence of varying amounts of complementary dNTP has been used to determine the rates associated with finger subdomain movement during nucleotide binding [8, 51, 52].

For the primer strand to be transferred from the polymerase to the exonuclease active site the fingers sub-domain is proposed to, at least partially, open in the pre-translocation state. Mutations affecting the rates of fingers opening and closing in this instance cannot be perturbed by addition of an experimental variable. I propose that a three color FRET system that describes the spatial position of the DNA polymerase on the DNA substrate, with single nucleotide spatial resolution, could operate to describe which translocation state the DNAP is in during dynamic movement of the fingers subdomain. The design of such an experiment could utilize one FRET pair to measure the location of the DNAP with respect to the primer-template duplex while the other FRET pair would be placed in a manner consistent with the pair utilized by Joyce et. al., reporting on the fingers subdomain movement with respect to the template.

Additional structural data could provide insight into the mechanisms of DNAP fidelity. More structural data could inform the observations from nanopore experiments. The nanopore studies presented herein describe changes in kinetic rates as phi29 DNAP transitions between covalently identical states in a single nucleotide addition cycle. The rate changes described are associated with changes in the energy landscape across a given transition and not a change in the covalent nature of the substrates and enzyme. In these instances, changes in the shape of an energy well indicate a change in the stability of a complex. The changes in the rates determined in the preceding

nanopore studies describe a shift in the equilibrium between low energy states of a dynamic complex.

The structural data describing phi29 DNAP is limited when considering the extensive body of biochemical and biophysical data regarding this polymerase [6, 9, 11, 12, 13, 15, 24, 25, 26, 27, 29, 34, 35, 36, 42, 49, 58, 69, 73, 74, 85, 86, 89, 95]. Several conserved residues have been identified as significant in the mechanisms of fidelity for this DNAP. Among the residues explored with the nanopore data are D12, N62, D66, Y226, Y254, Y390 [24, 25, 26, 27]. D12 and D66 are the aspartic acid residues which act as ligands for Me^{2+} binding in the exonuclease active site [9, 42]. N62 is associated with the primer strand transfer pathway between the polymerase and exonuclease active sites [58]. Y226 and Y390 were shown to affect substrate selection and replication [11, 95]. As well, Y254 is responsible for selection of complementary dNTP over rNTP and ground state stabilization of dNTP binding [6, 14, 86, 15]. Structural comparison of WT phi29 to the various mutants used in the nanopore studies in complex with DNA, Me^{++} (Ca^{++}, Mg^{++} and Mn^{++}) and complementary dNTP or rNTP substrates could inform the relationship that exists between structure and function for low energy states resolved in structures and kinetic data determined by nanopore experiments to provide a more robust mechanistic description of fidelity in DNA polymerases.

A specific example where more structural data is needed to understand the physical nature of the kinetic data provided by nanopore experiments is the

D12A/D66A mutant of phi 29 DNAP. Nanopore experiments have demonstrated the D12A/D66A mutant of phi29 DNAP has a higher affinity for nucleotide substrate than the WT enzyme [24, 25, 26, 67]. The exonuclease deficient mutant of phi29 DNAP (D12A/D66A) has a direct effect on the catalytic activity of the exonuclease active site and a pleiotropic effect on nucleotide binding in the polymerase active site, both of which contribute to the fidelity of DNAPs. This indicates that the structural study describing translocation in B-family DNAPs, using phi29 DNAP as a model [6], could have a unique ternary complex that may not reflect changes across the translocation step in the natural binary and ternary complexes.

Ensemble kinetic studies to probe mechanisms not measured by the nanopore. Finally as an orthogonal approach, pre-steady state kinetics can contribute important mechanistic information regarding not only the specific observations from nanopore data but also steps in the mechanism that may not be resolved by the nanopore. Previously such experiments have been instructional in the dynamic mechanism of bacteriophage DNAPs from T7 & RB69 [78, 108, 97], Klenow Fragment of *E. coli* pol-I [60, 61], and eukaryotic DNAPs β, γ, η, ι, δ & ε [1, 30, 105, 112, 45, 102, 103, 31, 112]. For example the rates of DNAP-DNA association, complementary and non-complementary dNTP binding and phosphodiester bond formation have been measured for some of these DNA polymerases.

Phosphodiester bond formation is described by the catalytic rate constant k_{cat}, and may be affected in each of the phi29 DNAP mutants described herein. Similarly k_{cat} could be perturbed when comparing the rates of rNTP and dNTP incorporation for the wild type enzyme. If the rate of chemistry for aberrant rNTP substrate is distinct from the canonical dNTP substrate then another mechanism for DNAP fidelity could be evoked beyond what was determined by the nanopore experiments. An alternative example is the case of the exonuclease deficient mutant of phi29 DNAP (D12A/D66A). This mutant shows an increase in the affinity for complementary nucleotide [24, 25]. The increase in the binding affinity (decrease in K_d) may or may not be associated with a concomitant change in the rate of chemistry. Furthermore, k_{cat} may be affected by the identity of the metal ion species when comparing the two catalytically proficient divalent metal ions described, Mg^{2+} and Mn^{2+}.

Another important mechanistic transition the nanopore was unable to resolve, with the exception of the Y226F mutation of phi29 DNAP, is the binding and release of pyrophosphate (PPi). PPi is the product of dNTP incorporation to the end of the primer terminus in the polymerase active site. The release of PPi has been proposed as a potential mechanistic determinant between the Brownian Ratchet and Power Stroke models for DNAP translocation. The data collected with the nanopore indicate that the polymerase fluctuates across the translocation step by Brownian motion, however, this does not mean that PPi release has no effect on the rates of translocation. The

nanopore data collected with the Y226F mutation of phi29 DNAP, captured in the presence of PPi, indicate that PPi changes the rate of transition from the pre to the post-translocation state. This observation was unique to this mutant, which also significantly increased the affinity for complementary nucleotide [25]. Without the ability to measure the binding and release of PPi and translocation simultaneously in the wild type background an accurate description for the mechanism of translocation by replicative DNAPs during DNA synthesis remains elusive. Pre-steady state kinetic experiments provide an alternative approach to probe this essential mechanistic distinction, a distinction that is intractable with current nanopore technology.

Supplemental Files

For Chapter 3: Supporting Information Text (Mathematical derivations); Figures S1-S4. This material is available free of charge via the Internet at http://pubs.acs.org.

Bibliography

1. Ahn, J., Werneburg B. G., and Tsai M. D., (1997) DNA polymerase : structure-fidelity relationship from pre-steady-state kinetic analyses of all possible correct and incorrect base pairs for wild type and R283A mutant. *Biochemistry* **36**, 1100–1107

2. Akeson, M., Branton, D., Kasianowicz, J. J., Brandin, E., and Deamer, D. W. (1999) Microsecond time-scale discrimination among polycytidylic acid, polyadenylic acid, and polyuridylic acid as homopolymers or as segments within single RNA molecules. *Biophys J* **77**, 3227–3233

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

4. Batra, V. K., Beard, W. A., Shock, D. D., Krahn, J. M., Pedersen, L. C., and Wilson, S. H. (2006) Magnesium-induced assembly of a complete DNA polymerase catalytic complex. *Structure* **14**, 757–766

5. Benner, S., Chen, R. J. A., Wilson, N. A., Abu-Shumays, R., Hurt, N., Lieberman, K. R., Deamer, D. W., Dunbar, W. B., and Akeson, M. (2007) Sequence-specific detection of individual DNA polymerase complexes in real time using a nanopore. *Nat Nanotechnol* **2**, 718–724

6. Berman, A. J., Kamtekar, S., Goodman, J. L., Lázaro, J. E. M., de Vega, M., Blanco, L., Salas, M., and Steitz, T. A. (2007) Structures of Φ29 DNA polymerase complexed with substrate: the mechanism of translocation in B-family polymerases. *Embo J* **26**, 3494–3505

7. Bermek, O., Grindley, N. D. F., and Joyce, C. M. (2010) Distinct roles of the active-site Mg2+ ligands, Asp882 and Asp705, of DNA polymerase I (Klenow fragment) during the prechemistry conformational transitions. *J Biol Chem* **286**, 3755–3766

8. Bermek, O., Grindley, N. D. F., and Joyce, C. M. (2013) Prechemistry nucleotide selection checkpoints in the reaction pathway of DNA polymerase I and roles of glu710 and tyr766. *Biochemistry* **52**, 6258–6274

9. Bernad, A., Blanco, L., Lázaro, J. M., Mart i n, G., and Salas, M. (1989) A conserved 3′-5′ exonuclease active site in prokaryotic and eukaryotic DNA polymerases. *Cell* **59**, 219–228

10. Beese, L. S., and Steitz, T. A. (1991) Structural basis for the $3'$-$5'$ exonuclease activity of Escherichia coli DNA polymerase I: a two metal ion mechanism. *Embo J* **10**, 25–33

11. Blanco, L., Bernad, A., Blasco, M. A., and Salas, M. (1991) A general structure for DNA-dependent DNA polymerases. *Gene* **100**, 27–38

12. Blanco, L., Bernad, A., Lázaro, J. M., Mart i n, G., Garmendia, C., and Salas, M. (1989) Highly efficient DNA synthesis by the phage Φ29 DNA polymerase. Symmetrical mode of DNA replication. *J Biol Chem* **264**, 8935–8940

13. Blanco, L., and Salas, M. (1996) Relating structure to function in Φ29 DNA polymerase. *J Biol Chem* **271**, 8509–8512

14. Blasco, M. A., Lázaro, J. M., Bernad, A., Blanco, L., and Salas, M. (1992) Φ29 DNA polymerase active site. Mutants in conserved residues Tyr254 and Tyr390 are affected in dNTP binding. *J Biol Chem* **267**, 19427–19434

15. Bonnin, A., Lázaro, J. M., Blanco, L., and Salas, M. (1999) A single tyrosine prevents insertion of ribonucleotides in the eukaryotic-type phi29 DNA polymerase. *J Mol Biol* **290**, 241–251

16. Brown, J. A., and Suo, Z. (2011) Unlocking the sugar "steric gate" of DNA polymerases. *Biochemistry* **50**, 1135–1142

17. Caldecott, K. W. (2014) Molecular biology. Ribose--an internal threat to DNA. *Science* **343**, 260–261

18. Capson, T. L., Peliska, J. A., Kaboord, B. F., Frey, M. W., Lively, C., Dahlberg, M., and Benkovic, S. J. (1992) Kinetic characterization of the polymerase and exonuclease activities of the gene 43 protein of bacteriophage T4. *Biochem.* **31**, 10984–10994

19. Cavanaugh, N. A., Beard, W. A., Batra, V. K., Perera, L., Pedersen, L. G., and Wilson, S. H. (2011) Molecular insights into DNA polymerase deterrents for ribonucleotide insertion. *J Biol Chem* **286**, 31650–31660

20. Cherf, G. M., Lieberman, K. R., Rashid, H., Lam, C. E., Karplus, K., and Akeson, M. (2012) Automated forward and reverse ratcheting of DNA in a nanopore at 5-Å precision. *Nat Biotechnol* **30**, 344–348

21. Clark, A. B., Lujan, S. A., Kissling, G. E., and Kunkel, T. A. (2011) Mismatch repair-independent tandem repeat sequence instability resulting

from ribonucleotide incorporation by DNA polymerase ε. *DNA Repair (Amst)* **10**, 476–482

22. Clausen, A. R., Zhang, S., Burgers, P. M., Lee, M. Y., and Kunkel, T. A. (2013) Ribonucleotide incorporation, proofreading and bypass by human DNA polymerase δ. *DNA Repair (Amst)* **12**, 121–127

23. Colquhoun, D., and Sigworth, F. J. (1995) in *Single-channel recording.* (Sakmann, B., and Neher, E., eds.), pp. 483–587, Plenum Press, New York

24. Dahl, J. M., Mai, A. H., Cherf, G. M., Jetha, N. N., Garalde, D. R., Marziali, A., Akeson, M., Wang, H., and Lieberman, K. R. (2012) Direct observation of translocation in individual DNA polymerase complexes. *J Biol Chem* **287**, 13407–13421

25. Dahl, J. M., Wang, H., Lázaro, J. M., Salas, M., and Lieberman, K. R. (2014) Dynamics of translocation and substrate binding in individual complexes formed with active site mutants of phi29 DNA polymerase. *J Biol Chem* **289**, 6350–6361

26. Dahl, J. M., Wang, H., Lázaro, J. M., Salas, M., and Lieberman, K. R. (2014) Kinetic mechanisms governing stable ribonucleotide incorporation in individual DNA polymerase complexes. *Biochemistry* **53**, 8061–8076

27. Dahl, J.M., Wang, H., Lieberman K.R., (2016) Modulation of DNA Polymerase Noncovalent Kinetic Transitions by Divalent Cations. *J Biol Chem* **291**, 6456-6470

28. Dalgaard, J. Z. (2012) Causes and consequences of ribonucleotide incorporation into nuclear DNA. *Trends in Genetics* **28**, 592–597

29. de Vega, M., Blanco, L., and Salas, M. (1999) Processive proofreading and the spatial relationship between polymerase and exonuclease active sites of bacteriophage phi29 DNA polymerase. *J Mol Biol* **292**, 39–51

30. Deamer, D., Akerson, M., and Branton, D., (2016) Three decades of nanopore sequencing. *Nat. Nanotech*, **35**(5), 518-524

31. Dieckman, L. M., Johnson, R. E., Prakash, S., and Washington, M. T. (2010) Pre-steady state kinetic studies of the fidelity of nucleotide incorporation by yeast DNA polymerase delta. *Biochemistry* **49**, 7344–7350

32. Doublié, S., Tabor, S., Long, A. M., Richardson, C. C., and Ellenberger, T. (1998) Crystal structure of a bacteriophage T7 DNA replication complex at 2.2 Angstrom resolution. *Nature* **391**, 251–258

33. Eder, P. S., Walder, R. Y., and Walder, J. A. (1993) Substrate specificity of human RNase H1 and its role in excision repair of ribose residues misincorporated in DNA. *Biochemistry* **75**, 123–126

34. Esteban, J. A., Bernad, A., Salas, M., and Blanco, L. (1992) Metal activation of synthetic and degradative activities of phi 29 DNA polymerase, a model enzyme for protein-primed DNA replication. *Biochemistry* **31**, 350–359

35. Esteban, J. A., Salas, M., and Blanco, L. (1993) Fidelity of phi 29 DNA polymerase. Comparison between protein-primed initiation and DNA polymerization. *J Biol Chem* **268**, 2719–2726

36. Esteban, J. A., Soengas, M. S., Salas, M., and Blanco, L. (1994) 3-->5′ exonuclease active site of Φ29 DNA polymerase. Evidence favoring a metal ion-assisted reaction mechanism. *J Biol Chem* **269**, 31946–31954

37. Eom, S. H., Wang, J., and Steitz, T. A. (1996) Structure of Taq polymerase with DNA at the polymerase active site. *Nature* **382**, 278–281

38. Fersht, A. (1985) *Enzyme Structure and Mechanism*, (Second Edition), W.H. Freeman, New York

39. Ferraro, P., Franzolin, E., Pontarin, G., Reichard, P., and Bianchi, V. (2010) Quantitation of cellular deoxynucleoside triphosphates. *Nucleic Acids Res* **38**, e85

40. Franklin, M. C., Wang, J., and Steitz, T. A. (2001) Structure of the replicating complex of a pol alpha family DNA polymerase. *Cell* **105**, 657–667

41. Garalde, D. R., Simon, C. A., Dahl, J. M., Wang, H., Akeson, M., and Lieberman, K. R. (2011) Distinct complexes of DNA polymerase I (Klenow fragment) for base and sugar discrimination during nucleotide substrate selection. *J Biol Chem* **286**, 14480–14492.

42. Garmendia, C., Bernad, A., Esteban, J. A., Blanco, L., and Salas, M. (1992) The bacteriophage phi 29 DNA polymerase, a proofreading enzyme. *J Biol Chem* **267**, 2594–2599.

43. Ghodgaonkar, M. M., Lazzaro, F., Olivera-Pimentel, M., Artola-Borán, M., Cejka, P., Reijns, M. A., Jackson, A. P., Plevani, P., Muzi-Falconi, M., and Jiricny, J. (2013) Ribonucleotides misincorporated into DNA act as strand-discrimination signals in eukaryotic mismatch repair. *Mol Cell* **50**, 323–332.

44. Göksenin, A. Y., Zahurancik, W., LeCompte, K. G., Taggart, D. J., Suo, Z., and Pursell, Z. F. (2012) Human DNA polymerase ε is able to efficiently extend from multiple consecutive ribonucleotides. *J Biol Chem* **287**, 42675–42684.

45. Graves, S. W., Johnson, A. A., Johnson, K. A. (1998) Expression purification and initial kinetic characterization of the large subunit of the mitochondrial DNA polymerase. *Biochemistry*, **37**, 6050-6068

46. Hanson J, Huxley H. E., (1955) The structural basis of contraction in striated muscle. *Symp Soc Exp Biol* **9**, 228–264

47. Hogg, M., Wallace, S. S., Doublié, S. (2004) Crystallographic snapshots of a replicative DNA polymerase encountering an abasic lesion. *EMBO J*, **23**(7), 1483-1493

48. Hogg, M., Osterman, P., Bylund, G. O., Ganai, R. A., Lundström, E.-B., Sauer-Eriksson, A. E., and Johansson, E. (2014) Structural basis for processive DNA synthesis by yeast DNA polymerase ε. *Nat Struct Mol Biol* **21**, 49–55.

49. Ibarra, B., Chemla, Y. R., Playsunov, S., Smith, S.B., Lázaro, J. M., Salas, M., and Bustamante, C. (2009) Proofreading dynamics of a processive DNA polymerase *EMBO J.* **28**, 2794-2802

50. Joyce, C. M. (1989) How DNA travels between the separate polymerase and 3-5-exonuclease sites the DNA polymerase I (Klenow Fragment). *J Biol Chem*, **264**, 10858-10866

51. Joyce, C. M. (1997) Choosing the right sugar: how polymerases select a nucleotide substrate. *Proc Natl Acad Sci U S A* **94**, 1619–1622.

52. Joyce, C. M., Potapova, O., Delucia, A. M., Huang X., Basu V. P., Grindly N.D.F., (2008) Fingers-closing and other rapid conformational changes in DNA polymerase-I (Klenow Fragment) and their role in nucleotide selectivity. *Biochemistry*, **47**(23):6103-6116

53. Johansson, E., and Macneill, S. A. (2010) The eukaryotic replicative DNA polymerases take shape. *Trends Biochem Sci* **35**, 339–347.

54. Johansson, E., and Dixon, N. (2013) Replicative DNA polymerases. *Cold Spring Harb Perspect Biol 5:a012799*, 1-14.

55. Johnson, K. A. (1993) Conformational coupling in DNA polymerase fidelity. *Annu Rev Biochem* **62**, 685–713.

56. Johnson, K. A. (2010) The kinetic and chemical mechanism of high-fidelity DNA polymerases. *Biochim Biophys Acta* **1804**, 1041–1048

57. Johnson, S. J., Taylor, J. S., and Beese, L. S. (2003) Processive DNA synthesis observed in a polymerase crystal suggests a mechanism for the prevention of frameshift mutations. *Proc. Natl. Acad. Sci. USA* **100**, 3895–3900

58. Kamtekar, S., Berman, A. J., Wang, J., Lázaro, J. E. M., de Vega, M., Blanco, L., Salas, M., and Steitz, T. A. (2004) Insights into strand displacement and processivity from the crystal structure of the protein-primed DNA polymerase of bacteriophage Φ29. *Mol Cell* **16**, 609–618

59. Kirby, T. W., DeRose, E. F., Cavanaugh, N. A., Beard, W. A., Shock, D. D., Mueller, G. A., Wilson, S. H., and London, R. E. (2012) Metal-induced DNA translocation leads to DNA polymerase conformational activation. *Nucleic Acids Res* **40**, 2974–2983

60. Kuchta R.D., Mizrahi V., Benkovik PA, Johnson K. A., Benkovik S. J., (1987) Kinetic Mechanism of DNA Polymerase I (Klenow Fragment). Biochemistry **26**, 8410-8417

61. Kuchta, R. D., Benkovic, P., and Benkovic, S. J., (1988) Kinetic mechanism whereby DNA polymerase I (Klenow) replicates DNA with high fidelity. Biochemistry **27**, 6716–6725.

62. Kunkel, T. A. (2004) DNA replication fidelity. *J Biol Chem,* **279,** 16895-16898

63. Kunkel, T. A. (2009) Evolving views of DNA replication (in)fidelity. *Cold Spring Harb Symp Quant Biol* **74**, 91–101

64. Lee, H. R., Wang, M., and Konigsberg, W. (2009) The Reopening Rate of the Fingers Domain Is a Determinant of Base Selectivity for RB69 DNA Polymerase. *Biochemistry* **48**, 2087–2098

65. Lieberman, K. R., Cherf, G. M., Doody, M. J., Olasagasti, F., Kolodji, Y., and Akeson, M. (2010) Processive Replication of Single DNA Molecules in a Nanopore Catalyzed by Φ29 DNA Polymerase. *J Am Chem Soc* **132**,

17961–17972

66. Lieberman, K. R., Dahl, J. M., Mai, A. H., Akeson, M., and Wang, H. (2012) Dynamics of the translocation step measured in individual DNA polymerase complexes. *J Am Chem Soc* **134**, 18816 –18823.

67. Lieberman, K. R., Dahl, J. M., Mai, A. H., Cox, A., Akeson, M., and Wang, H. (2013) Kinetic Mechanism of Translocation and dNTP Binding in Individual DNA Polymerase Complexes. *J Am Chem Soc* **135,** 9149 – 9155

68. Lieberman, K. R., Dahl, J. M., and Wang, H. (2014) Kinetic mechanism at the branchpoint between the DNA synthesis and editing pathways in individual DNA polymerase complexes. *J Am Chem Soc* **136**, 7117–7131.

69. Lázaro, J. M., Blanco, L., and Salas, M. (1995) Purification of bacteriophage Φ29 DNA polymerase. *Methods Enzymol* **262***,* 42–49

70. Lazzaro, F., Novarina, D., Amara, F., Watt, D. L., Stone, J. E., Costanzo, V., Burgers, P. M., Kunkel, T. A., Plevani, P., and Muzi-Falconi, M. (2012) RNase H and postreplication repair protect cells from ribonucleotides incorporated in DNA. *Mol Cell* **45**, 99 –110.

71. Lin, T. C., Wang, C. X., Joyce, C. M., and Konigsberg, W. H. (2001) 3"-5" Exonucleolytic activity of DNA polymerases: structural features that allow kinetic discrimination between ribo- and deoxyribonucleotide residues. *Biochem.* **40**, 8749 – 8755.

72. Lujan, S. A., Williams, J. S., Clausen, A. R., Clark, A. B., and Kunkel, T. A. (2013) Ribonucleotides are signals for mismatch repair of leading-strand replication errors. *Mol Cell* **50**, 437–443.

73. Morin, J. A., Cao, F. J., Lázaro, J. M., Arias-Gonzalez, J. R., Valpuesta, J. M., Carrascosa, J. L., Salas, M., and Ibarra, B. (2012) Active DNA unwinding dynamics during processive DNA replication. *Proc. Natl. Acad. Sci. USA* **109**, 8115–8120

74. Morin, J. A., Cao, F. J., Lázaro, J. M., Arias-Gonzales, J. R., Valpuesta, J. M., Carrascosa, J. L., Salas, M., and Ibarra, B. (2015) Mechano-chemical kinetics of DNA replication: identification of the translocation step of a replicative DNA polymerase. Nucleic Acids Research. **43**(7):3643-3652.

75. Miyabe, I., Kunkel, T. A., and Carr, A. M. (2011) The major roles of DNA polymerases epsilon and delta at the eukaryotic replication fork are evolutionarily conserved. *PLoS Genet* **7**, e1002407

76. Nick McElhinny, S. A., Watts, B. E., Kumar, D., Watt, D. L., Lundström, E.B., Burgers, P. M. J., Johansson, E., Chabes, A., and Kunkel, T. A. (2010) Abundant ribonucleotide incorporation into DNA by yeast replicative polymerases. *Proc. Natl. Acad. Sci. USA* **107**, 4949–4954.

77. Nick McElhinny, S. A., Kumar, D., Clark, A. B., Watt, D. L., Watts, B. E., Lundström, E.-B., Johansson, E., Chabes, A., and Kunkel, T. A. (2010) Genome instability due to ribonucleotide incorporation into DNA. *Nat. Chem. Biol.* **6**, 774–781.

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

79. Pelletier, H., Sawaya, M. R., Kumar, A., Wilson, S. H., and Kraut, J. (1994) Structures of ternary complexes of rat DNA polymerase beta, a DNA template-primer, and ddCTP. *Science* **264**, 1891–1903

80. Pursell, Z. F., Isoz, I., Lundström, E.-B., Johansson, E., and Kunkel, T. A. (2007) Yeast DNA polymerase epsilon participates in leading-strand DNA replication. *Science* **317**, 127–130

81. Rashidian M., Doizer J. K., Distfano M. D., (2013) Chemoenzymatic Labeling of Proteins: Techniques and Approaches. Bioconjug Chem, **24**(8): 1277-1294

82. Reha-Krantz, L. J. (2010) DNA polymerase proofreading: Multiple roles maintain genome stability. *Biochim Biophys Acta* **1804**, 1049–1063.

83. Reijns, M. A. M., Rabe, B., Rigby, R. E., Mill, P., Astell, K. R., Lettice, L. A., Boyle, S., Leitch, A., Keighren, M., Kilanowski, F., Devenney, P. S., Sexton, D., Grimes, G., Holt, I. J., Hill, R. E., Taylor, M. S., Lawson, K. A., Dorin, J. R., and Jackson, A. P. (2012) Enzymatic removal of ribonucleotides from DNA is essential for mammalian genome integrity and development. *Cell* **149**, 1008–1022.

84. Rydberg, B., and Game, J. (2002) Excision of misincorporated ribonucleotides in DNA by RNase H (type 2) and FEN-1 in cell-free extracts. *Proc Natl Acad Sci U S A* **99**, 16654–16659.

85. Salas, M., Blanco, L., Lázaro, J. E. M., and de Vega, M. (2008) The bacteriophage Φ29 DNA polymerase. *IUBMB Life* **60**, 82–85

86. Saturno, J., Blanco, L., Salas, M., and Esteban, J. A. (1995) A novel kinetic analysis to calculate nucleotide affinity of proofreading DNA

polymerases. Application to Φ29 DNA polymerase fidelity mutants. *J Biol Chem* **270**, 31235–31243

87. Shamoo, Y., and Steitz, T. A. (1999) Building a replisome from interacting pieces: sliding clamp complexed to a peptide from DNA polymerase and a polymerase editing complex. *Cell* **99**, 155–166.

88. Simon SM, Peskin CS, Oster GF (1992) What drives the translocation of proteins? *Proc Natl Acad Sci USA* **89**: 3770–3774

89. Soengas, M. S., Esteban, J. A., Lázaro, J. M., Bernad, A., Blasco, M. A., Salas, M., and Blanco, L. (1992) Site-directed mutagenesis at the Exo III motif of phi 29 DNA polymerase; overlapping structural domains for the 3′-5′ exonuclease and strand-displacement activities. *Embo J* **11**, 4227–4237.

90. Sparks, J. L., Chon, H., Cerritelli, S. M., Kunkel, T. A., Johansson, E., Crouch, R. J., and Burgers, P. M. (2012) RNase H2-initiated ribonucleotide excision repair. *Mol Cell* **47**, 980–986.

91. Steitz, T. A. (1993) DNA-Dependent and RNA-Dependent Dna-Polymerases. *Curr Opin Struct Biol* **3**, 31–38

92. Steitz, T. A., and Steitz, J. A. (1993) A general two-metal-ion mechanism for catalytic RNA. *Proc Natl Acad Sci U S A* **90**, 6498–6502

93. Swan, M. K., Johnson, R. E., Prakash, L., Prakash, S., and Aggarwal, A. K. (2009) Structural basis of high-fidelity DNA synthesis by yeast DNA polymerase delta. *Nat Struct Mol Biol* **16**, 979–986.

94. Traut, T. W. (1994) Physiological concentrations of purines and pyrimidines. *Mol. Cell. Biochem.* **140**, 1–22

95. Truniger, V., Lázaro, J. M., Salas, M., and Blanco, L. (1996) A DNA binding motif coordinating synthesis and degradation in proofreading DNA polymerases. *Embo J* **15**, 3430–3441

96. Vaisman, A., McDonald, J. P., Noll, S., Huston, D., Loeb, G., Goodman, M. F., and Woodgate, R. (2014) Investigating the mechanisms of ribonucleotide excision repair in Escherichia coli. *Mutat Res Fundam Mol Mech Mutagen* **761**, 21–33

97. Wang C. X., Zakharova, E., Li, J., Joyce, C. M., Wang, J., Koingsberg, W. (2004) Pre-steady state kinetics of RB69 DNA polymerase and its exo

domain mutants: effect of pH and thiophosphoryl linkages on 3'-5' exonuclease activity. *Biochemistry* **43**, 3853-3861

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

99. Wang, M., Lee, H. R., and Konigsberg, W. (2009) Effect of A and B Metal Ion Site Occupancy on Conformational Changes in an RB69 DNA Polymerase Ternary Complex. *Biochemistry* **48**, 2075–2086

100. Wang, M., Xia, S., Blaha, G., Steitz, T. A., Konigsberg, W. H., and Wang, J. (2010) Insights into Base Selectivity from the 1.8 Å Resolution Structure of an RB69 DNA Polymerase Ternary Complex. *Biochemistry* **50**, 581–590

101. Wang, W., Wu, E. Y., Hellinga, H. W., and Beese, L. S. (2012) Structural factors that determine selectivity of a high fidelity DNA polymerase for deoxy-, dideoxy-, and ribonucleotides. *J Biol Chem* **287**, 28215–28226

102. Washington, M. T., Prakash L., and Prakash S. (2001) Yeast DNA polymerase utilizes an induced fit mechanism of nucleotide incorporation. Cell **107**, 917–927

103. Washington, M. T., Johnson R. E., Prakash, L., Prakash, S. (2004) Human DNA Polymerase Utilizes Different Nucleotide Incorporation Mechanisms Dependent upon the Template Base. *Mol Cell Biol*, **24**(2), 936-943

104. Watt, D. L., Johansson, E., Burgers, P. M., and Kunzel, T. A. (2011) Replication of ribonucleotide-containing DNA templates by yeast replicative polymerases. *DNA Repair (Amst)* **10**, 897–902

105. Werneburg, B. G., J. Ahn, X. Zhong, R. J. Hondal, V. S. Kraynov, and M.-D. Tsai. 1996. DNA polymerase : pre-steady-state kinetic analysis and roles of arginine-283 in catalysis and fidelity. *Biochemistry* **35**, 7041–7050

106. Williams, J. S., Clausen, A. R., Nick McElhinny, S. A., Watts, B. E., Johansson, E., and Kunkel, T. A. (2012) Proofreading of ribonucleotides inserted into DNA by yeast DNA polymerase ε. *DNA Repair (Amst)* **11**, 649–656

107. Williams, J. S., and Kunkel, T. A. (2014) Ribonucleotides in DNA: Origins, repair and consequences. *DNA Repair (Amst)*, **19**, 27-37

108. Yang, G., Franklin, M., Li, J., Lin, T. C., and Konigsberg, W. (2002) A conserved Tyr residue is required for sugar selectivity in a Pol alpha DNA polymerase. *Biochem.* **41**, 10256–10261

109. Yang, W., Lee, J. Y., and Nowotny, M. (2006) Making and breaking nucleic acids: two-Mg^{2+}-ion catalysis and substrate specificity. *Mol Cell* **22**, 5–13

110. Yao, N. Y., Schroeder, J. W., Yurieva, O., Simmons, L. A., and O'Donnell, M. E. (2013) Cost of rNTP/dNTP pool imbalance at the replication fork. *PNAS* **110**, 12942–12947

111. Zahurancik, W. J., Klein, S. J., and Suo, Z. (2013) Kinetic mechanism of DNA polymerization catalyzed by human DNA polymerase ε. *Biochemistry* **52**, 7041–7049

112. Zhong, X., S. S. Patel, B. G. Werneburg, and M. D. Tsai. 1997. DNA polymerase : multiple conformational changes in the mechanism of catalysis. *Biochemistry* **36**, 11891–11900