DISSERTATION

STRUCTURAL AND BIOCHEMICAL CHARACTERIZATION OF POLIOVIRUS RNA-DEPENDENT RNA POLYMERASE FIDELITY AND TRANSLOCATION.

Submitted by
Aaron J. Sholders
Department of Biochemistry and Molecular Biology

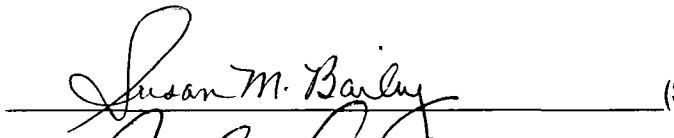
In partial fulfillment of the requirements
for the Degree of Doctor of Philosophy
Colorado State University
Fort Collins, Colorado
Summer 2006

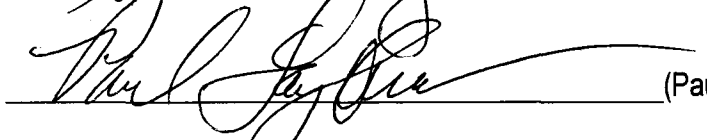
COLORADO STATE UNIVERSITY

June 13, 2006

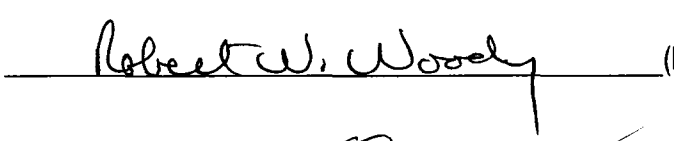
WE HEREBY RECOMMEND THAT THE DISSERTATION PREPARED UNDER OUR SUPERVISION BY AARON J. SHOLDERS ENTITLED "STRUCTURAL AND BIOCHEMICAL CHARACTERIZATION OF POLIOVIRUS RNA-DEPENDENT RNA POLYMERASE FIDELITY AND TRANSLOCATION" BE ACCEPTED AS FULFILLING IN PART REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

Committee on Graduate Work


(Susan Bailey)

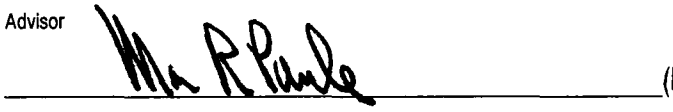

(Paul Laybourn)


(Karolin Luger)


(Robert Woody)


(Olve Peersen)

Advisor


(Marv Paule)

Department Head/Director

ABSTRACT OF DISSERTATION

STRUCTURAL AND BIOCHEMICAL CHARACTERIZATION OF POLIOVIRUS RNA-DEPENDENT RNA POLYMERASE FIDELITY AND TRANSLOCATION.

Picornaviruses are small positive-sense single-stranded RNA viruses that result in a number of pathological conditions including foot-and-mouth disease, heart disease, poliomyelitis, and the common cold. Upon infection the small RNA genome is translated into a single polypeptide that is subsequently cleaved by *cis*-acting proteases into ~12 viral proteins needed to carry out the viral life cycle. The last protein in the polypeptide is an RNA-dependent RNA polymerase whose primary function is genome replication. The N-terminus of the RNA-dependent RNA polymerase of poliovirus ($3D^{pol}$) is specifically buried within a tight pocket at the base of the fingers domain. Our lab has shown that the integrity of this pocket must be maintained for the proper positioning of Asp238, allowing it to form a hydrogen bond with the 2′-OH of the incoming NTP. If proper positioning of Asp238 is not maintained the polymerase will be unable to act as a catalyst for poliovirus genome replication. Recently the Kirkegaard and Andino laboratories isolated a poliovirus strain that was resistant to the effects of the antiviral nucleoside analog drug ribavirin (RTP). This strain contained a single point mutation within $3D^{pol}$, changing glycine 64 to a serine (G64S). Glycine 64 comprises a portion of the N-terminus binding pocket and forms two hydrogen bonds to glycine 1. The structure of the $3D^{pol}$ G64S mutant demonstrates that the loss of conformational flexibility at residue 64 results in the expansion of the exterior portion of the N-terminal binding pocket. Structural alignments of the G64S structure with the native protein also reveal larger structural changes at a helix located in the index finger of the polymerase. In addition to movements within the index finger, the Gly64 mutant exhibited density suggestive of an alternate conformation for a loop consisting of residue 288-292. Biochemical and structural investigation of

these residues indicates that the loop does indeed exist in two separate and stable conformations. Furthermore, the ability of the loop to switch between these two conformations is essential for proper functioning of the polymerase and mutations designed to inhibit loop flexibility lead to translocation defects. These data, in combination with the Gly64 data, are providing us with a better understanding of the molecular events in a complete catalytic cycle of the enzyme.

Aaron J. Sholders
Department of Biochemistry
and Molecular Biology
Colorado State University
Fort Collins, CO 80523
Summer 2006

ACKNOWLEDGEMENTS

I would first like to thank my mentor, Olve Peersen. He has been an excellent resource and a constant source of motivation and creativity in my scientific endeavors. He has fostered a laboratory environment that encourages each individual to pursue understanding in not only science but many academic disciplines. On a more personal note Olve has been extremely understanding of my commitments to my family and my desire to pursue a career as an educator. As such he has allowed me to devote significant portions of time to teaching, taking away from my productivity in the lab and I thank you for the sacrifices you have made for my family and me.

Secondly, I thank present and former Peersen lab members who have been excellent resources and have made the trip just a little easier. I am truly sorry if I ran you over on my protein prep days but there was work to be done! In particular, I would like to thank Aaron Thompson who has been an excellent bench mate for many years. Our many conversations have shaped me and informed me of both my true and false presupposition and I will never forget them, Thank you.

Thirdly, I thank my parents, my sisters, and all my in-laws. My gratitude to you all cannot be expressed in such a short space! You have supported me, you have loved me and you instructed and taught me and you have laughed and cried with me. I love you and cannot thank you enough.

Fourthly, I thank Isaiah and Ian. You have been and continue to be such an enormous blessing and such source of enjoyment to me. I could not think of a better way to forget the stresses of each day by coming home and seeing the joyful look on your faces simply because I have walked through the door. May I bless you as you have blessed me.

Fifthly, thank you to my wife, my best-friend and companion. In this life it is a very rare thing to find a soul-mate. Someone who will walk with you through thick and thin, who will love you and sacrifice themselves for your betterment. This describes my wife who has laughed with, cried with me and has shared in the stresses of each and everyday. I could not have dreamed of a better mate, you are a gift from God who knew that this could not be accomplished without your support and encouragement throughout! Becky, I cannot express to you what you mean to me, I can only say thank you from the bottom of my heart. We made it!

Finally, I must thank my Lord and Savior, Jesus Christ to whom all honor and glory must be given. Your grace has made this possible.

DEDICATION

I dedicate this dissertation and my life's work to my blessed Lord and Savior, Jesus Christ. "For by Him all things were created that are in heaven and that are on earth, visible and invisible whether thrones or dominions or principalities or powers. All things were created through Him and for Him. And He is before all things and in Him all things consist." (Col. 1:16-17)

I love you and am deeply honored that you have chosen to glorify Yourself through my life. Thank you for all things, my life is for You.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter I

Introduction

Poliovirus lifecycle

Epidemiology

In 1988 the World Health Organization put forth resolution 41.28 describing an international effort to eradicate poliovirus. Since that time the number of polio cases reported world-wide has dropped 99% and endemic wild poliovirus has been eliminated from all but four countries: Nigeria, India, Pakistan and Afghanistan (World Health Organization, 2005). Despite these promising numbers, the risk of a polio viral out break remains a distinct possibly in large part due to the use of the oral polio vaccine (OPV) that has resulted in a number of vaccine-derived poliovirus cases in Hispaniola, the Philippines, Madagascar and China (World Health Organization, 2005). These cases can be attributed to rapid viral evolution, resulting in reversion of two of the three mutations introduced to attenuate the live virus used in the vaccine. These reversions result in a vaccine-derived virus every bit as infective as the wild-type virus, leading to paralysis or death (Abraham et al., 1993). Currently it is estimated that there is a 65-90% chance of a vaccine-derived polio viral outbreak within the first year following OPV cessation (World Health Organization, 2005). Given this data, it is safe to say that research leading to drug development for poliovirus continues to be a relevant endeavor.

Poliovirus Cell-entry, Translation and Proteolytic Cleavage

Poliovirus is an enterovirus and a member of the *Picornaviridae* family of viruses whose hallmark feature is a small single-stranded positive-sense RNA genome. Members of this family of viruses include the heart-disease causing coxsackievirus, the aphthoviruses that cause a number of animal diseases such as foot-and-mouth disease, and the rhinoviruses that are the cause of the common cold.

Polio viral infection is mediated through an interaction between the virion, made of structural proteins VP1, VP2, VP3 and VP4, and an immunoglobulin cell surface receptor, Pvr/CD155. Current models of genomic cell entry suggest the formation of a pore through multiple interactions with cell surface receptors resulting in translocation of the small RNA genome into the host cell (Figure 1.1) (Hogle, 2002).

The polio viral genome is 7,441 nucleotides in length containing several complicated RNA structures found throughout the genome (Figure 1.2). Of particular interest are the 5′-cloverleaf, the 3′ stem-loop, the internal ribosome-entry site (IRES), and the CRE that are all crucial for various aspects of translation and replication of the genome (Barton et al., 2001; Mirmomeni et al., 1997; Paul et al., 2000; Pelletier and Sonenberg, 1988). Once inside the cell the genome is translated in a cap-independent manner by host ribosomes that recognize the IRES found in the 5′ UTR (Pelletier and Sonenberg, 1988). Translation of the polio viral genome results in a 247-kDa protein that is cleaved co- and post-translationally by viral proteases (Krausslich et al., 1988; Lawson and Semler, 1990), producing 11 viral proteins that fall into three broad categories: P1, P2 and P3 (Figure 1.1 and Figure 1.2). The P1 region consists of structural peptides that form the viral capsid and are important for cell surface recognition and viral infectivity. The P2 region contains proteins that shut down host protein synthesis and wreak havoc on the internal environment of the cell, destroying membranous compartments and reorganizing them into

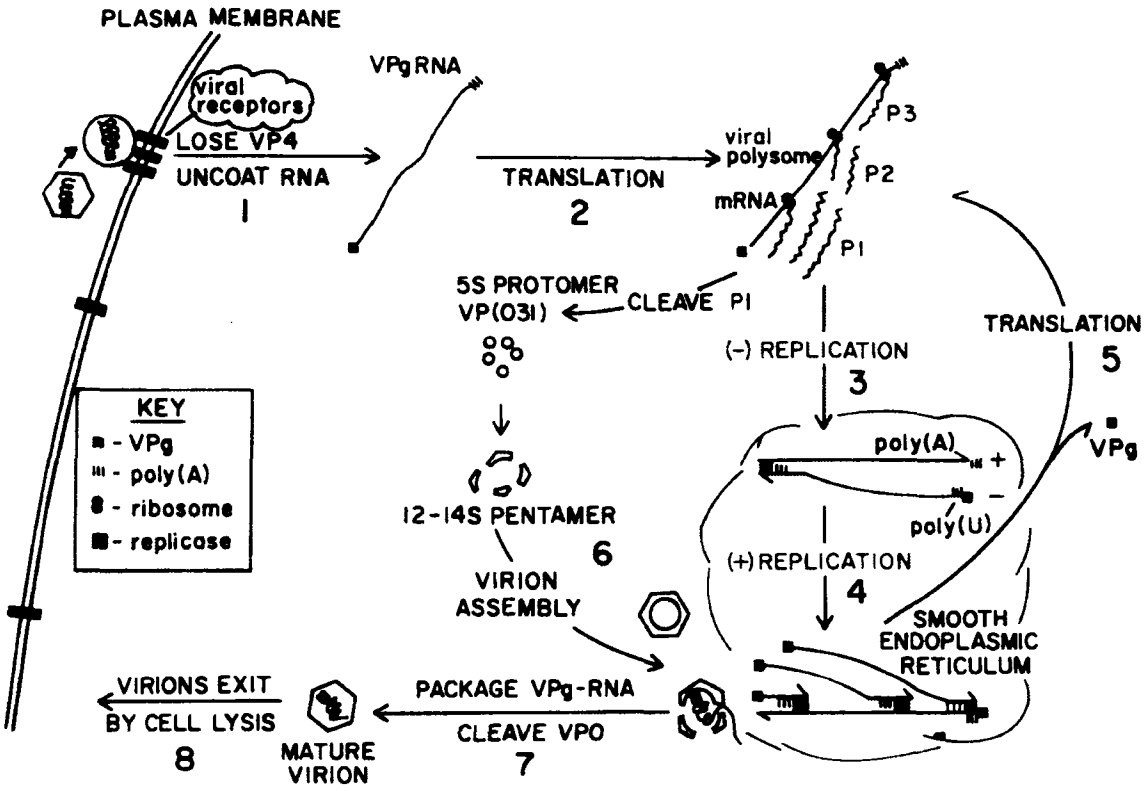


Figure 1.1: The *Picornaviridae* life cycle. The life cycle of poliovirus begins with host-cell attachment and RNA viral genome insertion (1). Once in the cell the small RNA genome is translated (2) making the viral proteins necessary to carry out the rest of the life-cycle. Translation is followed by rounds of minus-strand (3) and positive-strand synthesis (4), both of which occur in large membrane-bound replication centers (orange). The newly synthesized positive strands are then packaged with viral capsid proteins (7), resulting in mature virions that then exit the cell (8).

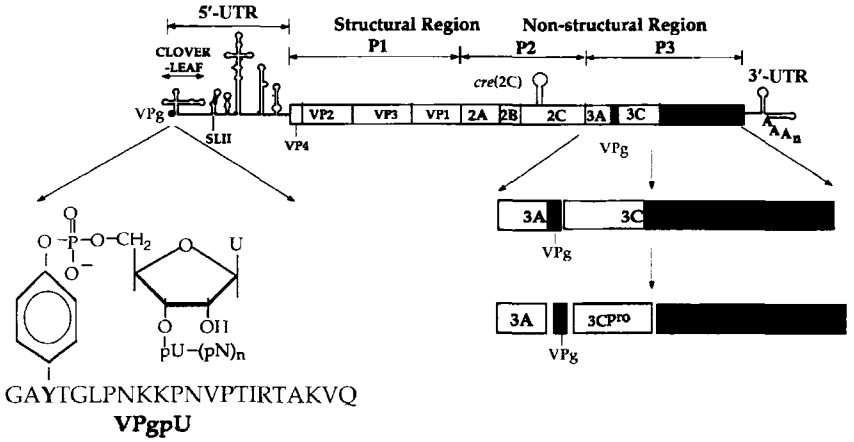


Figure 1.2: The Poliovirus Genome. RNA cloverleaf and hairpin structures, important for both replication and translation, are shown in the 5′ and 3′ -UTR's as well as CRE, a single RNA hairpin in the open reading frame (colored green) which is required for viral growth and propagation. The small 22-amino acid peptide, VPg or 3B (light blue), is shown attached to the 5′ -UTR and in the P3 coding region of the open reading frame. The opening reading frame is divided into 11 boxes each of which correspond to a separate viral protein. The last four proteins of the genome are involved in genome replication and include the RNA-Dependent RNA polymerase, shown in red, which is cleaved by viral protease from its precursor protein 3CD.

vesicles on which viral replication centers will assemble (Leong et al., 2002). Finally, the P3 region consists of several proteins involved in the replication of the viral genome (Figure 1.2). The P3 region is initially cleaved into two precursor proteins, 3AB and 3CD. 3CD is a protease and an RNA-binding protein that has been shown to interact with the 5′-cloverleaf, the 3′ stem-loop and the CRE (Gamarnik and Andino, 2000; Harris et al., 1994; Pathak et al., 2002). This protein is a "master regulator" of the polio viral genome, controlling both replication and translation, and is thought to lead to the circularization of the viral genome (Andino et al., 1993; Murray and Barton, 2003). 3CD is subsequently cleaved into $3C^{pro}$, a protease, and $3D^{pol}$, an RNA-dependent RNA polymerase (RdRP) responsible for the complete replication of the viral genome. Interestingly, 3CD is completely dead as a polymerase and has remarkably different specificities as a protease when compared to 3C (Flanegan and Van Dyke, 1979; Parsley et al., 1999; Ypma-Wong et al., 1988). 3AB cleavage results in 3B, also called VPg, a 22-amino acid peptide responsible for priming positive-strand RNA synthesis. 3B is found attached to every negative- and positive-sense RNA strand and must be cleaved from the genome by host proteases before translation can occur (Figure 1.2) (Ambros and Baltimore, 1980; Ambros et al., 1978).

Poliovirus Replication

Following polio viral protein synthesis and proteolytic cleavage, the RNA genome goes through replication and assembly at viral replication centers (Figure 1.1). Replication center are large membrane-bound vesicles made from the endoplasmic reticulum that contain viral proteins $3D^{pol}$, 3CD, 3AB and 2C. In 2003 the Barton lab proposed a model for the replication cycle whereby 3CD acts to circularize the genome through an ill-defined RNA-protein-protein-RNA interaction with the 5′-cloverleaf and the 3′ stem-loop found in the UTR's (Figure 1.3). 3CD then cleaves both itself

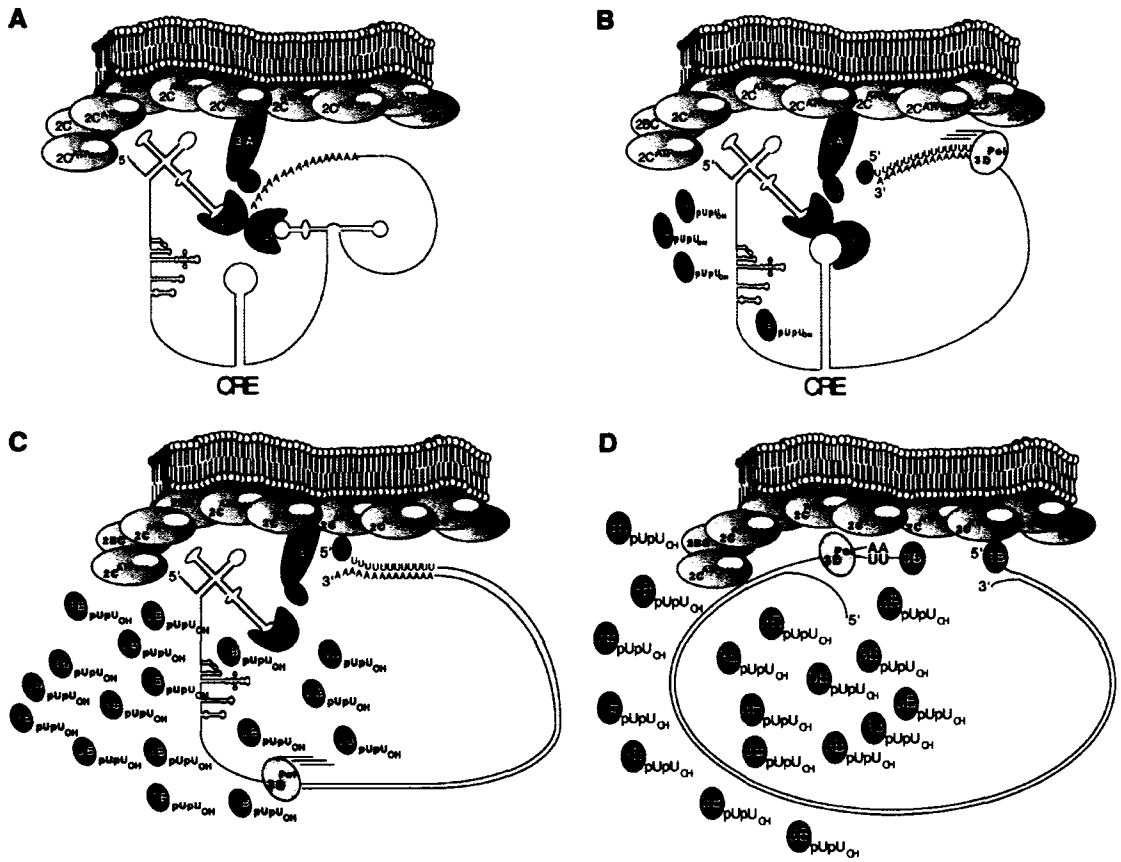


Figure 1.3: A Model of the Poliovirus Replication Cycle. The positive-sense genome is associated with membranes in the replication centers and negative-strand synthesis initiation occurs at the poly-A tail (A) followed by 3CD association with the CRE resulting in VPg uridylylation (B). VPg-pUpU production is terminated following the disruption of the 3CD-CRE complex by $3D^{pol}$, which is carrying out negative-strand synthesis (C). Once negative-strand synthesis is terminated the VPg-pUpU molecules are used to prime positive-strand synthesis using the newly generated negative-strand as a template (D) (figure from Murray and Barton, 2003) .

and 3AB, which has been found in a ribonucleoprotein complex at the 5′ UTR (Andino et al., 1993; Andino et al., 1990), releasing VPg and the RdRP $3D^{pol}$ at the poly-A tail. VPg acts to prime negative-strand synthesis, which begins at the poly-A tail working its way through the genome to the 5′ end.

In this model, initiation of negative-strand synthesis results in the disruption of the ribonucleoprotein complex at the 3′-UTR, allowing the 5′-UTR-bound 3CD to interact with a third 3CD bound to the CRE. CRE is a *cis*-acting RNA element found in the 2C coding region of the genome that is crucial for viral growth by acting as a template for VPg uridylylation (Goodfellow et al., 2000; Rieder et al., 2000). VPg uridylylation is a process by which $3D^{pol}$ adds two uridine residues to the hydroxyl of Tyr3, forming VPg-pUpU, which then acts as a primer for positive-strand RNA synthesis (Pathak et al., 2002). The formation of the ribonucleoprotein complex at the CRE will stimulate VPg uridylylation, producing a number of positive-strand RNA primers for the ensuing round of positive-strand synthesis (Paul et al., 2000; Rieder et al., 2000). Indeed, Murray and Barton demonstrate that ~500 VPg-pUpU molecules can be produced in the time expected to elapse for the completion of a single-round of negative-strand synthesis (Murray and Barton, 2003). Once negative-strand synthesis has been completed, positive-strand synthesis begins using the VPg-pUpU molecules as primers, resulting in a large number of positive-sense RNA genomes that are subsequently packaged into viral capsids that exit the cell (Murray and Barton, 2003). A hallmark feature of poliovirus infection is the asymmetry of replication where positive-strands outnumber the negative-strands by ~50:1 (Novak and Kirkegaard, 1991). The model presented by the Barton lab provides a molecular explanation for the asymmetry of replication and is consistent with a large body of data. However, further biochemical and structural investigation of RNA-protein-protein-RNA complexes needs to be done for fuller understanding of viral replication

and support of Barton's model. The Peersen lab has undertaken a structural and biochemical investigation of $3D^{pol}$, and other viral RdRP.

The Poliovirus RNA-Dependent RNA Polymerase

RNA-Dependent RNA Polymerases

To date, the structures of at least nine different RNA-Dependent RNA polymerases (RdRP's) have been solved, including the full-length structure of the RdRP from poliovirus (Thompson and Peersen, 2004). Comparison of these polymerases reveals striking similarities both structurally and biochemically. Each polymerase structure follows the right-hand analogy first used by Tom Steitz to describe the structure of the DNA polymerase I Klenow fragment from *Escherichia coli* (Ollis et al., 1985). The analogy compares the polymerase with the anatomical features of a cupped right-hand, dividing the protein into fingers, palm and thumb structural domains (Figure 1.4).

The active site is contained within the palm domain where two Asp residues act to coordinate the magnesium cations necessary for phosphoryl transfer. The active site, unlike those of larger polymerases, is said to be enclosed due to an interaction between the tip of the fingers domain and the top of the thumb. Many in the field hypothesize that this interaction likely prohibits large structural rearrangements that have been observed in other classes of polymerases, in particular the DNA-dependent RNA polymerase of bacteriophage T7 that has structural elements which move up to 70 Å when switching from the initiation to the elongation complex (Yin and Steitz, 2002). The RNA/polymerase complexes from foot-and-mouth-disease virus (FMDV) (Ferrer-Orta et al., 2004), bacteriophage ϕ6 (Butcher et al., 2001; Salgado et al., 2004) and hepatitis C virus (HCV) (O'Farrell et al., 2003) all reveal that the fingers domain forms a structural platform for RNA

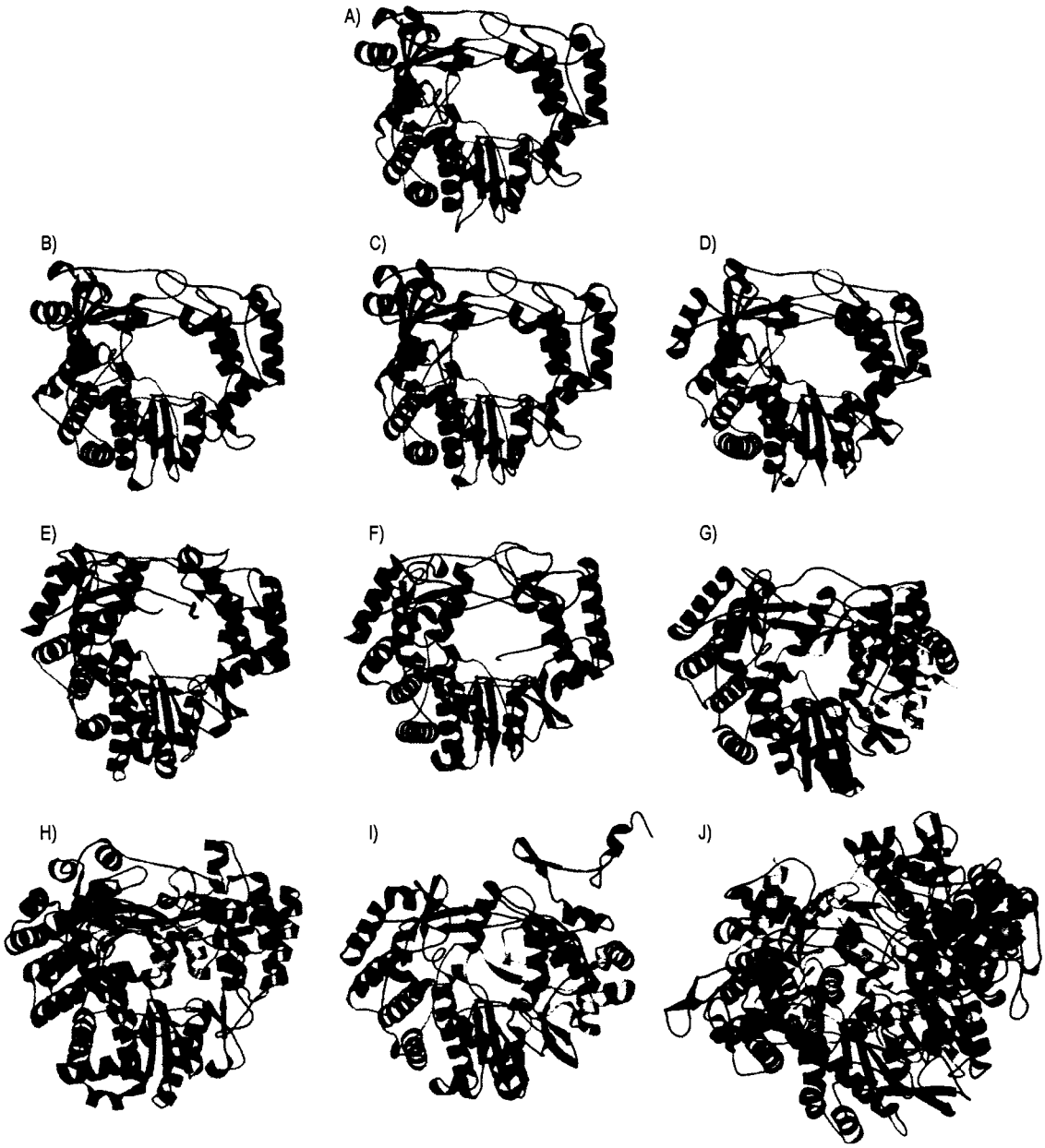


Figure 1.4: Structures of 10 Different RNA-Dependent RNA Polymerases. Each polymerase follows the right-hand analogy with thumb (blue), palm (gray) and finger (red) domains. An emerging feature of RNA-dependent RNA polymerases is an enclosed active site (magenta) due to an interaction between the fingers and thumb domain. Polymerases capable of primer-independent initiation, **(G)** HCV (Ago et al., 1999 (1QUV); Bressanelli et al., 1999 (1CSJ); Lesburg et al., 1999 (1C2P); O'Farrell et al., 2003 (1NB6)), **(H)** ϕ6 (Butcher and Grimes et al., 2001 (1HHT)),**(I)** BVDV (Choi et al., 2004 (1S48)) and **(J)** reovirus (Tao et al., 2002 (1MUK)), contain a large C-terminal domain referred to as the priming platform (yellow). Structures lacking the priming platform; **(A)** poliovirus (Thompson and Peersen, 2004 (1RA6)),**(B)** coxsackie virus (Campagnola et al., in press), **(C)** rhinovirus (Love et al., 2004 (1XR5)), **(D)** FMDV (Ferrer-Orta et al., 2004 (1WNE)), **(E)** rabbit hemorrhagic disease virus (Ng et al., 2002 (1KHV)), and **(F)** Norwalk virus (Ng et al., 2004 (1SH2)), require a protein or nucleotide primer for replication initiation (figure from Thompson, 2006).

template binding and that the interaction between the fingers and thumb forms a positively charged NTP entry tunnel.

In addition to the three structural motifs of the right-hand analogy, RdRPs from HCV (Ago et al., 1999; Bressanelli et al., 1999; Lesburg et al., 1999; O'Farrell et al., 2003), bovine viral diarrhea virus (BVDV) (Choi et al., 2004), bacteriophage ϕ6 (Butcher et al., 2001; Salgado et al., 2004) and reovirus (Tao et al., 2002) all contain a large C-terminal domain called the priming platform (Figure 1.4). The priming platform gives the structural basis for *de novo* initiation of RNA synthesis by providing a Tyr or Trp that interacts with the first base via a π-stacking interaction in the priming site. Recruitment of the first base to the priming site provides the 3′-OH necessary for phosphodiester bond formation with the second base bound to the catalytic site, resulting in primer independent initiation of RNA synthesis (Butcher et al., 2001).

As discussed above, $3D^{pol}$ and other RdRP's of the *Picornaviridae* family of viruses require a small protein primer called VPg to initiate either negative-strand or positive-strand synthesis. Recently the Verdaguer lab solved a structure of the RdRP from FMDV in complex with VPg (Ferrer-Orta et al., 2006). This structure demonstrates that the C-terminal end of VPg interacts with the RNA-binding groove similar to the RNA. The N-terminal end of VPg is bound in the NTP-entry tunnel and Tyr3, which is uridylyated, is located directly above the active site of the enzyme (Figure 1.5). Structural alignments of the RdRP's from rhinovirus and poliovirus show very high homology, suggesting that all picornaviral polymerase can bind VPg in this fashion (Ferrer-Orta et al., 2006).

The Structure of $3D^{pol}$ and the N-terminus Binding Pocket

The structure of the RdRP from poliovirus, $3D^{pol}$, demonstrates many of the features already discussed for the RdRP's in general. Furthermore, the picornaviral fingers domains can be further

Figure 1.5: *Picornaviridae* RNA-dependent RNA Polymerase VPg Binding. A) Cartoon model of the RdRP from FMDV following uridylylation of VPg. The VPg molecule has a UMP covalently bound to Tyr3, which is situation directly over the active site (magenta) and catalytic Mg^{2+} cations (red spheres). The thumb is shown in blue, the fingers shown in red and the palm is shown in gray. **B)** Surface representation of VPg-pU bound FMDV RdRP. A large portion of the fingers domain has been removed to reveal the VPg molecule's interaction with the palm (Ferrer-Orta et al., 2006 (2F8E)).

divided by analogy to anatomical features of a primate hand consisting of an index finger (residue 1-68), a middle finger (residues 269-285), a ring finger (residues 150-179) and a pinky finger (residues 96-146 and 180-190) (Figure 1.6). The active site, found in the palm domain, is enclosed due to a contact between index finger and the thumb where Phe 30 and 34 are buried in a hydrophobic pocket at the tip of the thumb (Figure 1.6) (Thompson and Peersen, 2004).

An important structural feature of this polymerase is the burial of the N-terminus in an intricate pocket at the base of the fingers domain. This pocket provides a number of hydrogen bonds to the amino group of Gly-1 which help anchor the N-terminus in place (Figure 1.7). Recently our lab has shown that the proper burial of the N-terminus plays an essential role in the catalytic ability of the enzyme. Replacement of Gly1 with an Ala or a Ser leads to 2- and 50-fold reductions in elongation activity, respectively. Similarly, the removal or addition of a single amino acid at the N-terminus completely abolishes elongation activity. The decrease in activity for these mutants has been attributed to an allosteric effect whereby the burial of the N-terminus correctly positions Asp238 relative to the active site to allow it to form a hydrogen bond to the 2′-OH of the incoming NTP (Thompson and Peersen, 2004). Interestingly, the burial of the N-terminus appears to be a common feature among picornaviral RdRP's, suggesting that it plays a fundamental role in the activation of all these enzymes (Figure 1.7).

While this data has given a structural basis for the observed proteolytic activation of $3D^{pol}$ upon cleavage of the precursor protein 3CD, it gives only limited insight into the molecular details of catalysis. Based upon data from a number of kinetic studies performed on the various classes of polymerases, including poliovirus $3D^{pol}$ (Arnold and Cameron, 2004), it is expected that for catalysis to occur the RdRP will have to undergo significant allosteric rearrangements. Indeed, the structure of $3D^{pol}$ bound to GTP indicates that the NTP initially binds in a site that is unsuitable

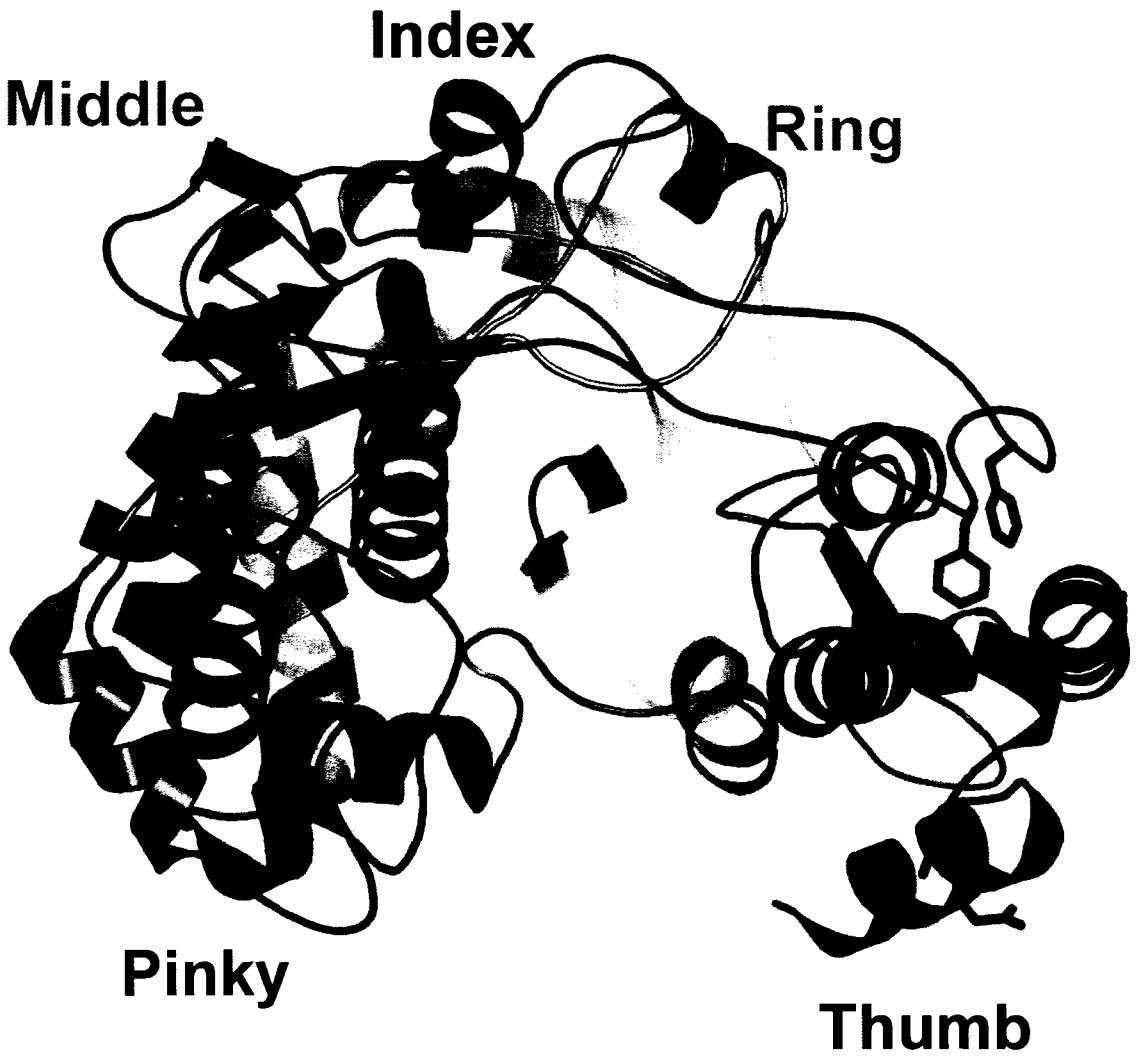


Figure 1.6: Top View of the Poliovirus RNA-Dependent RNA Polymerase. The fingers domain of the RdRP from poliovirus has been divided into 4 separate structural motifs: the pinky finger (red), the ring finger (yellow), the middle finger (orange) and the index finger (green). This coloring scheme will be the common depiction of $3D^{pol}$ throughout the thesis. The index finger reaches across the polymerase to insert Phe 30 and 34 into a hydrophobic pocket at the top of the thumb (blue) enclosing the active site (magenta). The N-terminus is shown as a blue sphere that is buried in an intricate pocket at the base of the fingers domain (figure from Thompson, 2006) (Thompson and Peersen, 2004 (1RA6)).

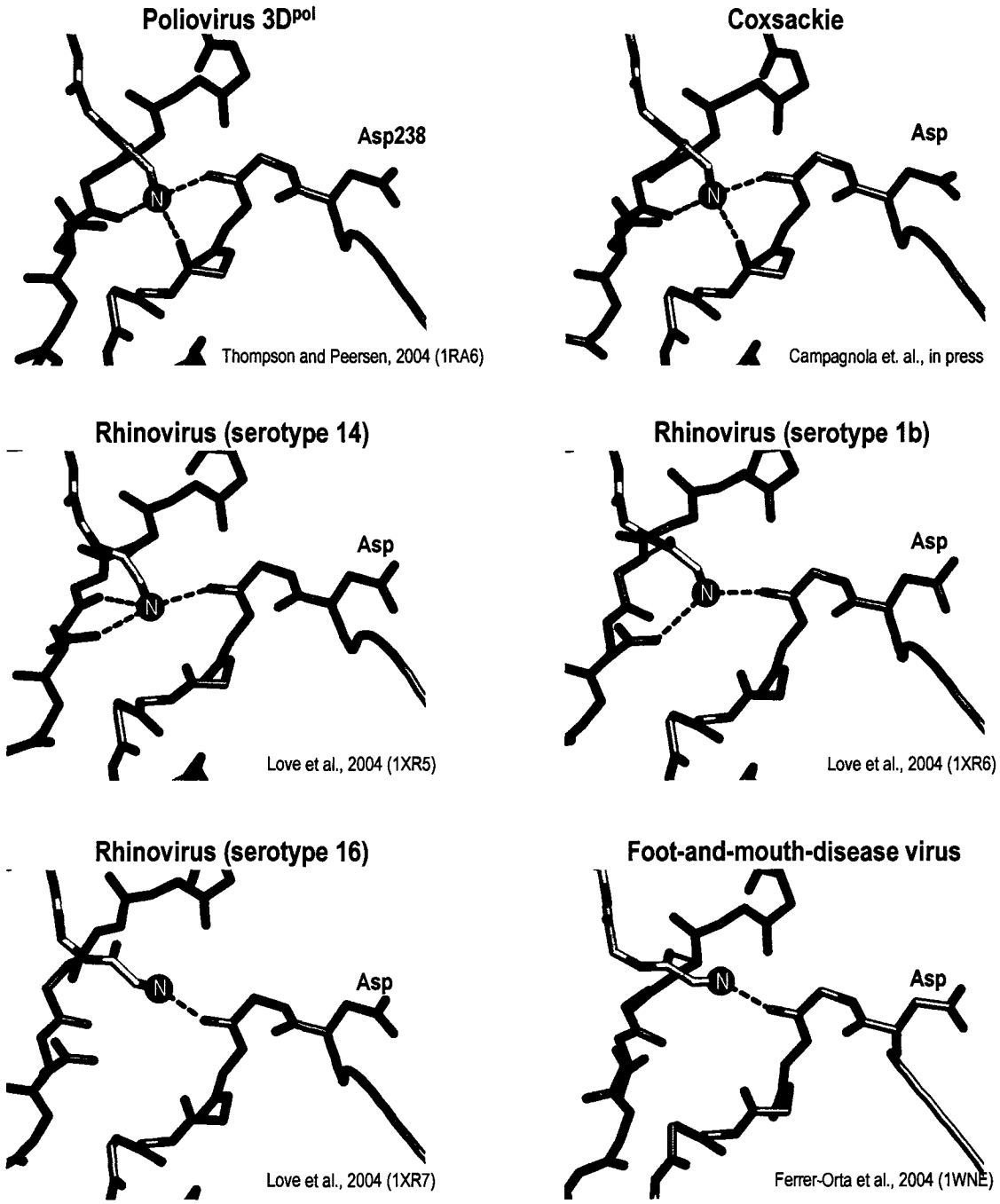


Figure 1.7: The Buried N-terminus of *Picornaviridae* RNA-Dependent RNA Polymerases. The N-terminus (blue sphere) from each *Picornaviridae* RdRP is shown buried in an intricate pocket at the base of the fingers domain. Each N-terminus is hydrogen bonded to the backbones and side chains of residues found in the finger (green) and the palm (gray) domains. Asp238 analogs are labeled for each structure. The Peersen lab has shown how the proper burial of the N-terminus in poliovirus is crucial for the activity of the polymerase due to correct positioning of Asp238 with respects to the active site (Thompson and Peersen, 2004).

for base pairing with the template strand and phosphoryl transfer (Thompson and Peersen, 2004). This structure indicates that at the very least the protein must undergo a conformational change that places the NTP in a site suitable for catalysis. This type of conformational change is yet to be observed in structural studies of any RdRP.

Description of this Dissertation

The work described in this dissertation is aimed at producing a better understanding of the molecular details of catalysis for the poliovirus RdRP. These investigations include the structural and biochemical analysis of the N-terminus binding pocket, specifically addressing the role of Gly64. Interestingly, mutation of this residue to a serine results in poliovirus strains that exhibit resistance to the antiviral effects of ribavirin (Pfeiffer and Kirkegaard, 2003; Vignuzzi et al., 2006). Results described in this dissertation have contributed to a greater understanding of the dynamic interrelationship between the N-terminus and possible structural rearrangements within the fingers domain.

Investigation of N-terminus-binding pocket mutants led to the discovery of dynamics within a separate portion of the enzyme centered at a loop comprised of residue 288-292 in the palm domain. Structural and biochemical analysis of this loop indicate that these residues exist in two stable conformations and that the ability of the loop to switch between these conformations is required for the translocation of the template/product strand following phosphoryl transfer. These results, in combination with results obtained from N-terminus-binding pocket mutants, are providing a better understanding of the molecular details of catalysis for all RdRP's.

Chapter II

Structure of the G64S Ribavirin Resistant Poliovirus Shows an Altered Binding Pocket for the Buried N-terminus.

Introduction

Viral RNA-dependent RNA polymerases are small single-subunit proteins exhibiting rather high mutation rates due to the lack of a proofreading mechanism seen in larger polymerases (Holland et al., 1982; Steinhauer et al., 1992). This results in a selective advantage commonly seen among RNA viruses whereby the viral genome does not exist as a single genotype, but rather as a distribution of genetically related sequences collectively called a quasi-species (Eigen et al., 1988; Holland et al., 1982; Steinhauer et al., 1989). This genomic diversity supplies a collection of potentially advantageous mutations that allow the virus to adapt to different cellular environments and escape host immune responses or clinical intervention (Domingo et al., 1997). The high mutation rate and existence of this pool of quasi-species continues to frustrate efforts to develop effective therapeutics for this class of viruses. For example, the attenuated poliovirus currently being used in the oral vaccine has been found to revert two of its three point mutations in a matter of days to produce a virus that is essentially as infective as wild-type (Abraham et al., 1993). This observation offers a vivid picture of the effect that the mutation rate has on a viral population and the establishment of quasi-species, and presents a major obstacle for the World Health Organization effort's to eliminate poliovirus from the planet.

While genetic diversity and high mutation rates provide selective advantages, they are also a danger to the virus in that small increases in the mutation rate can lead to the production of non-viable genomes. Indeed, Crotty et al. (2000) were able to show that ribavirin, a nucleoside analog with broad-spectrum antiviral activity, acts as a mutagen that is incorporated into the viral RNA genome opposite of cytosine or uracil bases, increasing the mutation frequency and subsequently pushing cultured virus into "error catastrophe". The authors were able to show that poliovirus cultured in the presence of high concentrations of ribavirin displayed a 10-fold increase in mutation frequency accompanied by a 2000-fold reduction in viral production and a 140-fold reduction in viable genomes (Crotty et al., 2001; Crotty et al., 2000). These results give direct molecular evidence linking mutation frequency with genomic viability and demonstrate ribavirin's ability to act as a lethal mutagen. Recently, two separate labs were able to generate a strain of poliovirus that was less sensitive to the mutagenic effects of ribavirin. Interestingly, both studies showed that this resistance was conferred by a single point mutation within the RNA-dependent RNA polymerase ($3D^{pol}$) coding region that changed glycine 64 to a serine (G64S) (Pfeiffer and Kirkegaard, 2003; Vignuzzi et al., 2006). The G64S mutant $3D^{pol}$ displays a greater fidelity, resulting in an average of ~0.3 mutations per genome per generation as compared to ~1.9 for wild-type. These mutant viruses were more sensitive to other antiviral drugs, such as the capsid-binding WIN compounds, due to their inability to rapidly evolve resistance to them (Arnold et al., 2005). Furthermore, G64S mutant-containing viruses also showed reduced tissue tropism and altered pathogenesis in infected mice, resulting from a more homogeneous quasi-species population, providing a direct link between mutation rate, genomic diversity and viral fitness and pathogenesis (Pfeiffer and Kirkegaard, 2005; Vignuzzi et al., 2006).

The structure of poliovirus RNA-dependent RNA polymerase ($3D^{pol}$) provides insights into a possible mechanism of action by the G64S mutation. This residue forms a portion of an intricate

pocket located at the base of the fingers domain in which the very N-terminus must be buried to activate the enzyme. The pocket provides key hydrogen bonds between the N-terminus and residues of the palm domain which in turn correctly position Asp238 relative to the active so that it can form a very important hydrogen bond with the 2′-OH of incoming NTP (Figure 2.1). This interaction is necessary for catalysis in that changing Asp238 to an alanine completely abolishes activity while maintaining backbone geometry identical to native $3D^{pol}$ (Thompson and Peersen, 2004).

Gly64 forms the exterior portion of the N-terminus-binding pocket and has two reciprocal hydrogen bonds linking the backbone carbonyls and amide groups of Gly64 and Gly1. These interactions require unusual backbone geometry and Gly64 is therefore in the disallowed region of the Ramachandran plot. The interior portion of the pocket is formed by two additional hydrogen-bonds between the N-terminus and the carbonyl oxygen of residues 239 and 241 (Figure 2.1). Proper burial of the N-terminus in this pocket is essential for enzymatic activity in that the addition or removal of a single amino acid to the N-terminus will completely abolish activity while mutations of Gly1 will result in marked decreases in activity (Thompson and Peersen, 2004). For example, substitution of Gly1 with an alanine results in a 2-fold reduction in polymerase activity as measured by an oligo-dT/polyA extension assay (Appendix 1). This change in activity has been attributed to a ~0.9Å displacement of the N-terminus resulting in an expansion of the N-terminus-binding pocket, the loss of a hydrogen bond between the N-terminus and residue 241, and an accompanying movement of Asp238 away from the active site (Thompson and Peersen, 2004).

It was our hypothesis that the G64S mutation disrupted the intricate N-terminus-binding pocket by replacing Gly64 with a residue unlikely to adopt the constrained backbone torsion angles observed in the native structure, thus resulting in a decrease in the activity and overall rate of this

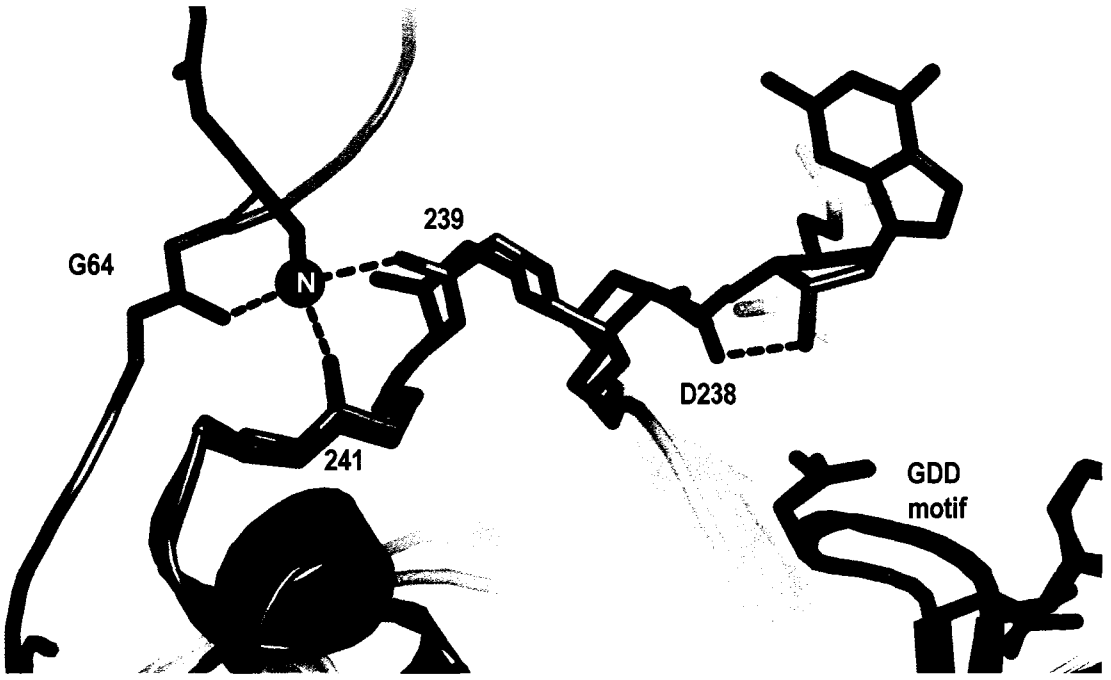


Figure 2.1: Asp238 moves 1.4 Å away from the active site in absence of an N-terminus. Superposition of full-length $3D^{pol}$ (gray and green) (Thompson and Peersen, 2004 (1RA6)) with $3D^{pol}$-Δ68 (pink) (Thompson and Peersen, 2004 (1RDR)), a construct missing the first 68 residues of the polymerase, demonstrates that the burial of the N-terminus (blue sphere) results in a movement of Asp238 towards the active site (magenta) allowing it to form a hydrogen bond with the 2′-OH of the incoming NTP. The N-terminus-binding pocket is made of residue 64, which forms two hydrogen bonds with the N-terminus, and residues 239 and 241. The active site is comprised of two aspartic acid residues that coordinate catalytic Mg^{2+} cations and is often referred to as the GDD motif.

enzyme. Here I present the crystal structures of native $3D^{pol}$ and $3D^{pol}$ G64S in the presence of ribavirin triphosphate (RTP). These structures demonstrate that RTP binds $3D^{pol}$ in the same site as canonical NTPs, supporting its role as a nucleotide analog. The structure of the G64S mutant displays not only a local change about the N-terminus-binding pocket, but also a shift in the entire α-helix comprising residues 52-62. Furthermore, the binding of RTP to the G64S mutant causes a shift in this α-helix back toward its native position. We also wanted to examine if there was something unique about the replacement of glycine with serine at residue 64 and therefore generated an alanine mutation of this residue (G64A). This structure is superimposable with the G64S structure and also results in a 4-fold decrease in activity, lending support to the view that the serine side-chain does not play a unique role from a structural standpoint. The combination of our results suggests a dynamic interplay between the N-terminus, its binding pocket, the surrounding structural elements, and the binding of NTP's all coming together to modulate polymerase fidelity and activity.

Results

Crystals and x-rays structures were obtained as reported by Thompson and Peersen (2004) with slight modifications to cyro-protectant and soaking procedures. In order to obtain base and ribose density for ribavirin tri-phosphate, the crystals were slowly transferred to a solution containing 250mM sodium acetate and 30% PEG400, as compared to 1.8M sodium acetate and 30% glycerol used in prior work (Thompson and Peersen, 2004). These lower ionic strength conditions lead to better overall density for the nucleoside analog and seemed to slightly improve resolution up to a typical limit of 2.1Å. This effect on resolution may have resulted from the binding of a PEG molecule at a crystal packing contact, leading to stabilization of the contacts and an overall increase in the order of crystal packing (Figure 2.2). Structures were refined with resolution

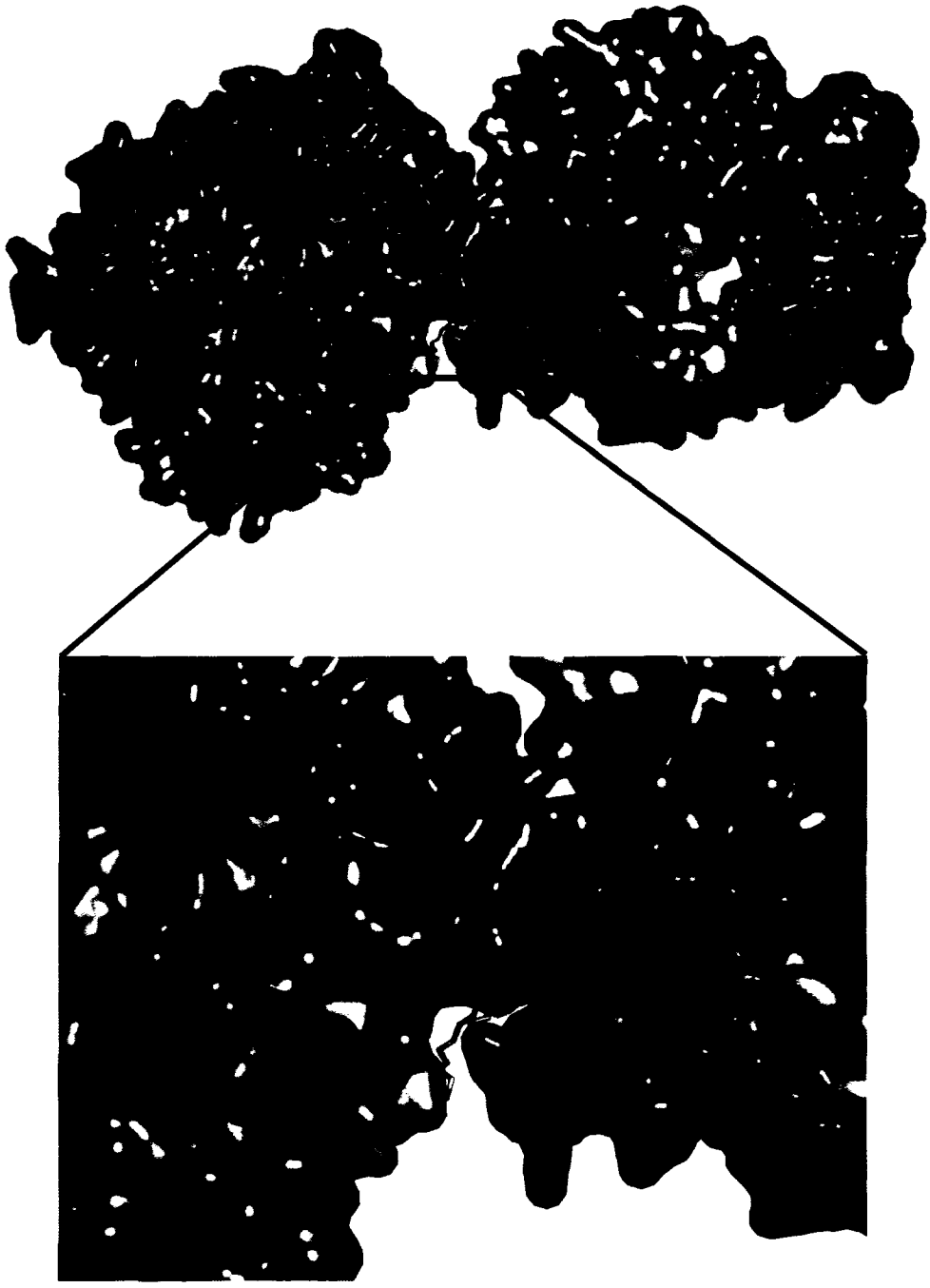


Figure 2.2: Poly (ethylene glycol) (PEG) binding at a crystal contact. Surface representation of two symmetry related molecules with a PEG molecule (yellow and red sticks) bound at their interface. The PEG molecule was not observed in the original crystal lattice and may be contributing to the small increase in resolution observed in the Gly64 mutant crystals as compared to the native crystals (Thompson and Peersen, 2004).

ranging from 2.1-2.6 Å and R and R_{free} values ranging from 21-24% and 23-26% respectively (Appendix 2, Tables A2.1 and A2.2).

The overall structures of the mutant polymerases do not significantly change as compared to the native $3D^{pol}$ structure. All major structural domains and folds are maintained throughout and changes resulting from each mutation are limited to small movements in the N-terminus binding pocket and the surrounding structural elements.

$3D^{pol}$ G64S Structure and Biochemistry

Immediate observation of the G64S mutant structure reveals a significant expansion of the exterior portion of the N-terminus-binding pocket that results in the loss of hydrogen bonds from the N-terminus to both the amide nitrogen and carbonyl oxygen of residue 64 (Figure 2.3). This expansion is not unexpected due to the improbability of finding a serine in the disallowed region of the Ramachandran plot. Despite these perturbations, the N-terminus displays only a slight displacement and maintains hydrogen bonds with the backbones of residue 239 and 241. It is also evident that no major structural change takes place in residues comprising the active site and that the position of Asp238 has not been greatly affected despite the 4-fold decrease in activity as measured by an oligo-dT/polyA extension assay (Table 2.1).

In light of these observations, we set out to compare this structure with an ensemble of $3D^{pol}$ structures solved by our lab using a maximum likelihood-based structural alignment algorithm developed by Doug Theobald at the University of Colorado and implemented in his program Theseus (Theobald and Wuttke, 2006). These alignments brought to light the fact that while mutations at Gly64 certainly effect the local structure surrounding the N-terminus itself, they also have a more global effect reaching into the fingers domain of the polymerase, in particular to the α-helix in the index finger that precedes the Gly64 mutation site (Figure 2.4). The α-helix,

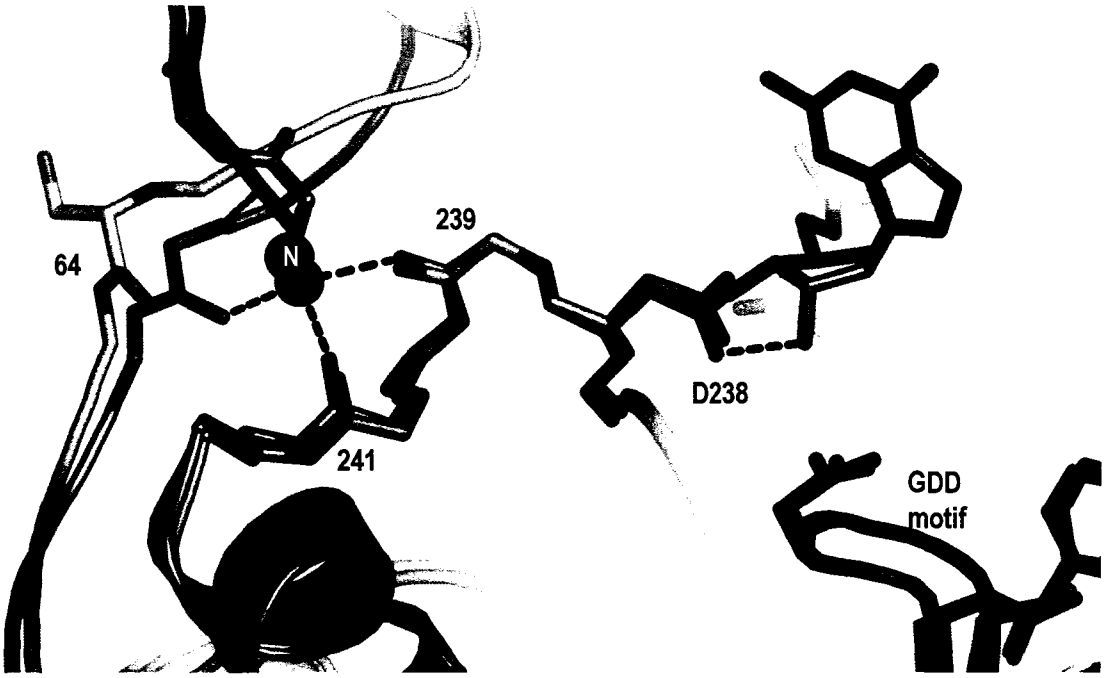


Figure 2.3: Structural alignment of $3D^{pol}$ and $3D^{pol}$ G64S. The structure of $3D^{pol}$ G64S (orange) shows that a serine at residue 64 will cause an expansion of the exterior portion of the N-terminus-binding pocket resulting in the loss of two hydrogen bonds to the N-terminus. For simplicity hydrogen bonds from the N-terminus to residue 239 and 241 are not shown for $3D^{pol}$ G64S. The N-terminus is slightly displaced compared to native $3D^{pol}$. However, Asp238 and the GDD motif do not show significant changes in conformation.

Table 2.1 – Polymerase extension activities for Gly64 mutants

Protein	% Activity
3DDD G64S	28 ± 2
3DDD G64A	26 ± 2
3DDD G64D	18 ± 3
3DDD G64N	33 ± 8

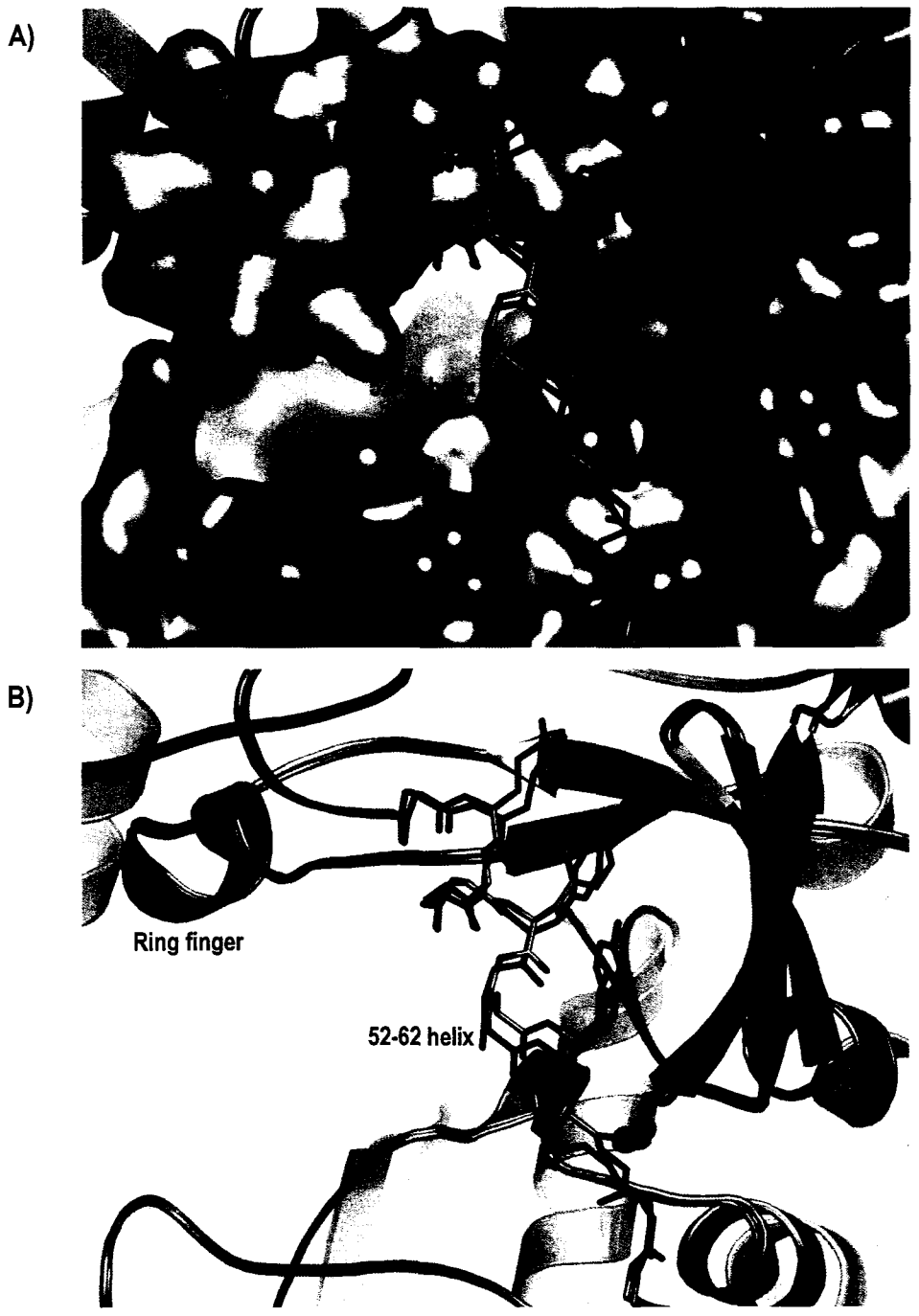


Figure 2.4: An α-helix comprised of residue 52-62 moves in response to the G64S mutation. **A)** Structural alignment of native $3D^{pol}$ (normal coloring scheme) with $3D^{pol}$ G64S (cyan) shows that the 52-62 α-helix, shown as sticks with residue 64 colored purple, moves up and away from the active site. Surface representations of the ring finger, the palm and the middle finger are shown displaying the tight interaction this α-helix has with these structural motifs. The N-terminus of each protein is shown as a blue sphere and the active site (magenta) can be seen in the background. **B)** Same view and color scheme as shown in A) however, all structural elements besides the α-helix are shown as cartoon models.

comprising residues 52-62, shifts up and away from the N-terminus-binding pocket relative to native $3D^{pol}$. Interestingly, the binding of RTP results in a shift of the helix back toward the N-terminus-binding pocket and this is accompanied by a marked improvement in the density surrounding the N-terminus, suggesting that the binding of an NTP locks the polymerase in a less dynamic state (Figure 2.5).

Why is Only Serine Found in Ribavirin Resistant Mutations?

We were also interested in determining if there was anything uniquely favorable about a serine substitution at Gly64 since two different laboratories recovered the same ribavirin resistance mutation. The lack of density and high B-factors for the side-chain of residue 64 do not suggest that the residue makes any specific structural contacts that would require a serine. To test whether the expansion of the pocket was due primarily to the loss in conformational freedom and not an intrinsic effect mediated by the serine side-chain, I expressed and crystallized a mutant that contained an alanine at residue 64. Alanine, like serine, will not readily adopt unfavorable phi and psi angles, but minimizes the potential for side-chain specific interactions. G64A is virtually identical with G64S in all aspects tested. The structure of G64A is superimposable with G64S, both having an expanded N-terminus-binding pocket accompanied by a movement in helix 52-62 (Figure 2.6). In addition this mutant shows a comparable ~4-fold decrease in activity (Table 2.1) and small displacements of the N-terminus and Asp238.

Ribavirin mutagenesis results in transversion mutations of the original GGU codon at residue 64, leading to three possible amino acids substitutions: serine (AGU), aspartic acid (GAU) and asparagine (AAU). Mutations of the wobble position to GGN retain the wild-type Gly. While it is statistically improbable to generate a double mutant resulting in asparagine at residue 64, both it (G64N) and the aspartic acid (G64D) mutations were generated and analyzed. Like the G64A

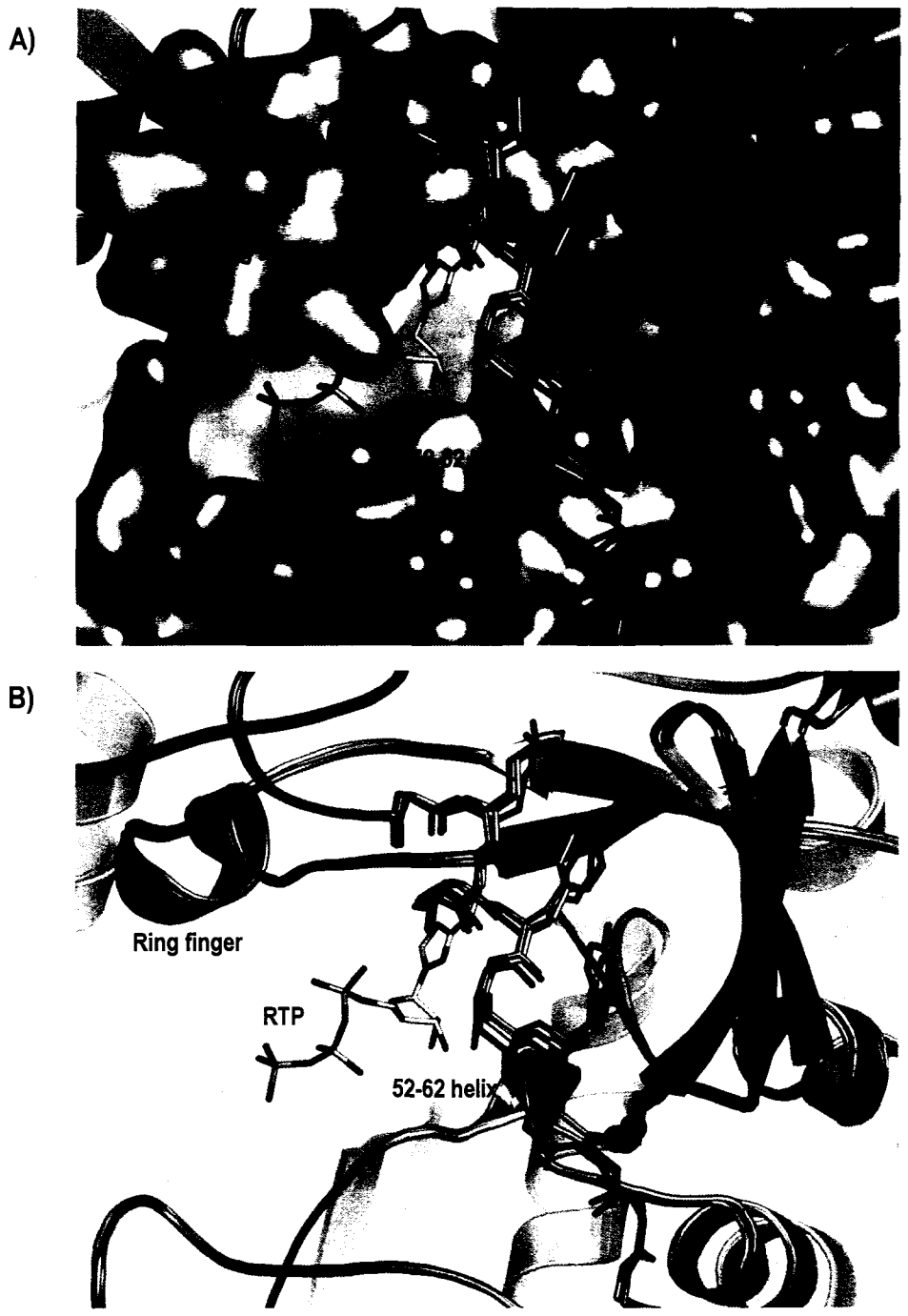


Figure 2.5: The binding of an NTP to the $3D^{pol}$ G64S mutant results in a shift of the 52-62 α-helix back toward native-like topology. **A)** Structural alignment of $3D^{pol}$ (normal coloring scheme), $3D^{pol}$ G64S (cyan), and RTP bound $3D^{pol}$ G64S (wheat). A stick model of the 52-62 α-helix is shown with Gly64 colored purple and the N-terminus is represented by a blue sphere. The binding of an NTP results in a small shift of the α-helix back toward native-like topology. **B)** Identical view and coloring scheme as shown in A) with the surface representation of the fingers and palm domains removed.

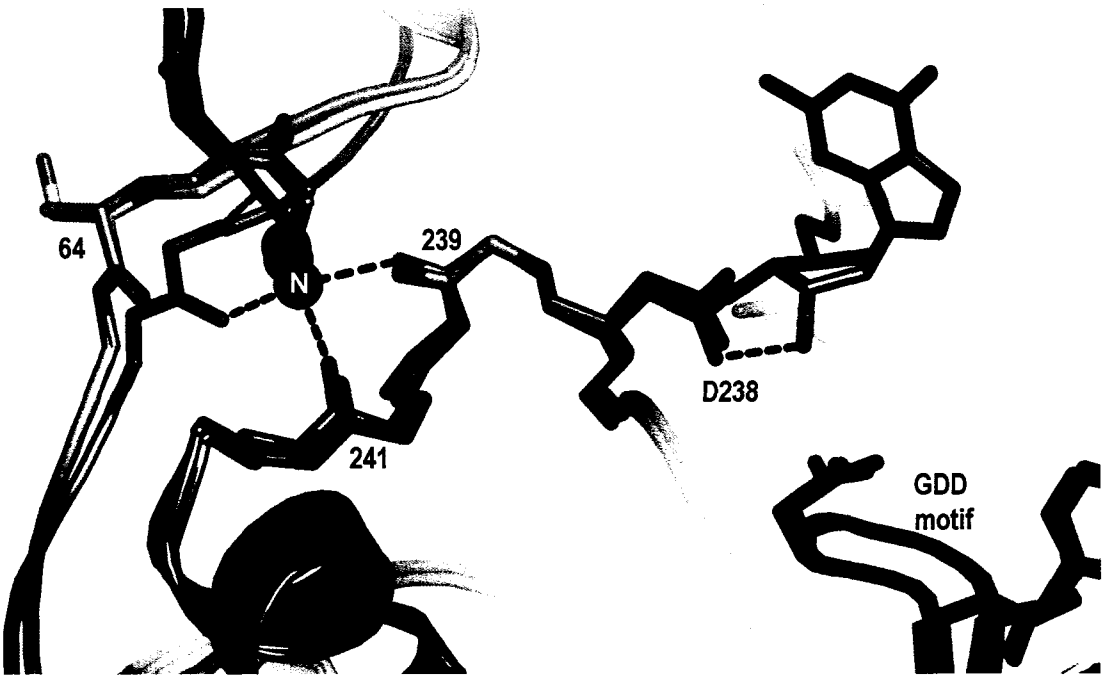


Figure 2.6: Structural alignment of $3D^{pol}$, $3D^{pol}$ G64A and $3D^{pol}$ G64S. Substitution of Gly64 with an alanine (cyan) results in a backbone geometry that is identical to $3D^{pol}$ G64S (orange) displaying an expansion of the N-terminus-binding pocket and the loss of two hydrogen bonds from the N-terminus to residue 64. The G64A mutant, like G64S, does not display a significant change in Asp238 and the GDD motif.

mutation, the structures of both G64N and G64D were virtually superimposable with the G64S structure and these mutants display ~4-fold and ~5-fold reductions in activity respectively (Figure 2.7) (Table 2.1).

X-ray Crystal Structures with Ribavirin tri-phosphate (RTP)

The crystal structures of native $3D^{pol}$ and G64S in the presence of RTP show that the antiviral drug is found to have identical interactions with the protein as does GTP, supporting the idea that RTP acts as a nucleoside analog (Figure 2.8) (Crotty et al., 2000). The position of the base is stabilized by both a stacking interaction with the side-chain of Arg174 and a hydrogen bond between the exocyclic carboxy-amide and the backbone carbonyl of residue 175. In addition, multiple hydrogen bonds are made between the ribose and phosphate moieties of RTP and the side-chains of residues 163, 167 and 174. The hydrogen bond between the 2′-OH of the RTP and Asp238 is conserved in both structures, an interaction shown to be crucial for NTP/dNTP selection (Gohara et al., 2000; Huang et al., 1997). The binding site of the RTP molecule, like that of GTP, is pre-catalytic in that the base-pairing face of the nucleotide is facing the protein and the phosphate moiety is not yet positioned over the Mg^{2+} ion-containing active site. The binding of the RTP molecule has little or no effect on the overall structure, other than the aforementioned change seen in the α-helix comprised of residues 52-62, and all other structural elements and folds remain identical to those seen in the apo-structures.

Discussion

Structural investigation of the RNA-dependent RNA polymerase of poliovirus has lead to a detailed understanding of the importance of proper proteolytic processing and subsequent burial of the N-terminus in a specific pocket at the base of the fingers domain. Our lab has shown that this

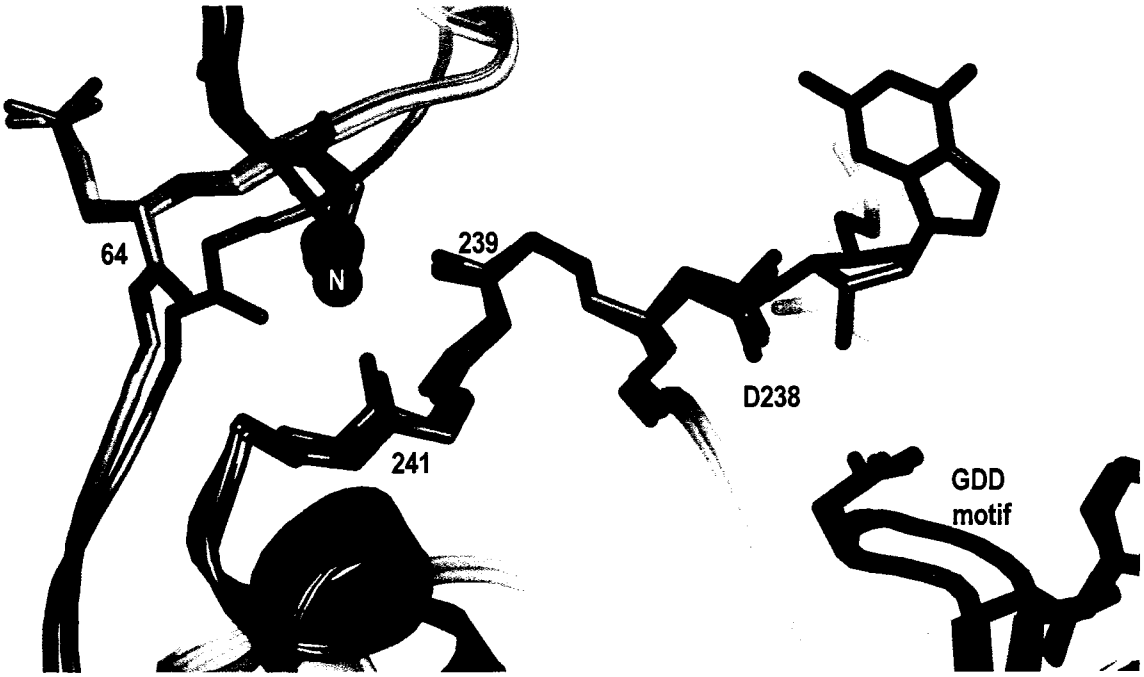


Figure 2.7: Structural alignment of $3D^{pol}$, $3D^{pol}$ G64A, G64D, G64N and G64S. Identical view of the N-terminus-binding pocket as that seen in Figure 2.6 with the addition of $3D^{pol}$ G64D (black) and G64N (salmon). Like $3D^{pol}$ G64A, the backbone geometry of $3D^{pol}$ G64D and G64N are identical to that of $3D^{pol}$ G64S.

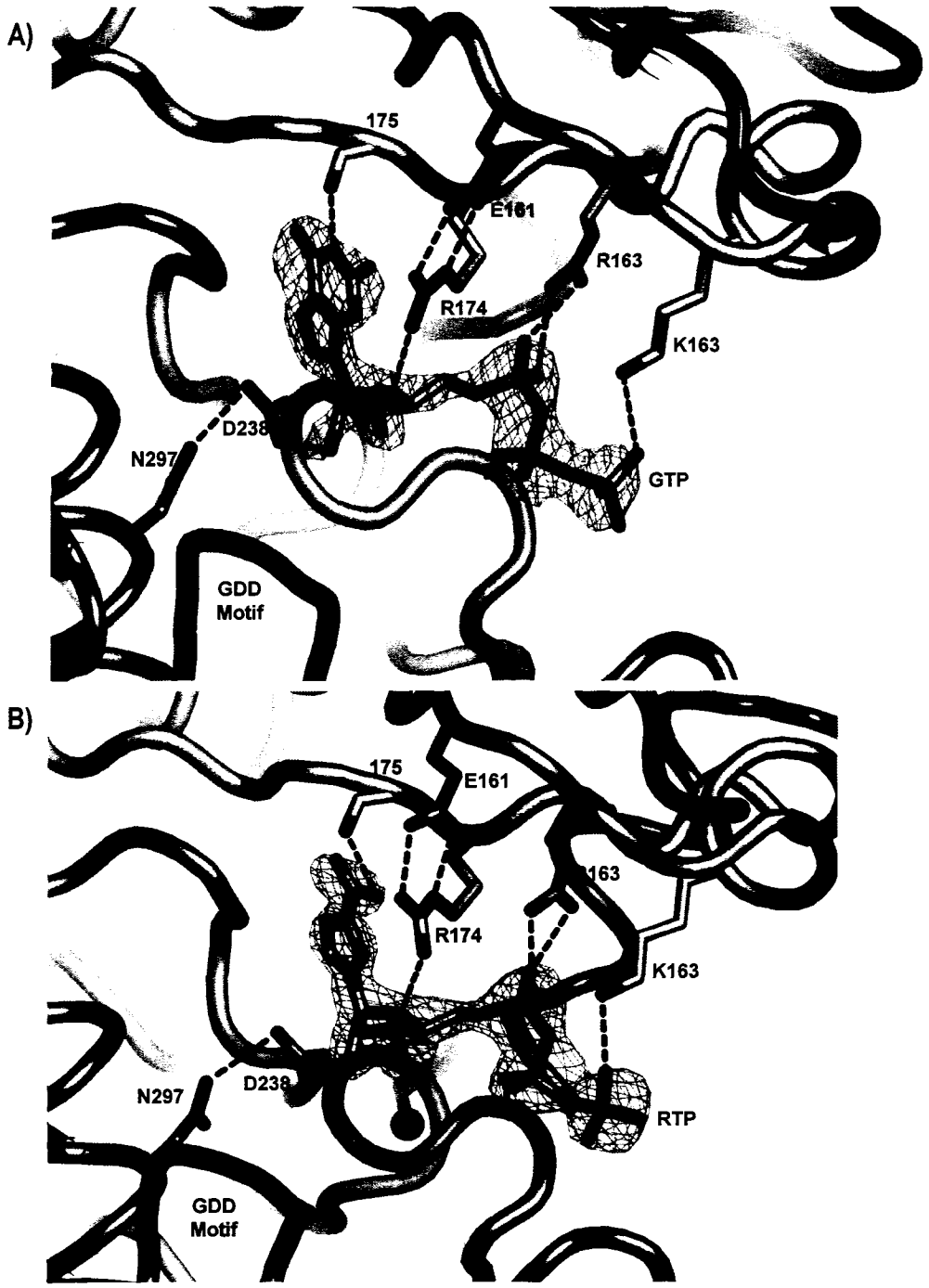


Figure 2.8: RTP 1000K composite simulated anneal omit map contoured at 1.6σ. A) GTP-bound $3D^{pol}$ showing multiple interactions between the ring finger and the NTP including the hydrogen bond between Asp238 and the 2′-OH (Thompson and Peersen, 2004 (1RA7)). **B)** RTP bound $3D^{pol}$ shows that RTP has identical interactions with the protein as a canonical NTP supporting its role as a nucleoside analog. Maps for both GTP and RTP are composite simulated anneal omit maps contoured at 1.6σ.

interaction is essential for the activity of the polymerase and proper positioning of Asp238 in the active site, allowing it to form a hydrogen bond with the 2′-OH of the incoming NTP (Thompson and Peersen, 2004). Structural studies of rhinovirus (Appleby et al., 2005; Love et al., 2004), foot and mouth disease virus (Ferrer-Orta et al., 2006; Ferrer-Orta et al., 2004) and coxsackie virus (Campagnola et al., in press) polymerases reveal an analogous feature, suggesting a similar link between polymerase activity, structural integrity of the polymerase active site, and proper positioning of the N-terminus across all picornaviruses. The isolation of a mutation in a residue comprising the outer edge of the N-terminus-binding pocket leading to higher fidelity and resistance to the antiviral drug ribavirin prompted us to investigate whether these affects could be accounted for structurally.

Interestingly, two separate labs found only a serine mutant arising as a viable ribavirin resistant mutant, begging the question whether the serine side-chain plays a unique structural role. For example, serine side-chains have been known to hydrogen bond back to the protein backbone, which in this case could be stabilizing unfavorable backbone geometry. However, in light of the essentially identical structure and activity of G64S and G64A mutants and the surface-exposed structure of the serine side-chain, it is unlikely that the hydrogen bonding potential of the serine side chain plays any significant structural role.

Instead, the serine mutation is likely the result of the limited set of mutations that can be obtained by a ribavirin-induced G to A transversion of the original GGU codon. Transversion of the first G to A in positive-strand synthesis results in an AGU codon and the observed Ser64 mutation, while mutation of the second G results in a GAU codon encoding an aspartic acid. Minus-strand mutagenesis of the A opposite the U in the Gly64 codon during viral replication would result in a GGC codon in the plus strand that would still encode the wild-type Gly64 residue. Thus, a serine

and an aspartic acid are the only residues possible with a single ribavirin-induced point mutation in the codon. Even though they are unlikely to be recovered in a growth-selection experiment, double mutations could result in asparagine (AAU), serine (AGC) or aspartic acid (GAC) codons in addition to reversions back to glycine. Therefore, mutants containing either an asparagine (G64N) or an aspartic acid (G64D) at residue 64 were also expressed and analyzed structurally and biochemically. Like G64A, both G64N and G64D mutants were identical in structure to the G64S mutant. The G64N mutant had a similar ~4-fold reduction in elongation activity while the G64D mutant displayed a ~5-fold reduction in activity (Table 2.1). These results imply that a minimum elongation rate may be necessary to support viral growth in cultured cells, a condition not met by the G64D mutant. Therefore, the lower activity of the G64D single mutant and the statistical improbability of isolating the G64N double mutant are both likely explanations for the isolation of only a G64S mutant in growth-selection experiments in the presence of ribavirin.

Ribavirin is currently being used to treat hepatitis C, severe acute respiratory viruses and Lassa fever virus (Andrei and De Clercq, 1993; Cummings et al., 2001; Krilov, 2001). Its mode of action, up until recently, was thought to be the inhibition of inosine monophosphate dehydrogenase (IMPDH), an enzyme required for the catabolism of GTP (Muller et al., 1977). Low intracellular levels of GTP may lead to a reduction in viral protein translation and RNA replication and a reduction in the efficiency of RNA capping (Bougie and Bisaillon, 2004; Hong and Cameron, 2002). Ribavirin has also been shown to directly inhibit the polymerases of hantaan virus, HCV, HIV and influenza virus (Fernandez-Larsson and Patterson, 1990; Maag et al., 2001; Severson et al., 2003; Vo et al., 2003; Wray et al., 1985). However, the Cameron lab has shown that upon entering the cell, ribavirin is converted to ribavirin triphosphate (RTP), which then acts as a nucleoside analog (Crotty et al., 2000). The structures of RTP bound to both the G64S mutant and native $3D^{pol}$ further support that it does indeed bind as a nucleotide analog and is found to have identical

interactions with the protein as compared to all other NTPs (Figure 2.8) (Thompson and Peersen, 2004; Thompson, 2006). It is also evident that binding of RTP in G64S and native $3D^{pol}$ is identical, supporting the nearly equal affinities each protein has for the nucleoside analog (Arnold et al., 2005). These findings strongly support the conclusion that it is an increase in fidelity of the G64S polymerase that leads to ribavirin resistance, and not simply a defect in RTP binding.

Craig Cameron's lab has done a full kinetic and thermodynamic analysis of G64S and has attributed the increase in fidelity to the slow kinetics of a conformational change preceding phosphoryl transfer. In this paper the authors observed a 3-fold decrease in the overall elongation rate of G64S as compared to native $3D^{pol}$. Furthermore, they observe a 3-fold decrease in the equilibrium constant for the conformational change that takes place after NTP binding but before phosphoryl transfer, implicating this as the rate-limiting step in catalysis for G64S (Arnold et al., 2005). We set out to provide a structural explanation for how the G64S mutation slows this conformational change and, while this remains an elusive goal, the structures reported here, in combination with already existing structures, lead us to believe that there is a dynamic interplay between the N-terminus, Asp238, and the binding and incorporation of NTP's.

The G64S mutant has a significant impact on the structural integrity of the N-terminus-binding pocket, showing a large expansion of the outer edge resulting in the loss of two hydrogen bonds from residue 64 to the N-terminus and a 4-fold decrease in catalytic activity as compared to native $3D^{pol}$. A nearly identical expansion of the N-terminus-binding pocket is observed when replacing Gly1 with an alanine (G1A), which is accompanied by a 2-fold decrease in activity (Figure 2.9). Our lab has attributed the loss of activity in G1A to the displacement of the N-terminus leading to altered positioning of Asp238 in the active site (Thompson and Peersen, 2004), and it was quite surprising to see that while the N-terminus of G64S was slightly displaced, Asp238 displayed no significant movement. In addition, Grace Campagnola of our lab has solved the structure of a

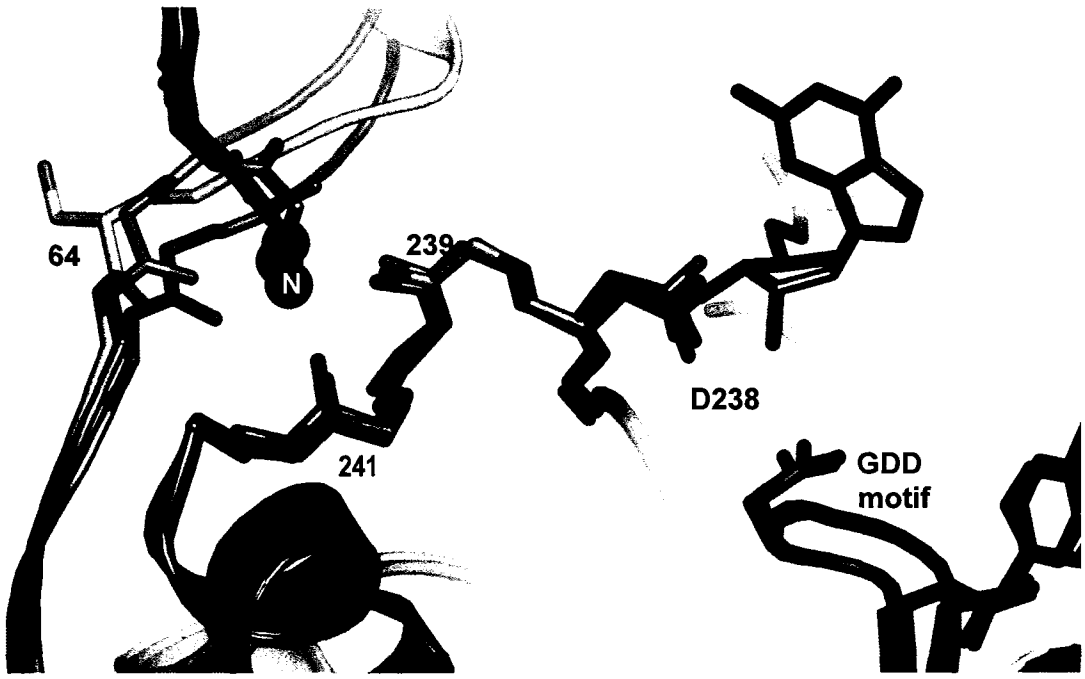


Figure 2.9: Structural alignment of $3D^{pol}$ G1A, $3D^{pol}$ G64S and native $3D^{pol}$. $3D^{pol}$ G1A, shown in pink, results in an expansion of the N-terminus-binding pocket similar to $3D^{pol}$ G64S and a small displacement of both the N-terminus and Asp238, resulting in a 2-fold decrease in activity (*Thompson and Peersen, 2004 (1TQL)*). Oddly, the movement of Asp238 is not observed in $3D^{pol}$ G64S despite a more drastic affect on the activity of the enzyme. Substitution of Gly1 with an alanine results in the loss of a hydrogen bond from residue 241 to the N-terminus. For the sake of simplicity, hydrogen bonds to the N-terminus are not shown.

polymerase in which Gln65 has been replaced with an alanine (N65A). This side chain is hydrogen bonded to the N-terminus in rhinovirus (Love et al., 2004) and FMDV (Ferrer-Orta et al., 2004) polymerase structures and may represent a second position for the N-terminus in poliovirus $3D^{pol}$. Interestingly, the N65A mutation causes a collapse, rather than an expansion, of the outer edge of the N-terminus-binding pocket that is accompanied by a 10-fold decrease in activity (Figure 2.10). The N-terminus of this structure is not well-ordered but appears to have been significantly displaced but, Asp238 again exhibits little or no displacement (Thompson, 2006).

How then can one correlate the clear loss in activity with the differing structural perturbations surrounding the N-terminus and Asp238 that arise from separate mutations? To answer this question we compared the G64S structures with an ensemble of structures solved in our lab using Theseus, a structural alignment program developed by Doug Theobald (Theobald and Wuttke, 2006). The strength of Theseus is in its use of maximum likelihood probability mathematics that allow it to down-weight the contribution from residues with clearly different relative coordinates. The result is an alignment that is dominated by domains of each protein which are clearly super-imposable, thereby highlighting those regions that differ.

Structural alignments of an ensemble of polymerase structures solved by our lab reveal that N65A, G1A, and all mutations at residue 64 lead to a movement of residues 52-62, comprising an α-helix in the index finger, up and away from the N-terminus (Figure 2.11). In the N65A structure, the movement in this α-helix is accompanied by a movement in adjacent residues at the tip of the ring finger (Figure 2.12). This finger, comprised of residues 150-179, forms the roof of the NTP entry tunnel and is involved in a fairly extensive interface with the index finger directly adjacent to the 52-62 α-helix. The fact that the ring finger moves in response to N-terminus-binding pocket perturbations demonstrates that the fingers domain is not a rigid body, but is rather quite flexible

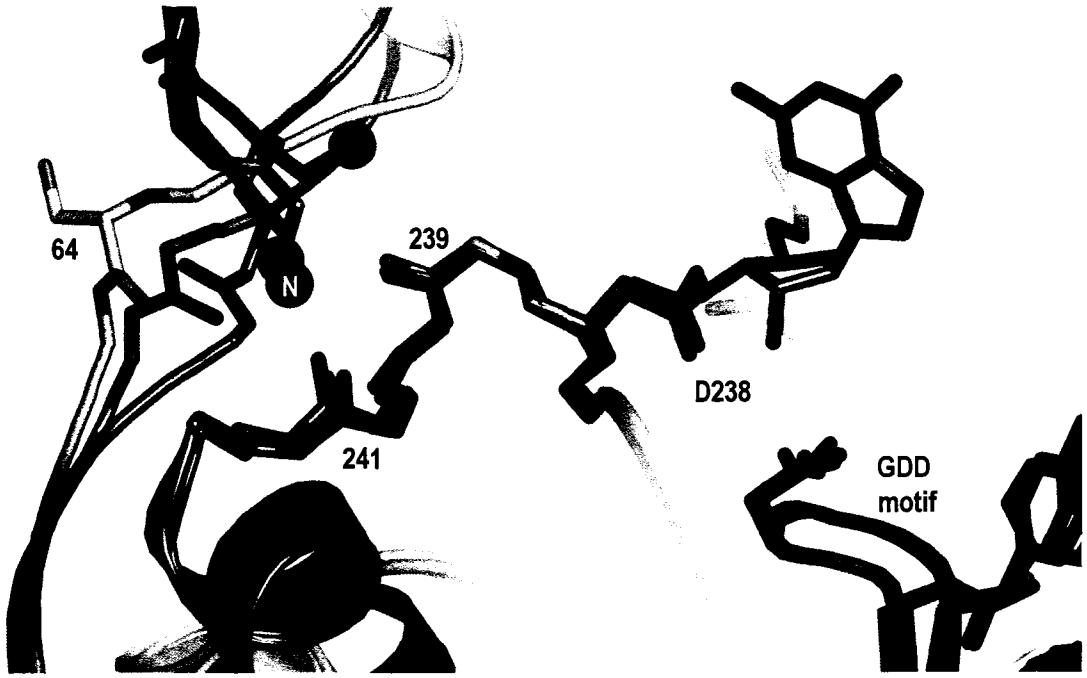


Figure 2.10: Structural alignment of $3D^{pol}$ N65A, $3D^{pol}$ G64S and native $3D^{pol}$. $3D^{pol}$ N65A, shown in lime, results in a collapse of the N-terminus-binding pocket and a large displacement of the N-terminus, and yet Asp238 and the GDD motif do not show significant alterations (Thompson, 2006). Substitution of Asn65 with an alanine results in the loss of all hydrogen bonds to the N-terminus except a single hydrogen bond from its carbonyl to the backbone amide of Gly64. For simplicity sake all hydrogen bonds to the N-terminus are not shown.

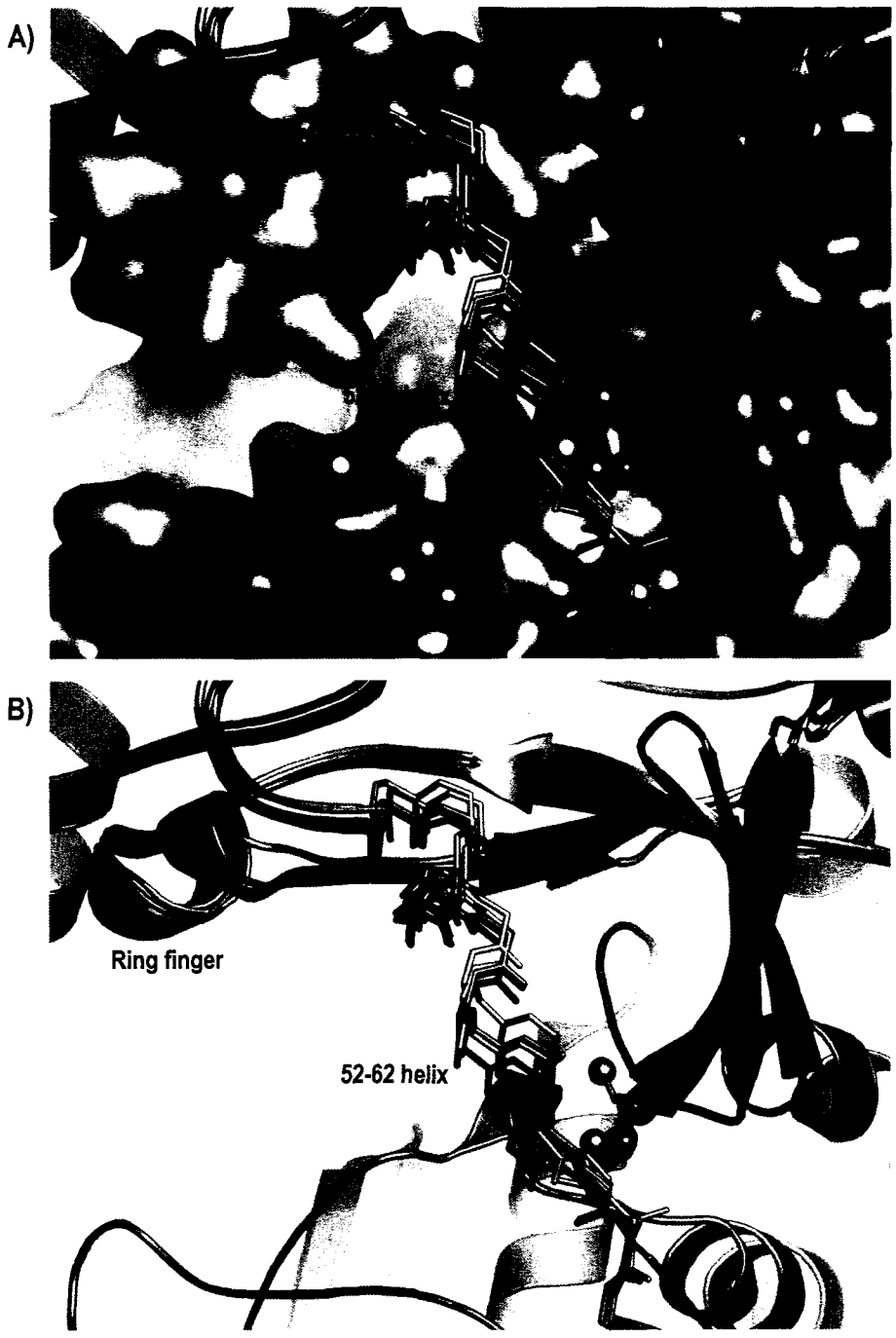


Figure 2.11: The movement of the 52-62 α-helix is a common structural feature observed among mutations that disrupt the N-terminus-binding pocket. A) Structural alignment of native $3D^{pol}$ (normal coloring scheme) with Gly64 mutants (cyan), G1A (wheat) and N65A (wheat) shows that the 52-62 α-helix (shown as sticks) moves up and away from the active site in all cases. The N-terminus of each protein is shown as a sphere and the active site (magenta) can be seen in the background. **B)** Same view and color scheme as shown in A) however, all structural elements besides the α-helix are shown as cartoon models.

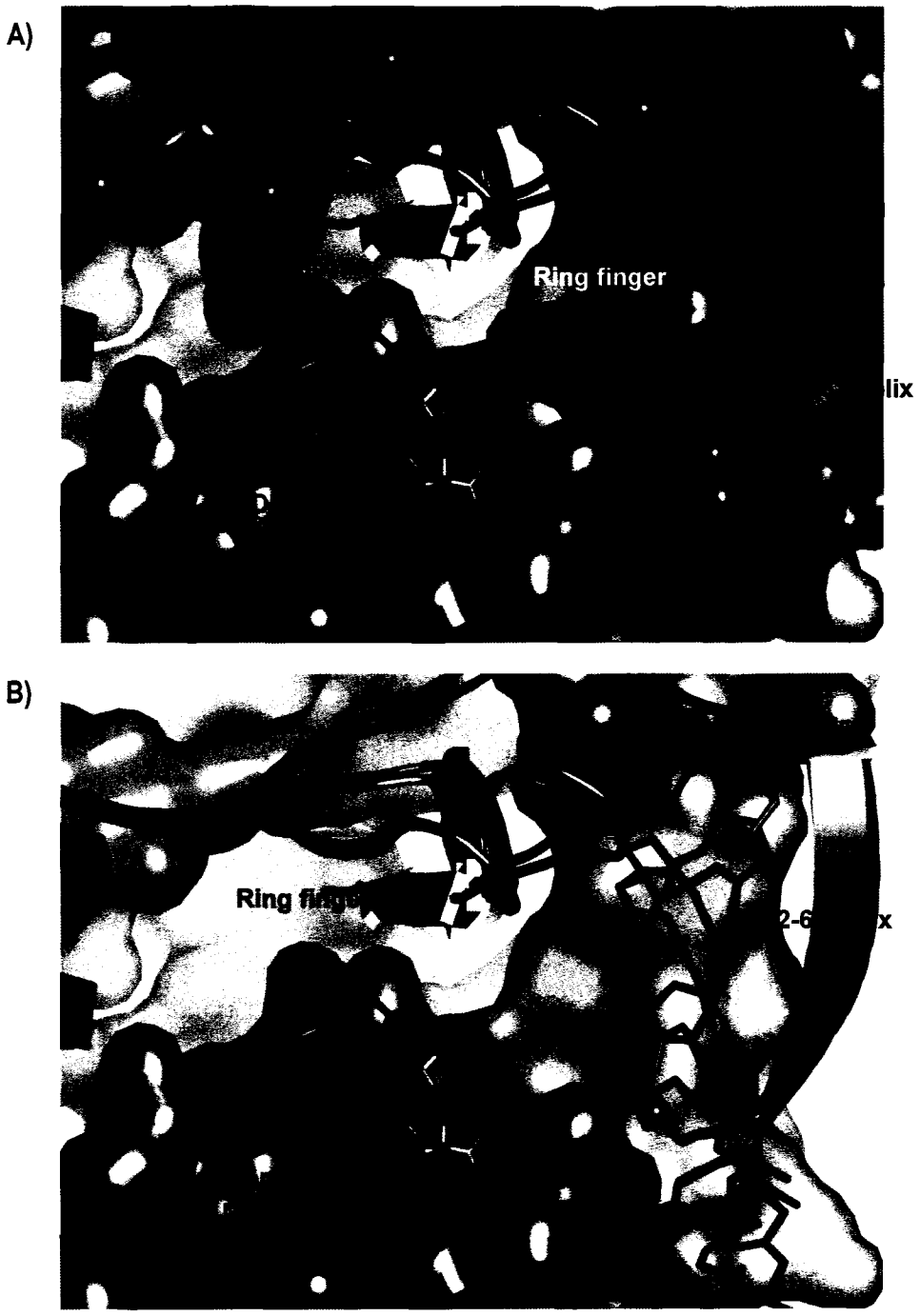


Figure 2.12: The 3D^{pol} N65A mutant affects the conformation of the ring finger. A) Cartoon representation of the ring finger from native 3D^{pol} (yellow) and 3D^{pol} N65A (slate) demonstrates a small movement in response to the movement of the 52-62 α-helix. Surface representation of the index finger demonstrate how the ring finger is tucked beneath a rather large structural platform. **B)** Identical view and coloring scheme to that seen in A). However, surface representations of the index finger (green) have been removed. The stick model of the 52-62 α-helix demonstrates the rather large movement made by this helix in response to the 3D^{pol} N65A mutation (slate). The NTP in both A) and B) has been modeled from GTP-bound 3D^{pol} (Thompson and Peersen, 2004).

and small movements in one location can result in movements in separate portions of the protein. Indeed, binding of an NTP to the G64S mutant results in a small movement of the 52-62 α-helix back to native-like topology, indicating that the index finger is responsive to events taking place in the NTP-entry tunnel and the active site.

The fact that the NTP is bound in such a manner as to prevent base-pairing and catalysis demonstrates the need for a conformational change of the nucleotide and the protein before catalysis can occur. An analogous situation is found in the RNA-polymerase of bacteriophage T7, where the Steitz lab has demonstrated that the NTP-binding O-helix undergoes a significant conformational change to place the incoming NTP over the active site (Yin and Steitz, 2004). Initially, NTP binds the T7 RNAP in a conformation unsuitable for catalysis in that the phosphate moiety is not positioned over the active site nor is the base able to base pair with the template. Following NTP-binding the O-helix undergoes a 22.5° rotation to position the NTP phosphates over the catalytic aspartate residues. Furthermore, the movement of the O-helix displaces Tyr639 from the base-binding site, allowing the incoming base to base pair with the template strand (Figure 2.13) (Yin and Steitz, 2004).

The Cameron lab has described the full kinetic mechanism for poliovirus $3D^{pol}$ and has shown that a conformational change preceding phosphoryl transfer is the rate-limiting step in $3D^{pol}$ G64S catalysis (Arnold et al., 2005). Many in the field, including the Cameron lab, believe that this conformational change is similar to the conformational change observed in the O-helix of T7 RNAP that positions the NTP over the active site. Without further structural information it is difficult to say how this conformational change occurs in $3D^{pol}$, but our data demonstrates that structural elements in the fingers domain are flexible and responsive to events taking place in the active site. Based off these data, and by analogy to the O-helix, I propose that a yet-to-be-seen movement of the 52-

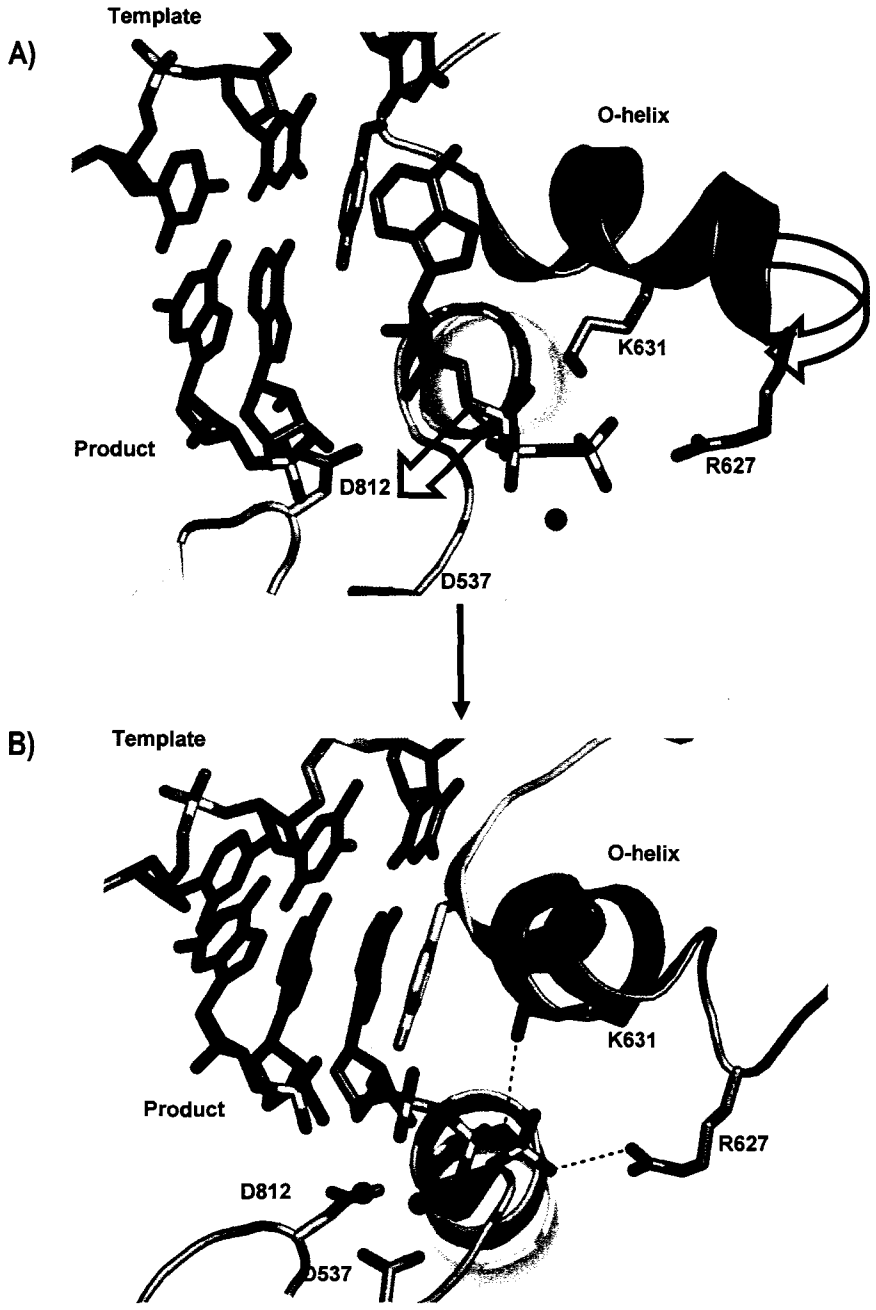


Figure 2.13: Structural basis for NTP positioning in T7 RNA-polymerase. **A)** Initial binding of the NTP to the O-helix results in a conformation that is unsuitable for catalysis, requiring a conformational change prior to phosphoryl transfer. The template and product strands of the nucleic acid have been labeled and catalytic aspartates (D812 and D537) are shown as sticks. Arrows demonstrate the conformational change that needs to occur for a catalytically competent conformation (Cheetham and Steitz, 1999, (1QLN)) **B)** Catalytically competent conformation of the O-helix and NTP with two Mg^{2+} cations (blue spheres) coordinated by active-site aspartates and the phosphate moiety of the NTP. The O-helix undergoes a 22.5° rotation to move the NTP over the active site (Yin and Steitz, 2004 (1S77)).

62 α-helix will result in a clockwise rotation of the ring finger that will subsequently push the NTP into the active site (Figure 2.14). By replacing a Gly with Ser, the G64S mutant constricts movement within this region of the protein, favoring a distorted conformation of the 52-62 α-helix, giving a possible explanation for the kinetic and thermodynamic effects observed by the Cameron lab (Arnold et al., 2005). Indeed, the N65A mutant demonstrates how a movement of the 52-62 α-helix, similar to that observed in G64S, results in a counterclockwise rotation of the ring finger (Figure 2.12). As a result, the 52-62 α-helix in the G64S mutant polymerase likely represents an off-pathway conformation that must be brought back to native topology before the catalytic cycle can be completed. Supporting this hypothesis is the observation that NTP-binding brings the 52-62 α-helix back to a conformation similar to native $3D^{pol}$. However, further movement of the α-helix and rearrangements in the fingers domain would be required before the polymerase was on a pathway leading to catalysis. Additionally, the unfavorable conformation of the 52-62 α-helix may also account for an increase in fidelity in that it slows the mutant polymerase and thus favors canonical base pairs due to an energetically more favorable association preventing premature NTP release.

The results presented in this chapter, in combination with the G1A and N65A mutations, are leading us away from the view that structural perturbations of the N-terminus-binding pocket effect Asp238 positioning alone, but rather have larger implications affecting the kinetics and thermodynamic parameters governing a conformational change in the ring finger and the polymerase mechanism as a whole. In conclusion we believe that a dynamic correlation exists between movements in the fingers domain, the position of the N-terminus and Asp238, and the binding and incorporation of NTP's. However, without further structural information the precise molecular details of these events remain elusive.

Figure 2.14: Proposed movement of the ring finger. Surface representation of $3D^{pol}$ demonstrates how the ring finger (yellow) makes multiple contacts with the index finger (green), in particular the 52-62 α-helix (cyan). Movements within the 52-62 α-helix may result in a clockwise rotation (demonstrated by the arrow) of the ring finger that push the NTP into the active site (magenta) much like the O-helix of T7 RNAP (Yin and Steitz, 2004).

Chapter III

Structural and Biochemical Analysis of the 288-292 Loop, Implications for Polymerase Activity and Mechanism.

Introduction

Incorporation of a single nucleotide into a growing nucleic acid is a complex process requiring much more than a simple phosphoryl transfer catalyzed by the active site of the enzyme. Kinetic and thermodynamic studies of different classes of polymerases from *Escherichia coli* (Dahlberg and Benkovic, 1991; Kuchta et al., 1987), bacteriophage T7 (Patel et al., 1991), human immunodeficiency virus (Hsieh et al., 1993; Kati et al., 1992) and poliovirus (Arnold and Cameron, 2004) all reveal that a mechanism invoking up to five steps is required for the addition of a single nucleotide. In all cases these steps include: 1) NTP binding 2) a conformational change 3) phosphodiester bond formation 4) a second conformational change and 5) pyrophosphate release (Figure 3.1). The conformational changes flanking the phosphodiester bond formation have been hypothesized to be correct positioning of the nucleoside triphosphate over the active site prior to chemical catalysis and translocation of the template/product strand following chemical catalysis (Dahlberg and Benkovic, 1991; Kuchta et al., 1987; Patel et al., 1991). In support of this hypothesis the Steitz lab published a series of papers describing the structures of the DNA-dependent RNA polymerase from bacteriophage T7 (T7 RNAP) in various phases of its catalytic cycle (Cheetham and Steitz, 1999; Tahirov et al., 2002; Yin and Steitz, 2002; Yin and Steitz, 2004).

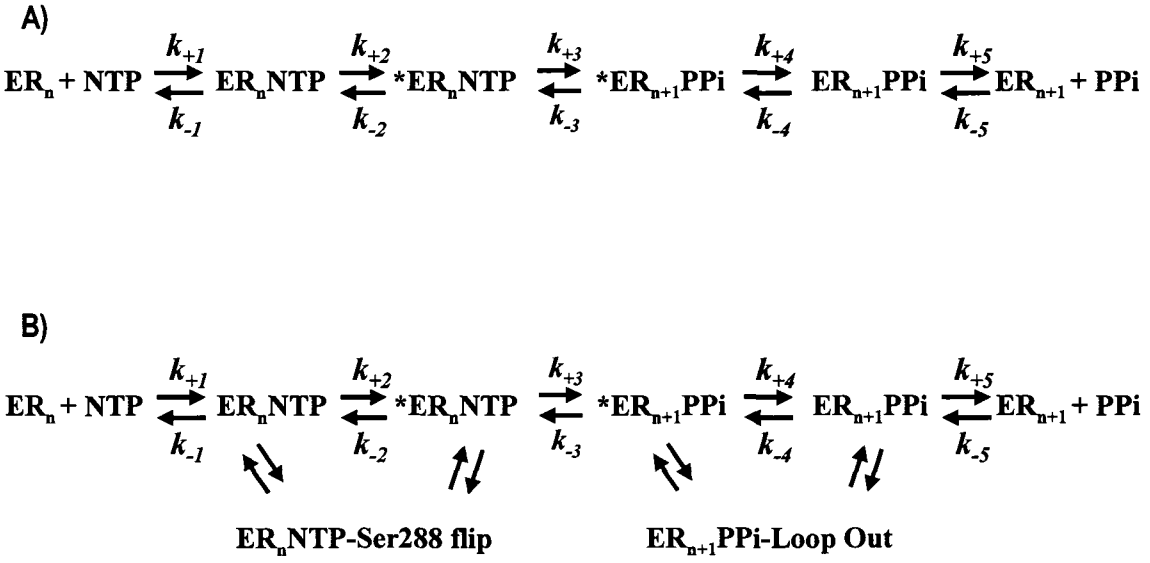


Figure 3.1: Kinetic scheme for the catalytic cycle of a polymerase. A) Each catalytic cycle begins with NTP binding to the Enzyme:RNA complex (ER_n). This event is followed by a conformational change ($^*ER_nNTP$) thought to position the NTP over the actives site. The conformational change is followed by phosphoryl transfer making the product strand one nucleotide longer ($^*ER_{n+1}PPi$). The complex then undergoes a second conformational change and inorganic phosphate (PPi) is released. **B)** The results presented in this chapter represent smaller conformational changes that lead to the larger conformational changes described kinetically in A). We propose that Ser288 of $3D^{pol}$ must first flip (ER_nNTP-Ser288 flip) down leading to a series of events that have been measured kinetically as the first conformational change. Following phosphoryl transfer, the 288-292 loop of $3D^{pol}$ will adopt a second conformation ($ER_{n+1}PPi$-Loop Out) along with a series of other yet-to-be-described events leading to the second conformational change. (Figure adapted from Arnold and Cameron, 2004)

These articles provide the molecular details of two conformational changes required for catalysis that do indeed correlate with NTP positioning and translocation.

In 1997, Steve Schultz's lab published the partial structure of the RNA-dependent RNA polymerase (RdRP) from poliovirus (Hansen et al., 1997), demonstrating that the structures of this class of polymerases could also be described using the right-hand analogy first used by Tom Steitz to describe a fragment of the DNA polymerase from *Escherichia coli* (Klenow fragment) (Ollis et al., 1985). Since this publication, the structures of at least nine separate RdRP have been solved, including the full-length structure of the RdRP from poliovirus. Observation of each structure reveals striking similarities between the members of this class of polymerases, all of which adopt the right-hand analogy with an enclosed active site characterized by a contact between the finger and thumb domains (Figure 1.4). In addition to the apo-structures, the structures of foot-and-mouth disease virus (Ferrer-Orta et al., 2004), bacteriophage ϕ6 (Butcher et al., 2001; Salgado et al., 2004), hepatitis C virus (O'Farrell et al., 2003) and reovirus λ3 (Tao et al., 2002) RdRP's have all been solved in complex with RNA. However, none of these structures have demonstrated the conformational changes required for NTP positioning or translocation as seen in the T7 RNAP structures.

Investigation of the ribavirin-resistant mutant polymerase $3D^{pol}$ G64S from poliovirus lead to the discovery of a patch of positive density surrounding residues 288-292 in the palm domain. This density was suggestive of an alternate conformation for this five-amino acid loop and it caught our attention because of possible implications on the translocation of the template/product strand, giving the first hint of one of the conformational changes required for the catalytic cycle of the enzyme.

We report here the structural and biochemical analysis of 16 different mutants in this loop. Our results indicate that Ser288 and the loop can exist in two stable conformations described by: 1) a 180° flip of Ser288 down towards the active site, and 2) accessibility of residue 290 to a hydrophobic pocket directly behind the loop. We demonstrate that the ability of the loop to undergo transitions between these separate conformations is modulated by the flexibility of a Gly at residue 289 and the nature of the side chain at residue 290. Our results indicate that the flexibility of this loop is essential for the elongation activity of the polymerase, and if restricted will result in translocation-deficient polymerases. From our results we hypothesize that for a complete catalytic cycle Ser288 must first flip down to form a hydrogen bond with Asp238, setting the stage for the conformational change required to correctly position the NTP over the active site. Following phosphoryl transfer the loop adopts a second conformation, moving residue 290 out of a hydrophobic pocket towards the active site, resulting in a steric clash with the template RNA that leads to the translocation of the nucleic acid (Figure 3.1).

Results

Crystallization and Structure Determination

Crystals and x-ray structures were obtained as reported by Thompson and Peersen (2004) with small modifications to cyro-protectant and soaking procedures as discussed in Chapter 2 and Appendix 1. All structures have at least one PEG molecule bound at a crystal-packing interface, and this appears to result in an increase in overall order of the crystal lattice and slightly higher resolution. Structures were refined to resolutions ranging from 2.0 Å to 2.8 Å with R and R_{free} values ranging from 20.7%-23.4% and 23.3%-25.6% respectively. The majority of structures were refined to a typical resolution of 2.3-2.4 Å and had R and R_{free} values separated by 2-3 percentage points (Tables 3-10, Appendix 2). Initial electron density maps were calculated using a model

missing residues 287-293, thus allowing for an unbiased calculation of the map surrounding this region of the protein. 1000K composite simulated anneal omit maps were calculated for the final structures to ensure the validity of each model (Appendix 3).

The overall structure of each mutant polymerase does not change significantly as compared to the native structure. All major structural folds and domains are maintained throughout and changes resulting from each mutation are isolated to movements within the 288-292 loop and surrounding structural elements, as revealed by structural alignments calculated by Theseus (Theobald and Wuttke, 2006). Changes in the 52-62 α-helix previously described for mutations surrounding the N-terminus-binding pocket (Chapter 2) are not observed in any of these structures, demonstrating that such movements can indeed be correlated with amino acid substitution and not merely a result of crystal packing or alternate soaking procedures.

The Initial Motivation

Structural investigation of $3D^{pol}$ G64S, a mutation resulting in ribavirin-resistant poliovirus, lead to the discovery of a good deal of positive density surrounding residues 288-292 that was suggestive of a well-defined alternate conformation. At the time we were unsure whether this density had arisen as a result of either mutations at residue 64 and/or complications arising from a dimethyl-arsenic adduct on residue 290. To control for this, crystals of native $3D^{pol}$ were grown and harvested in the exact conditions as those seen for the Gly64 mutants. These structures do indeed show hints of the alternate conformation, indicating that residues 288-292 are indeed flexible and that this flexibility is not a result of mutations surrounding the N-terminus-binding pocket. Interestingly, this density is not observed in the native $3D^{pol}$ structure solved by the Peersen lab in 2004 using crystals harvested in a 1.8M sodium acetate solution as opposed to the 250mM sodium acetate solution used for the Gly64 mutants. To further investigate the significance of the positive

density observed in the G64 mutant structures, 16 mutants of the 288-292 loop were generated, purified and analyzed structurally and biochemically.

The 288-292 Loop Conformation

An overview of all the mutant structures demonstrates that the conformation of residues 288-292 exists in two stable conformations (Figure 3.2). The first conformation is analogous to that which was originally described by Thompson and Peersen (2004) wherein the side chain of residue 290 is tucked into a hydrophobic pocket comprised of Val154, Met187, Tyr267, Val268 and Pro287. The second conformation is characterized by a ~5.0 Å movement of residue 290 out of the hydrophobic pocket and towards the active site (Figure 3.2 and Figure 3.3). The two conformations are distinguished by their relationship to the active site and the hydrophobic pocket and will therefore be referred to as the *in* conformation describing loops with the residue 290 tucked into the hydrophobic pocket and the *out* conformation describing loops with residue 290 proximal to the active site. The majority of mutant polymerase structures show predominant density in either the *in* or *out* conformation (Appendix 3). However, in some instances the conformation of the loop is complicated by a dimethyl-arsenic adduct at residue 290. This modification is often seen at cysteine residues in the original structure, including Cys290. The modification at Cys290 is tucked nicely into the hydrophobic pocket directly behind the loop and appears to not significantly alter the structure of the polymerase (Thompson and Peersen, 2004). Indeed, replacement of Cys290 with a serine results in virtually superimposable structures with only a small movement of Tyr267 to accommodate the lack of a dimethyl-arsenic adduct in the hydrophobic pocket (Figure 3.4). Mutant polymerases that retain the dimethyl-arsenic adduct display weaker overall density for residues 288-292 along with relatively high B-factors for the loop, both indicative of molecular disorder.

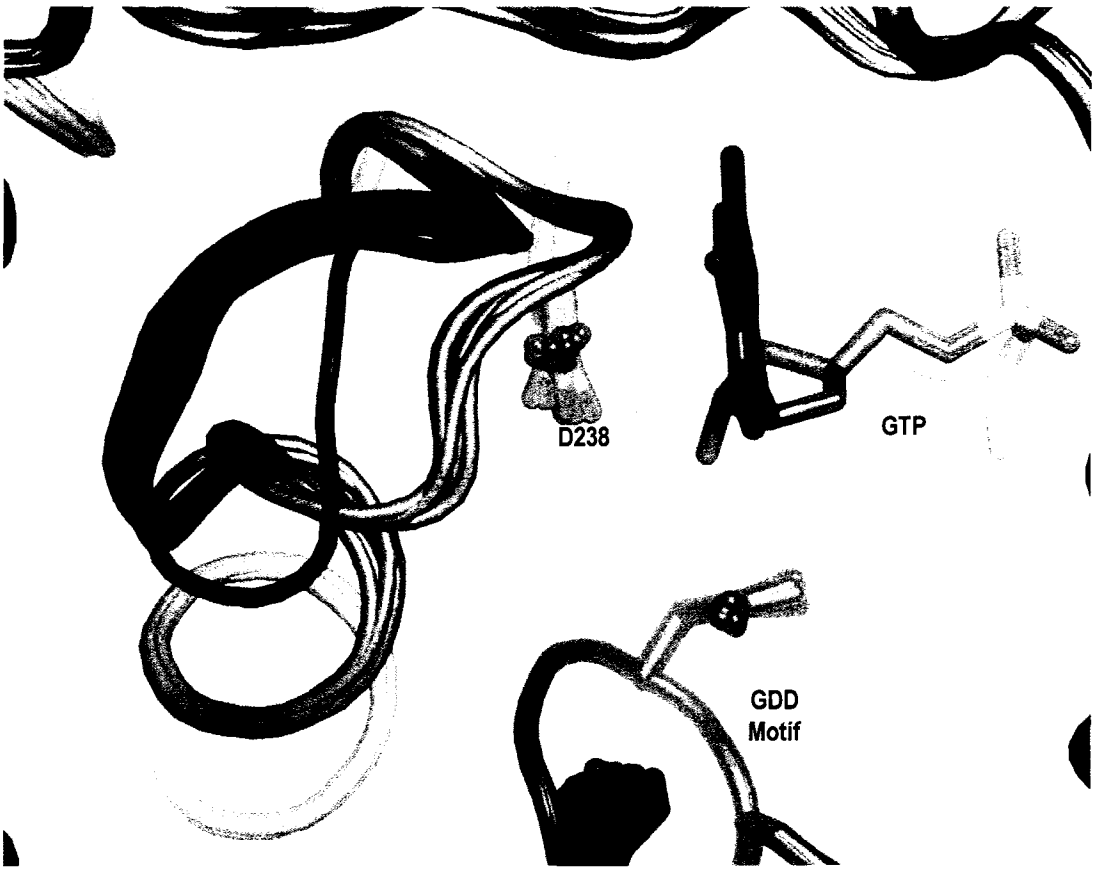


Figure 3.2: Structural alignment of 13 out of 16 loop mutations demonstrates how the 288-292 loop can exist in two stable conformations. The loops colored cyan represent the *in* conformation that was originally reported by Thompson and Peersen in 2004, whereas the loops colored orange represent the *out* conformation. The loop colored black has a distorted backbone geometry due to the complication of a dimethyl-arsenic adduct attached to residue 290 that is found on 4 out of 16 mutants. The active site is colored magenta and a stick model of Asp238 is shown for each mutant. The incoming NTP has been modeled, and loop structures represented here were solved in the absence of an NTP.

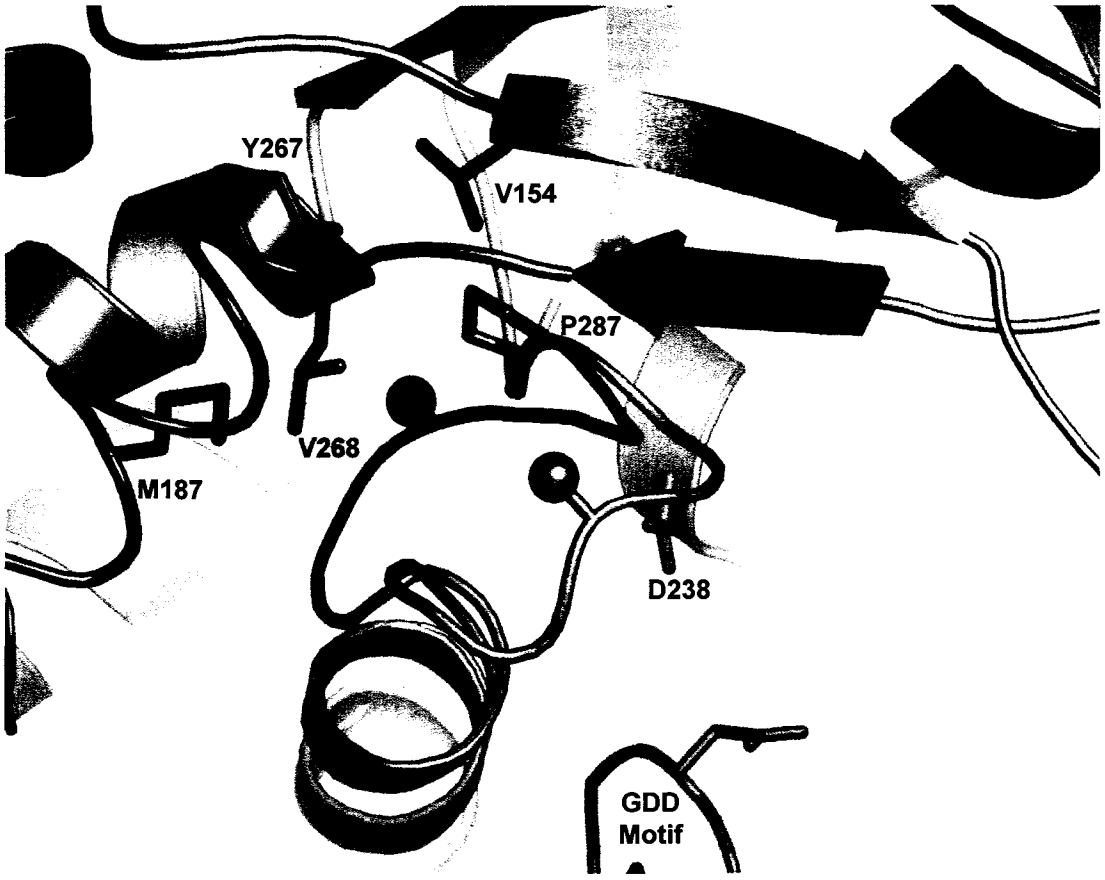


Figure 3.3: The *in* and the *out* conformations are differentiated by whether residue 290 is tucked into a hydrophobic pocket directly behind the loop. The *in* and *out* conformations are shown as either a cyan or orange loop respectively. Residue 290 is represented by a single ball and stick in the center of the loop and is tucked into a hydrophobic pocket comprised of V154, M187, Y267, V268 and P287 (each colored tan) in the *in* conformation. The alpha-carbon of residue 290 moves 5.0Å towards the active site (magenta) when moving from the *in* to the *out* conformation.

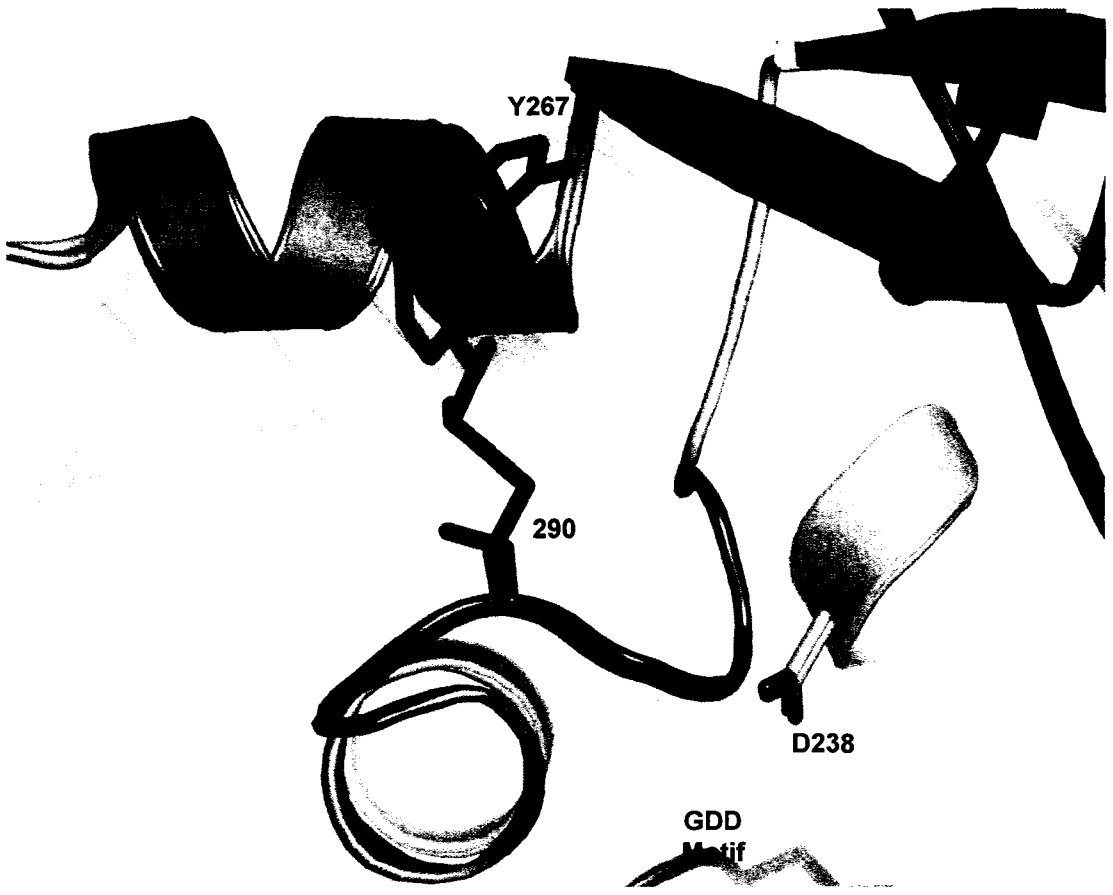


Figure 3.4: Tyrosine 267 moves in response to the removal of the dimethyl-arsenic adduct at residue 290. Structural alignment of the original structure (cyan) (Thompson and Peersen, 2004 (1RA6)) with the structure of C290S (tan) demonstrates how the removal of the dimethyl-arsenic adduct allows Tyr267 to move into the hydrophobic pocket. It is expected that the conformation of Tyr267 in C290S is its native conformation. The dimethyl-arsenic adduct is found on four of the mutant structures and significantly alters the backbone of the loop in a number of these cases.

In addition to the burial or removal of residue 290 from the hydrophobic pocket, each loop conformation can be characterized by several interactions with residues found at the base of the NTP-binding ring finger. In the *in* conformation, Gly289 forms two backbone hydrogen bonds linking its carbonyl oxygen and amide nitrogen with the amide nitrogen of residue 179 and the carbonyl oxygen of residue 177 respectively. These two hydrogen bonds are lost in the *out* conformation due to the ~5.0 Å movement of the loop towards the active site. Ser288 participates in two hydrogen bonds in both the *in* and *out* conformations. The first of these hydrogen bonds is conserved in both conformations and is between the Ser288 amide nitrogen and the side chain of Asp 177. The second hydrogen bond is conserved in that the transition from the *in* to the *out* conformation leads to an atomic "switch" whereby the amide nitrogen of Asp 177 is hydrogen bonded to either the Ser288 side chain (*in*) or the Ser288 carbonyl oxygen (*out*) (Figure 3.5). This switch is accompanied by a large change in the backbone torsion angles of Gly289 allowing the Ser288 carbonyl oxygen to fill the vacancy left by the Ser288 side chain during the *in* to *out* transition.

A further distinction between the *out* and *in* conformations can be made by observing the conformation of Ser288. When the loop is in the *out* conformation, this side chain points down towards the active site and is within hydrogen-bonding distance to the side chain of Asp238, the side chain that forms a key hydrogen bond to the 2′-OH of the incoming NTP. In the *in* conformation however, Ser288 points up away from the active site and forms a hydrogen bond with the backbone of residue 177 (Figure 3.5). The two conformations of Ser288 can be distinguished by whether the side chain points up towards the ring finger or down towards the active site and will be designated as *up* or *down* respectively.

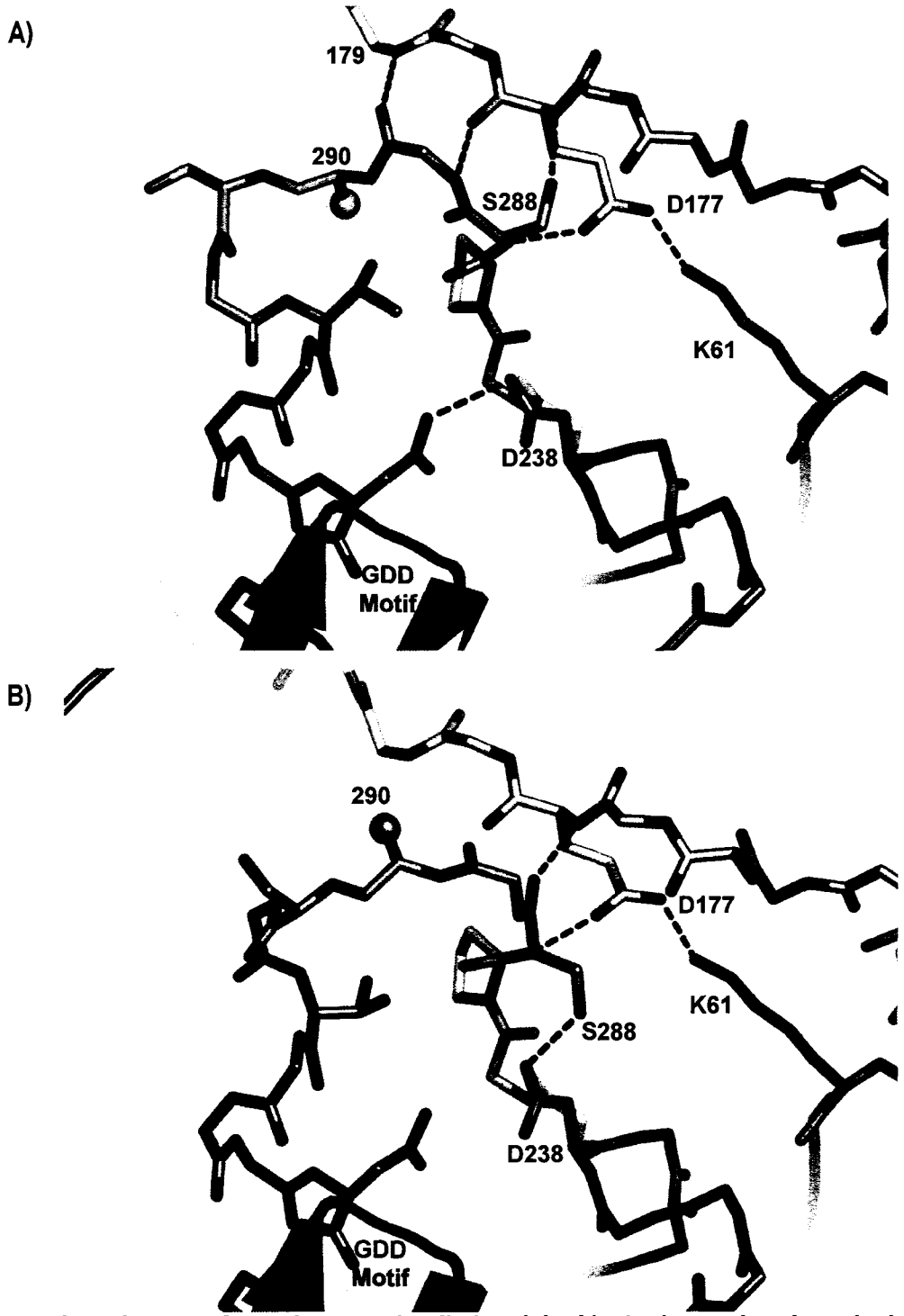


Figure 3.5: ***In*** **and** ***out*** **conformations can be distinguished by hydrogen bonds to the base of the ring finger and the conformation of Ser288. A)** The loop (orange) in the *in* conformation forms four hydrogen bonds with the ring finger (yellow) and the side chain of Ser288 is pointing up away from the active site (magenta) . **B)** The 5.0 Å move of residue 290 resulting in the *out* conformation breaks two hydrogen bonds with the ring finger and, interestingly, the carbonyl of Ser288 takes the place of its side chain, which is now pointing down towards the active site, forming a hydrogen bond with Asp238.

The above observations hold true for the majority of the structures solved, but notable exceptions do exist, resulting in slightly altered *in* or *out* conformations as will be discussed in sections below.

Loop Conformation does not Correlate with Activity

To analyze the biological significance of the loop, the enzymatic activity of each mutant polymerase was tested using an oligo-dT/polyA extension assay and reported as a percentage of the activity of the native polymerase. Activities measured ranged from ~0%-320%, clearly showing the importance of this loop for the overall mechanism of the polymerase (Table 3.1). Coupled with the structures, the activity results also demonstrate that there is not a strict correlation between the loop conformation observed in the crystal structure and the polymerase's ability to synthesize RNA from a primed template. For instance, mutants containing either a Phe or a Glu at residue 290 (C290F and C290E) both have loops described by the *out* conformation, yet display strikingly different elongation activities (Figure 3.6). Substitution of residue 290 with a Phe leads to an approximate 2.5-fold decrease in activity, whereas substitution with a Glu completely abolishes activity (Table 3.1). A similar correlation can be made for the *in* conformation where substitution of residue 290 with a Ser (C290S), a mutation that changes only a single atom, leads to a ~5-fold reduction in activity and excellent density describing the canonical *in* conformation. On the other hand, substitution of residue 290 with a Ser accompanied by a separate mutation changing Gly289 to an alanine (G289A/C290S) leads to a nearly identical loop conformation as C290S alone, yet this double mutation completely abolishes activity (Figure 3.6) (Table 3.1).

Table 3.1-Polymerase extension activities and RNA binding affinity for Loop mutants

Protein	% Activity	RNA binding Affinity (μM)
3DDD	100 ± 3	1.71 ± 0.06
3DDD C290D	1.0 ± 0.9	4.2 ± 0.6
3DDD C290E	0.9 ± 0.7	8 ± 1
3DDD C290F	39 ± 4	1.42 ± 0.09
3DDD C290I	170 ± 30	1.8 ± 0.10
3DDD C290M	21 ± 5	1.61 ± 0.03
3DDD C290S	20 ± 4	1.21 ± 0.04
3DDD C290V	320 ± 60	1.56 ± 0.05
3DDD G289A	-0.5 ± 2	2.8 ± 0.10
3DDD G289A/C290F	0.4 ± 0.1	3.25 ± 0.08
3DDD G289A/C290I	0.12 ± .05	2.4 ± 0.10
3DDD G289A/C290M	4.5 ± 0.3	1.73 ± 0.06
3DDD G289A/C290S	1.3 ± 0.2	1.39 ± 0.05
3DDD G289A/C290V	1.8 ± 0.7	3.15 ± 0.08
3DDD S288A	13 ± 5	4.0 ± 0.3
3DDD S291P	2 ± 1	2.39 ± 0.05
3DDD G292A	2 ± 1	7.1 ± 0.4

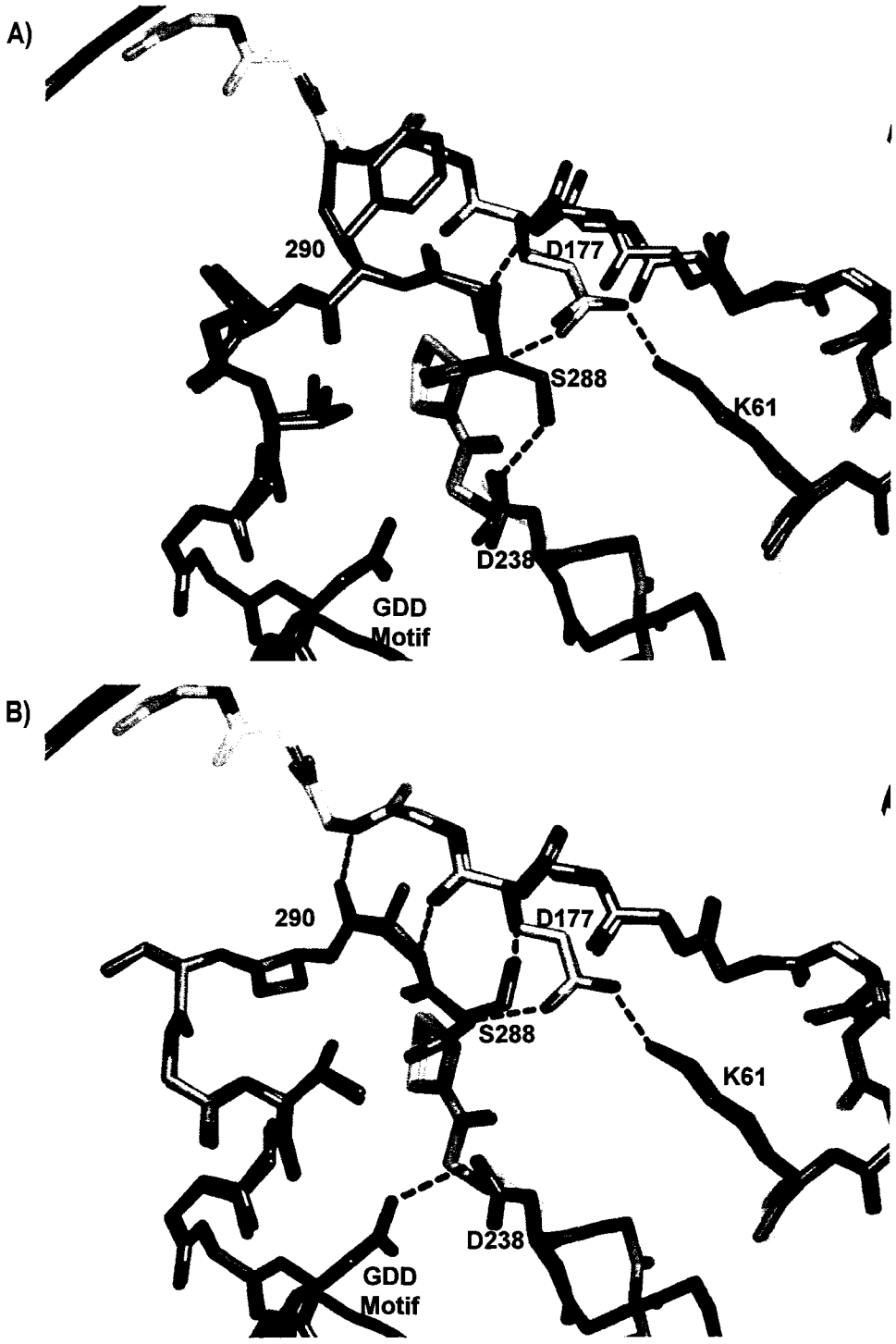


Figure 3.6: Loop activity does not correlate with loop conformation. A) Structural alignment of $3D^{pol}$ C290F (orange loop) with $3D^{pol}$C290E (tan loop) demonstrates identical *out* loop conformations and yet these two polymerases have strikingly different elongation activities. **B)** In accord with the results seen in A) the structures of $3D^{pol}$ C290S (orange loop) and the double mutant $3D^{pol}$ G289A/C290S (tan loop) display identical *in* conformations but have quite different elongation activity.

Activity Affects are not a Result of RNA Binding Defects

To test whether the changes in activity could be attributed to a change in the polymerase's affinity for RNA, the K_d of each polymerase for a defined RNA template was measured using a fluorescence polarization assay (Appendix 1). The results of these experiments indicate that all mutants, with the exception of mutations changing either Cys290 to a glutamate (C290E) or Gly292 to alanine (G292A), bind RNA with similar affinities as native $3D^{pol}$ (Table 3.1). In combination with the activity data, these results imply that mutations in the loop have either disrupted or enhanced the molecular dynamics necessary for catalysis and primer elongation.

To further investigate the activity of each mutant, I have developed a gel assay which allows me to determine whether the enzyme has the ability to add 1, 2 or 3 nucleotides to a short self-priming RNA hairpin template and also gives a qualitative estimation of the rate and fidelity with which it does so (Appendix 1). Interestingly, a large number of mutants tested have the ability to add only one nucleotide, suggesting that the catalytic site is intact and able to carry out phosphoryl transfer, but the enzyme is unable to translocate the product/template strand for the next round of catalysis. In addition, some of the mutant polymerases with activities reported as 0%, as measured by the oligo-dT/polyA extension assay, display the ability to add all three nucleotides but at a much lower rate as compared to the native polymerase (Figure 3.7). These apparently conflicting results are not surprising when considering that a signal in the oligo-dT/polyA extension assay is dependent upon total NTP incorporation on an ill-defined template. Due to the nature of the extension assay, polymerases possessing the ability to elongate long stretches of RNA in a short amount of time will give a large signal whereas kinetically slower polymerases, while showing incorporation of three nucleotides on a gel, simply do not incorporate enough radio-labeled NTP to give a signal in the extension assay.

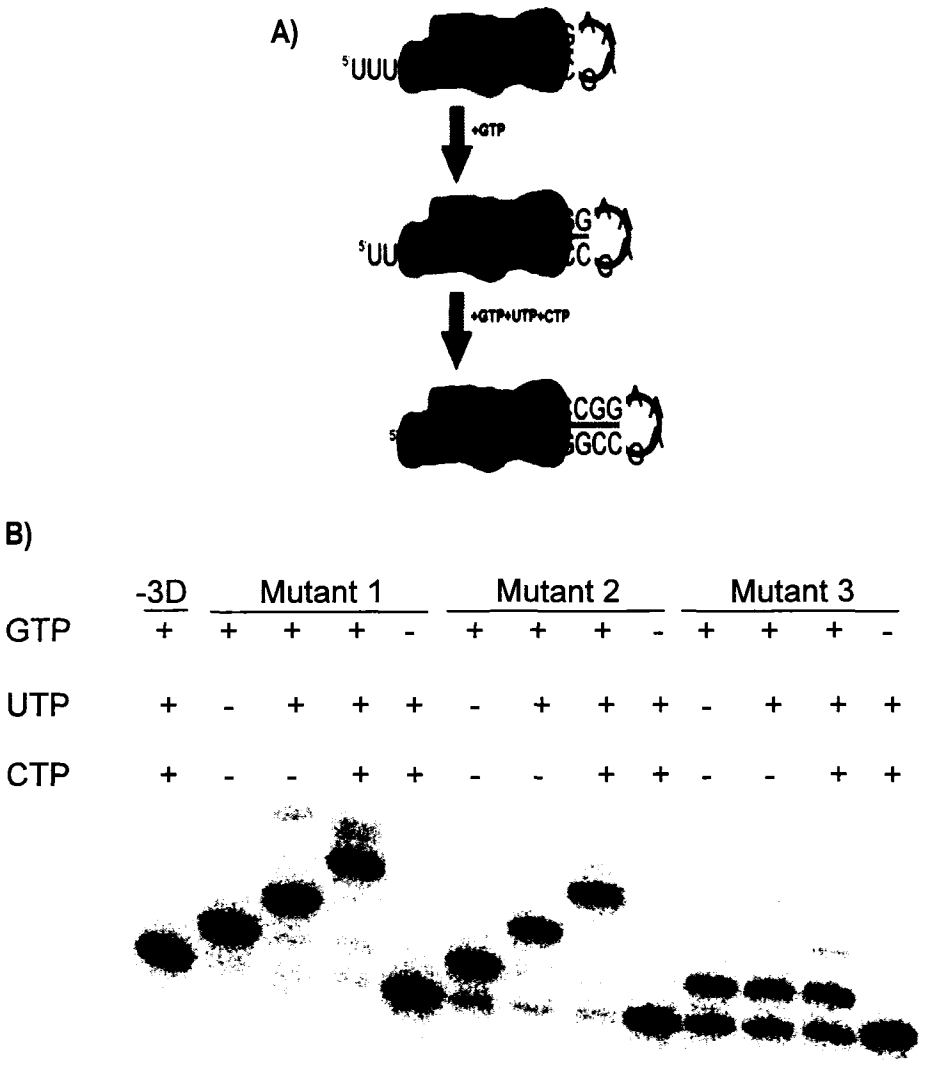


Figure 3.7: **Gel electrophoresis translocation assay. A)** The RNA template used in this assay consists of 26 bases of which 20 form an RNA hairpin resulting in a six base overhang. The first template base encodes a guanine followed by a uridine, cytosine and finally adenine. **B)** 2.5 μM mutant polymerase is mixed with 500nM 5'-end labeled RNA template and 25μM of each NTP. The reaction is incubated at room temperature for 45 minutes resulting in the reaction shown in A). Two μl of the each sample is then run over a 20% polyacrylamide gel containing 7M urea. As shown in this sample gel the addition of each nucleotide to the first two mutant polymerases results in an RNA strand one, two and three nucleotides longer than the initial template. If the initial guanine is left out of the sample, nucleotide addition does not occur. The third mutant is an example of a translocation deficient polymerase in that even when mixed with all three nucleotides the polymerase is still only capable of making a product one nucleotide longer than the original template.

GTP Binding Effects

In addition to apo-structures, the structure of each mutant polymerase was solved in complex with GTP. GTP was bound in exactly the same conformation as previously described having multiple interactions with basic residues in the ring finger (Thompson and Peersen, 2004). This GTP is in a pre-catalytic conformation in which its base pairing face is hydrogen-bonded to the back-bone of residue 175 and the phosphates are not positioned over the active site (Figure 3.8). Overall, the binding of GTP to mutant polymerases does not cause any notable conformational changes in the loop, with the majority of GTP-bound structures having exactly the same backbone geometry as the apo-structure. A notable exception to this observation is a small shift in the *out* conformation loops. In these cases the loop is offset slightly and moved away from the NTP. This is likely due to an interaction between the 2′-OH, Asp238 and Ser288 that form a small hydrogen-bonded network.

The density describing GTP binding in each mutant polymerase ranges from excellent to very poor (Figure 3.8). This may be an indication of the polymerase's affinity for GTP, which may be affected by the various mutations. It should be noted that the quality of the density for GTP does not correlate with activity of the mutants. In many cases, GTP density will be excellent while the mutant displays little or no activity, or GTP density is poor while the mutant displays relatively small changes in activity. These results suggest that NTP-binding defects are not responsible for the observed effects on activity, although it is impossible to rule out this possibility without determining NTP-binding constants for each mutant.

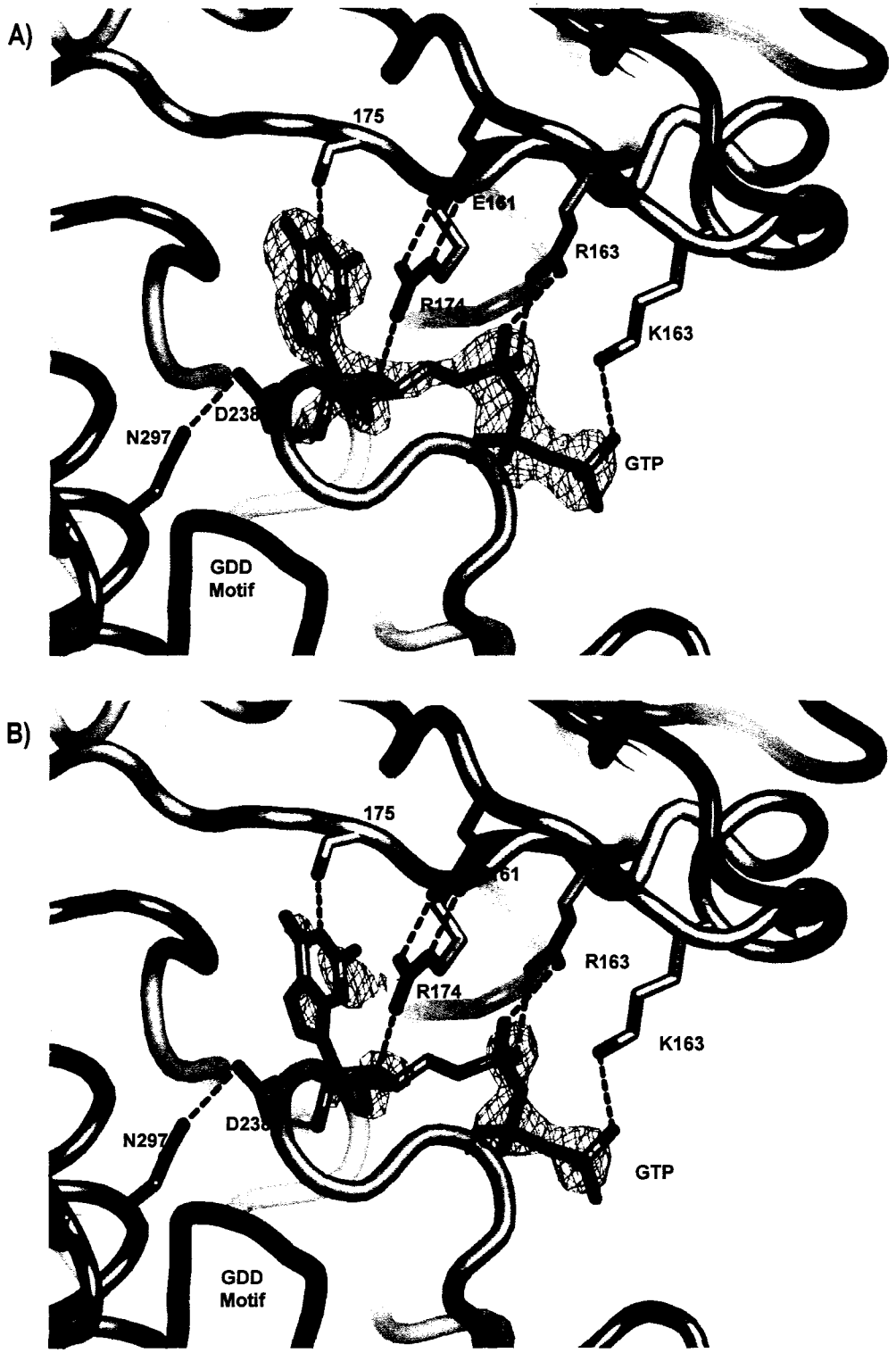


Figure 3.8: GTP 1000K composite simulated anneal omit map contoured at 1.6σ. The quality of the map for GTP in each mutant polymerase ranges from excellent (A) to poor (B) and is found, in all cases, bound in a fashion identical to that reported for the original structure (Thompson and Peersen, 2004 (1RA7)). Hydrogen bonds from the ring finger to the NTP are shown in black and the NTP is in a pre-catalytic conformation.

Small Hydrophobic Substitutions at Residue 290 Lead to a Distorted "in" Loop Conformation

Substitution of residue 290 with either Ile or Val (C290I and C290V) results in an *in* conformation with abnormal 288-292 backbone geometry as compared to native $3D^{pol}$ (Figure 3.9) (Appendix 3 panels 3 and 6). The most significant change in these structures is a 180° flip of Ser288, breaking the hydrogen bond between its side chain and the amide nitrogen of residue 177 in favor of a hydrogen bond between the Ser288 side chain and Asp238. To accommodate this flip, Gly289 undergoes a significant change in phi and psi, angles resulting in the loss of hydrogen bonds between the loop and the ring finger. Interestingly, the hydrogen bond between the amide nitrogen of Ser288 and the side chain of Asp177 has been broken in C290I and C290V, allowing Asp177 to move away from the loop, leading to the loss of a well-ordered salt-bridge with Lys61 in the index finger (Figure 3.9).

In addition to having a significant impact on the backbone geometry of the *in* loop conformation, the Val and Ile mutations lead to drastic increases in the activity of the polymerase. As would be expected, both mutants display the ability to add all three nucleotides to a defined template (Figure 3.10).

Large Hydrophobic Substitution at Residue 290 Lead to Canonical "out" Loop Conformation

To further investigate the significance of a hydrophobic substitution at residue 290, Phe and Met mutations were generated and crystallized (C290F and C290M). The structures of these mutant polymerases reveal a canonical *out* loop conformation characterized by 180° flip of Ser288 (Figure 3.11). The density surrounding each loop, while being strong enough to build a model, is weak compared to surrounding structural elements (Appendix 3 panels 2 and 4). This weak

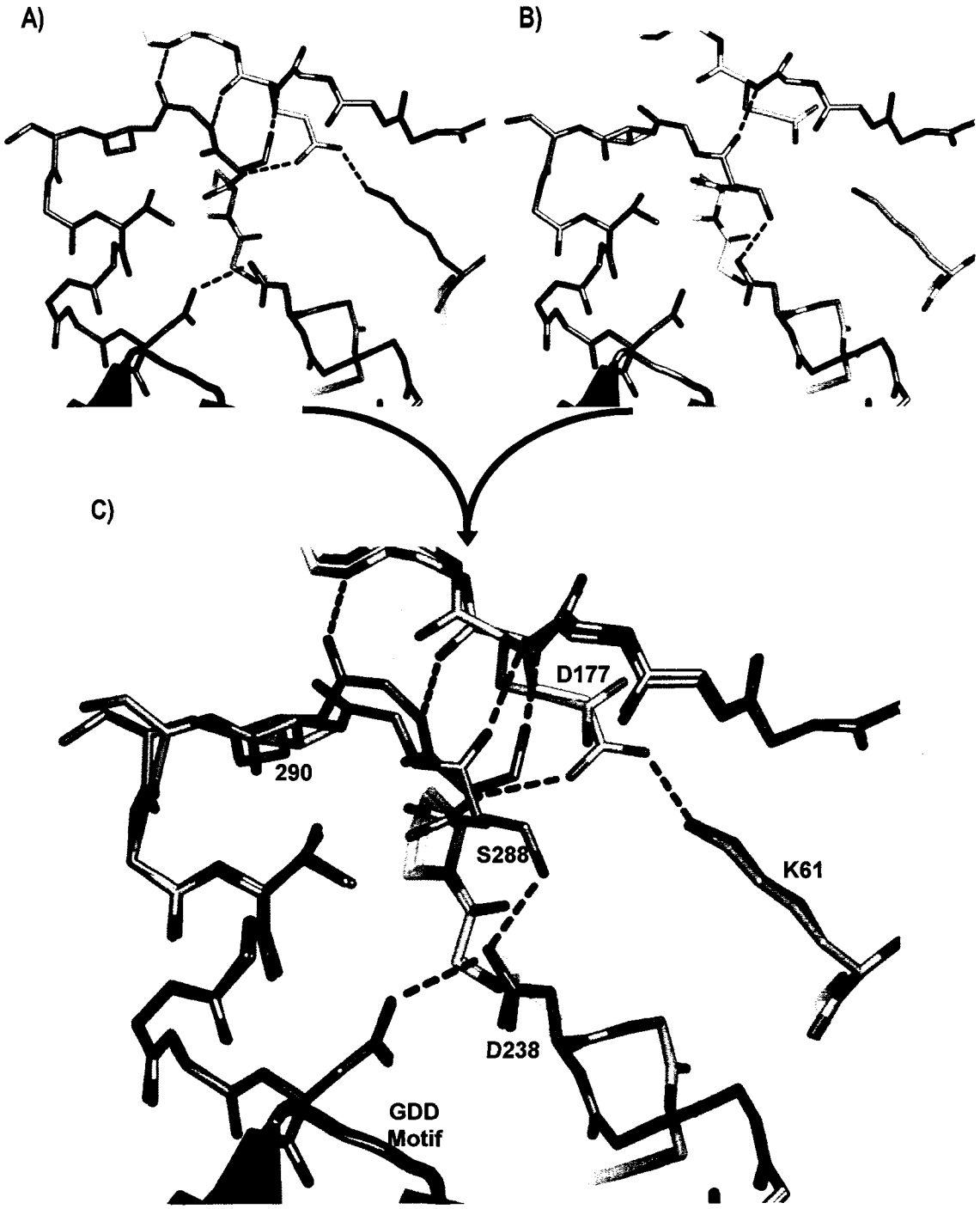


Figure 3.9: 3D^{pol} C290I and 3D^{pol} C290V are in an altered *in* conformation. **A)** Stick representation of 3D^{pol} C290S, which adopts the canonical *in* conformation with four hydrogen bonds between the loop and the ring finger and Ser288 pointing *up*. **B)** Stick representation of 3D^{pol} C290V that has only one hydrogen bond to the ring finger and Ser288 pointing *down* despite having Val290 buried in the hydrophobic pocket. Asp177 has moved away from the loop and Lys61 resulting in the loss of a salt bridge between the ring and index finger. **C)** Structural alignment of 3D^{pol} C290S and 3D^{pol} C290V, all coloring schemes have remained the same as those shown in A) and B). 3D^{pol} C290I and 3D^{pol} C290V backbone geometry is essentially identical and 3D^{pol} C290I is not shown.

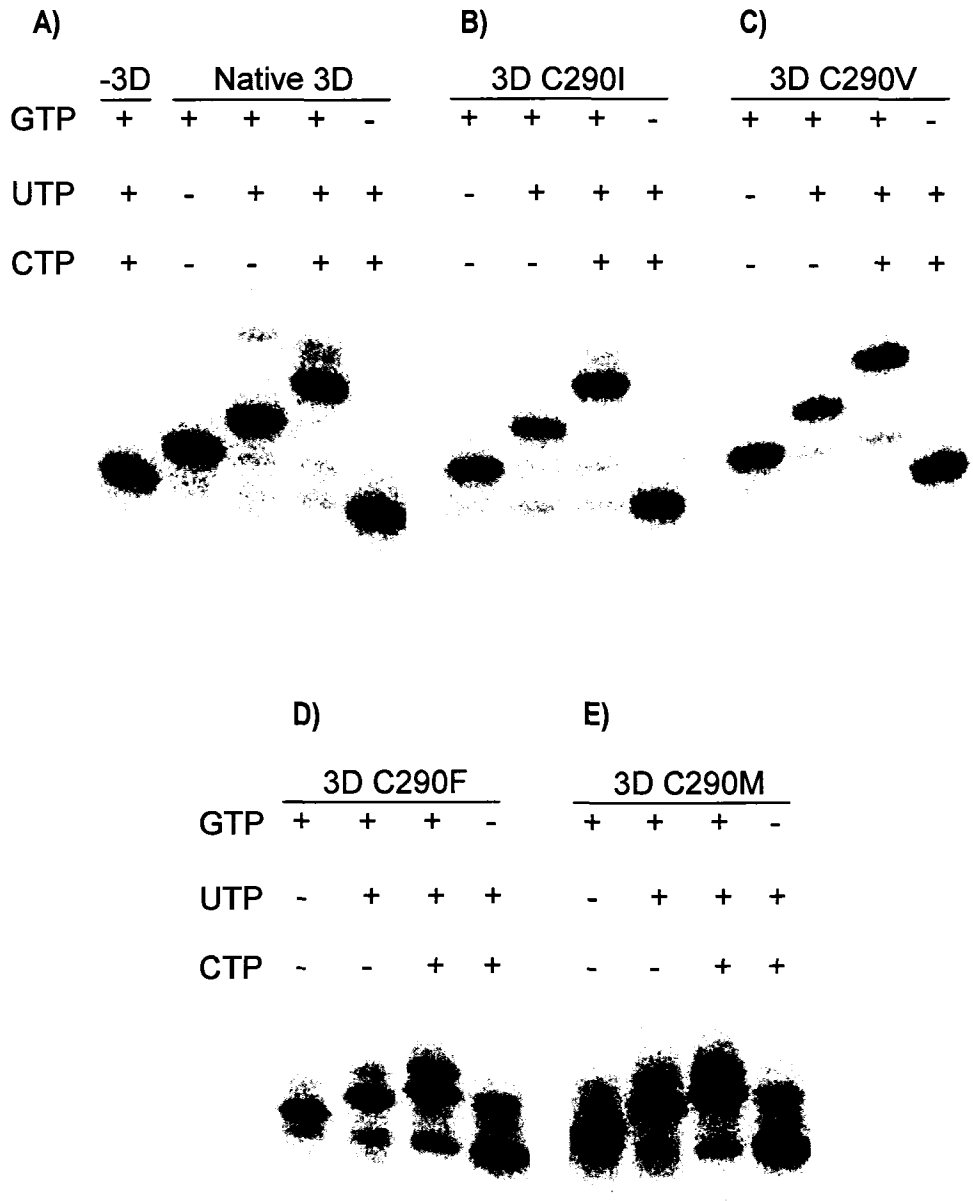


Figure 3.10: **Substitution of residue 290 with a hydrophobic side chain does not prevent translocation**. Gels A)-E) each demonstrates the ability of the mutant polymerase to add three nucleotides to a defined template. $3D^{pol}$ C290I (B) and $3D^{pol}$ C290V (C), like native $3D^{pol}$ (A), efficiently add all three nucleotides as demonstrated by the lack of template left behind following a 45-minute incubation. On the other hand, $3D^{pol}$ C290F (D) and $3D^{pol}$ C290M (E) leave small amounts of template behind, indicating that these polymerase may be kinetically slower. Assays were run as described in 3.7 and Appendix 1.

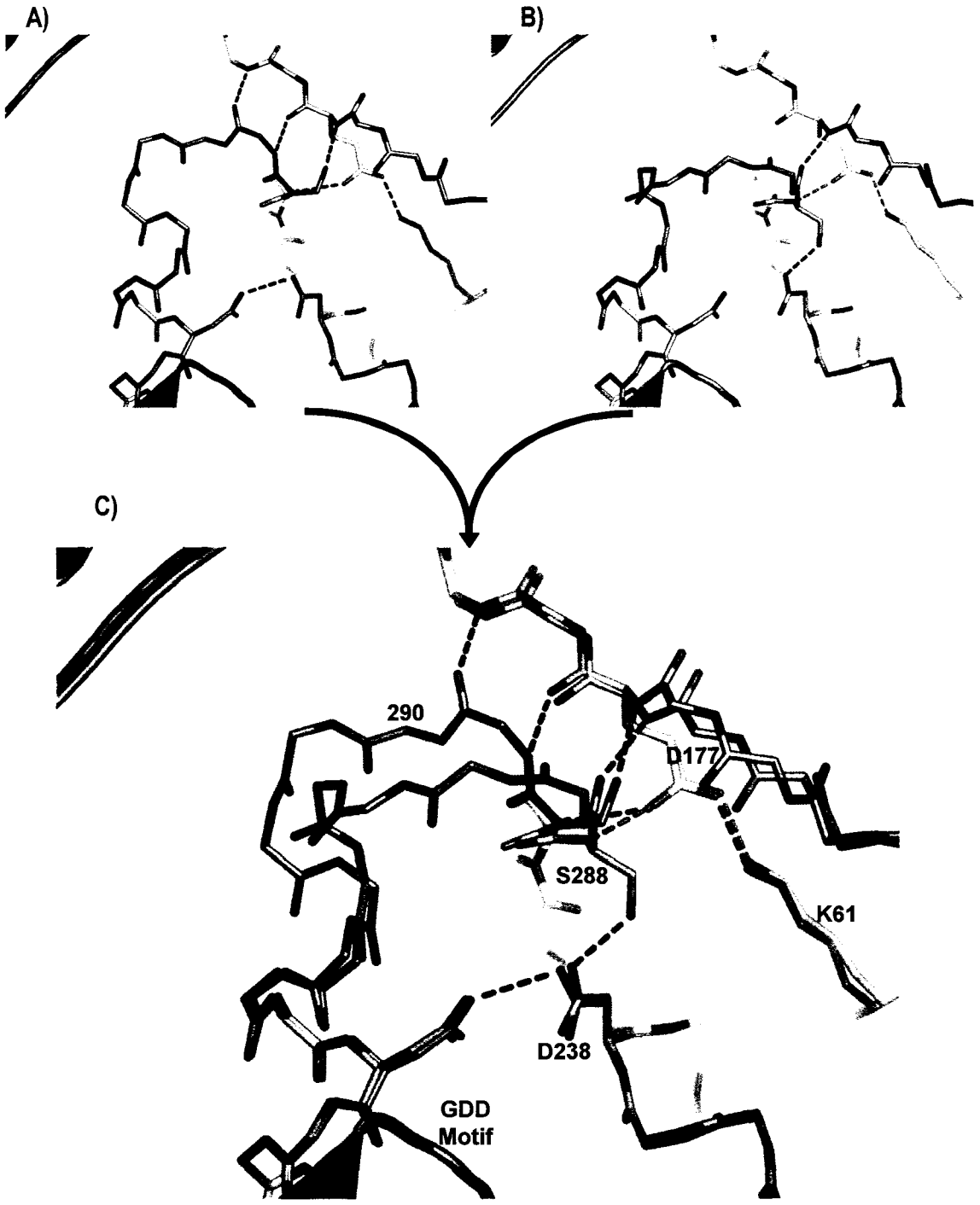


Figure 3.11: 3D^{pol} C290F and 3D^{pol} C290M are in the *out* conformation. **A)** Stick representation of 3D^{pol} C290S, which adopts the canonical *in* conformation with four hydrogen bonds between the loop and the ring finger and Ser288 pointing *up*. **B)** Stick representation of 3D^{pol} C290F that adopts the canonical *out* conformation with Ser288 pointing *down* towards the active site. The loop maintains two hydrogen bonds to the ring finger and the carbonyl of Ser288 takes the place of the serine side chain. **C)** Structural alignment of 3D^{pol} C290S and 3D^{pol} C290F, all coloring schemes have remained the same as those shown in A) and B). 3D^{pol} C290F and 3D^{pol} C290M backbone geometry is essentially identical and 3D^{pol} C290M is not shown. The side chains of the loop, with the exception of Ser288, are not shown.

density, in combination with slightly elevated B-factors, implies greater flexibility within these loops and is suggestive of multiple conformations.

Interestingly, the Phe and Met side chains have strikingly different conformations. The Met extends back toward the hydrophobic pocket, where a portion of its side chain is slightly buried. The Phe, on the other hand points up and away from the hydrophobic pocket, interacting instead with Ile176 leading to a conformational change in the ring finger as compared to native $3D^{pol}$ (Figure 3.12). This small conformational change in the ring finger is not observed in any of the other loop mutants and is a direct result of the presence of a Phe at residue 290. Indeed, binding of GTP to C290F shifts the loop and the Phe side chain away from the ring finger allowing it to come back to its native conformation.

Unlike the hyperactive C290V and C290I polymerases, the substitution of residue 290 with larger hydrophobic side chains leads to decreases in activity with C290F and C290M, exhibiting ~2 and ~4-fold reductions respectively. Each mutant retains the ability to incorporate all three nucleotides but appears to do so at a slower rate as measured qualitatively (Figure 3.10).

Substitution of Residue 290 with Hydrophilic Side chains Leads to Drastic Reductions in Activity

Mutation of Cys290 to Ser (C290S) removed the possibility of a dimethyl-arsenic adduct formation, which requires the presence of a free sulfhydryl group, and the structural complications that were observed in both the native $3D^{pol}$ and Gly64 mutants. It was not expected that the replacement of a single atom within the loop would have drastic effects on the polymerase, and indeed the structure of the C290S mutant is identical to native $3D^{pol}$, described by the canonical *in* conformation (Appendix 3 panel 5). However, despite the essentially identical structure, this mutation resulted in a 5-fold reduction in the activity of the polymerase, indicating the important

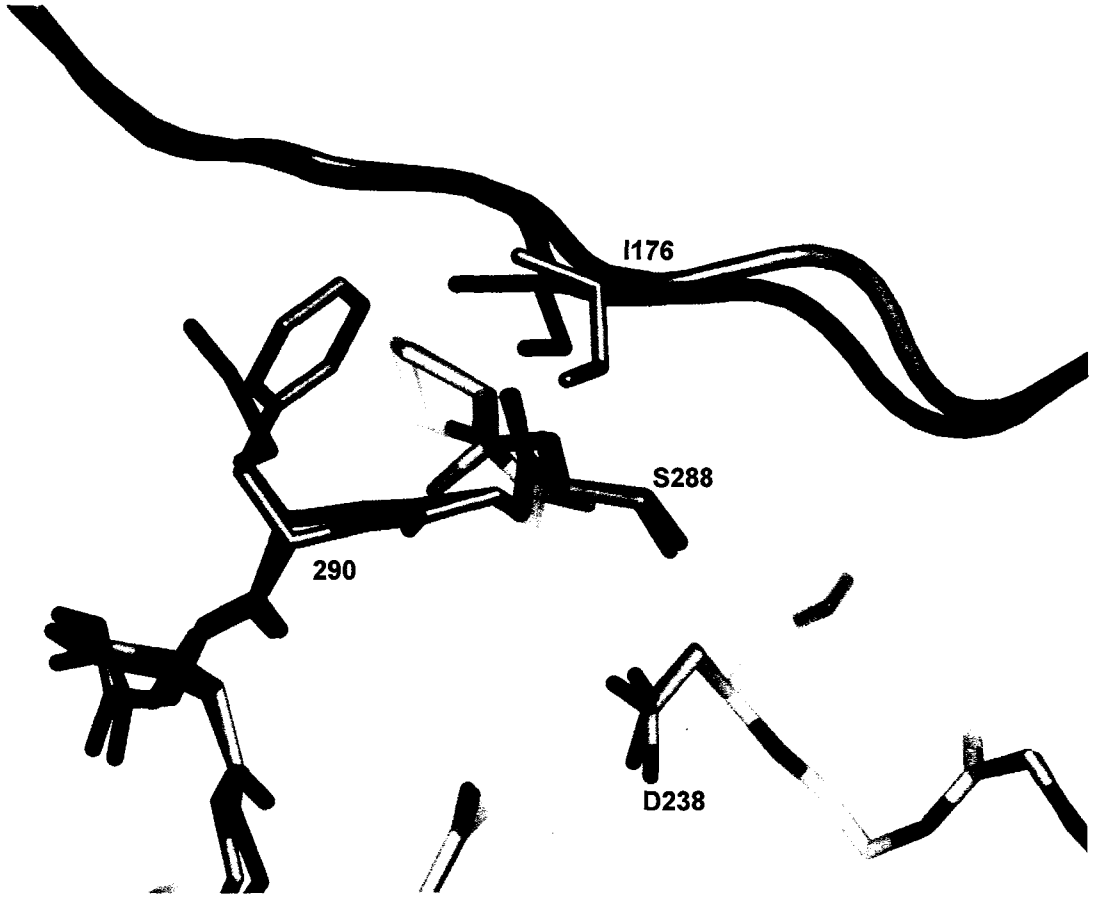


Figure 3.12: 3D^{pol} C290F alters the conformation of the ring finger. 3D^{pol} C290F (orange) and 3D^{pol} C290M (cyan) have very similar backbone geometries. However, the conformation of residue 290 is significantly different, with the methionine pointing back toward the hydrophobic pocket. The phenylalanine side chain forms an interaction with Ile 176 in the ring finger, slightly disrupting its backbone. This interaction is a result of the limited number of rotamers a phenylalanine can adopt and it is expected that the movement in the ring finger is a direct result of the phenylalanine side chain.

functional role of amino acid 290 (Table 3.1). Interestingly, replacing Cys290 with a glutamate (C290E) completely abolishes activity, yielding excellent density for the canonical *out* conformation (Appendix 3 panel 1) (Table 3.1). This mutant, and an aspartic acid substitution that did not crystallize (C290D), have lost the ability to incorporate even a single nucleotide (Figure 3.13).

Gly289 Flexibility is Required for Activity

We hypothesize that the large changes observed in loop conformation for C290V, C290I, C290F, C290M and C290E are dependent upon the flexibility of a Gly at residue 289, allowing for the adoption of the unfavorable phi and psi angles seen in the various structures. To limit this flexibility, Gly289 was converted to an alanine, a residue that limits backbone flexibility without providing significant potential for side chain interactions. The G289A mutation was made in combination with each of the above mutations at Cys290 (Appendix 3 panels 8-12). The combination of the 289 alanine mutation and the various Cys290 mutations leads to the canonical *in* conformations retaining all four hydrogen bonds between the loop and the ring finger. In the case of the Val and Ile 290 substitutions, the salt-bridge between Asp177 and Lys61 has been reformed (Figure 3.14). Interestingly, in all cases the G289A mutation completely abolishes the activity of the polymerase and, with the exception of G289A/C290M, prevents the addition +2 and +3 nucleotides (Figure 3.15) (Table 3.1).

To determine if the observed effects of alanine substitution at residue 289 could be attributed to this residue alone, the mutation was generated in the background of a native Cys290 (G289A). G289A also limits activity to a single nucleotide incorporation event and prevents the addition of +2 and +3 nucleotides, demonstrating the dominant effects this mutation has on the activity of the polymerase (Figure 3.15). The structure of G289A is complicated by the presence of a dimethyl-arsenic adduct on Cys290 resulting in a backbone geometry that cannot be described by either the

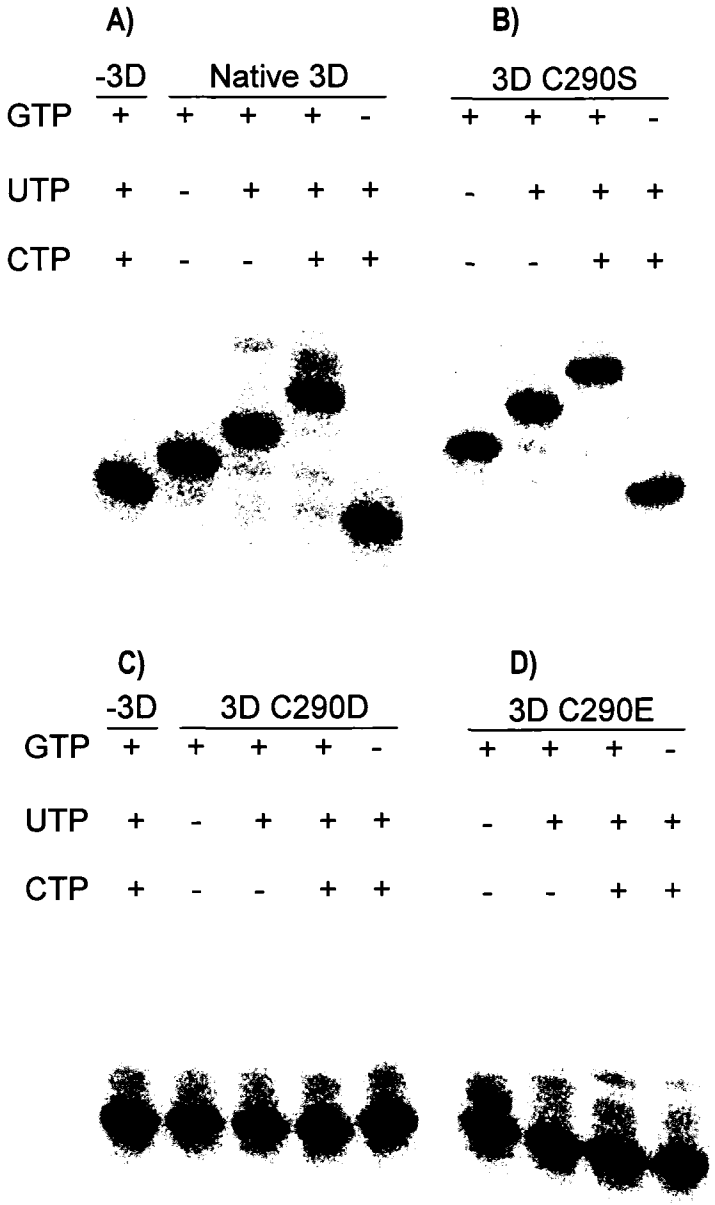


Figure 3.13: Substitution of residue 290 with a charged side chain completely abolishes polymerase activity. Gels A)-D) demonstrate the effect of a polar side chain placed at residue 290. The substitution of Cys290 with a serine (B) does not appear to effect the catalytic ability of the polymerase in any way, yet the elongation activity of the polymerase has decreased 5-fold. Substitution of Cys290 with either aspartate (C) or glutamate (D) affects the elongation activity of the polymerase by completely abolishing the ability of the polymerase to add a single nucleotide. Assays were run as described in Figure 3.7 and Appendix 1.

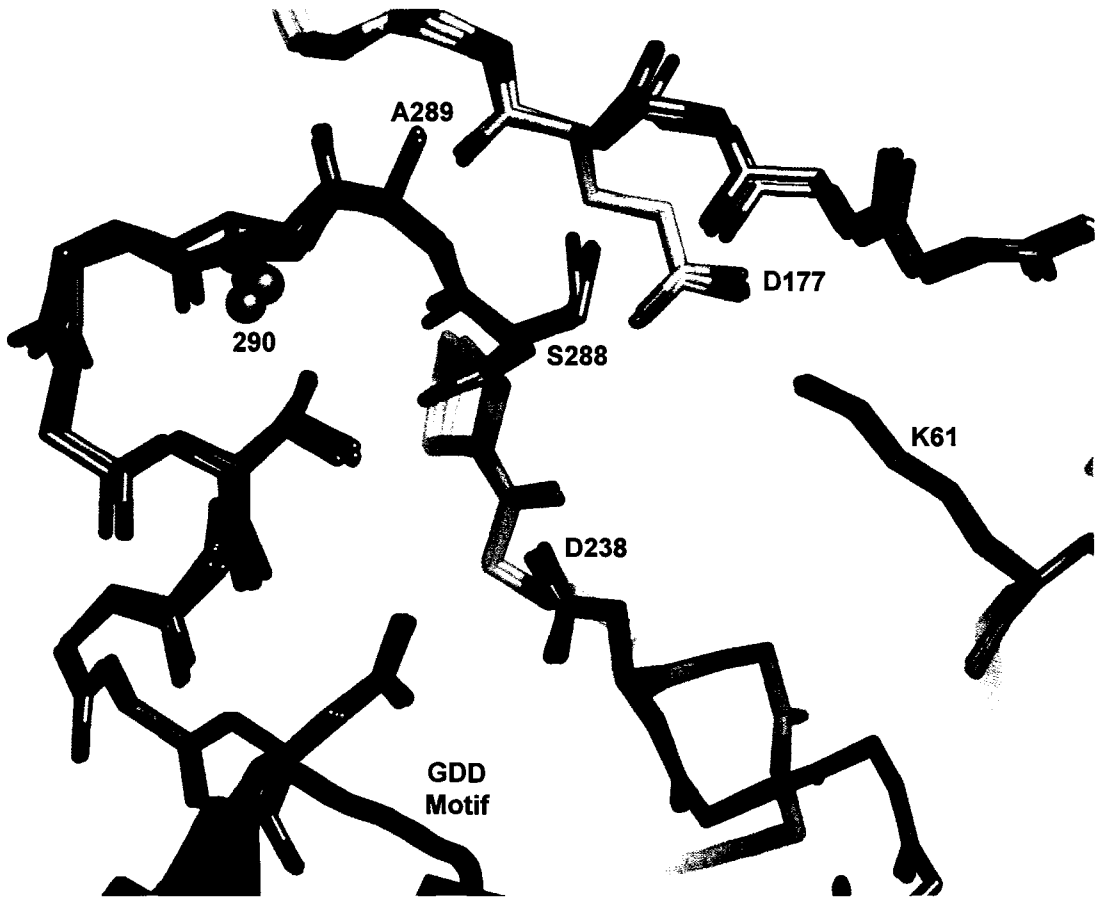


Figure 3.14: Substitution of Gly289 with an alanine results in the canonical *in* conformation. Stick figure of four separate double mutants all containing an alanine at residue 289 and different side chains at residue 290 (represented as a sphere for simplicity). The backbone geometry of each mutant is virtually identical to the others, Ser288 is *up* and residue 290 is tucked into the hydrophobic pocket. The salt bridge between Lys61 and Asp177 is maintained in all structures. Interestingly, single mutations at residue 290 that result in the *out* conformation, such as $3D^{pol}$ C290F, adopt the *in* conformation when combined with an alanine at residue 289.

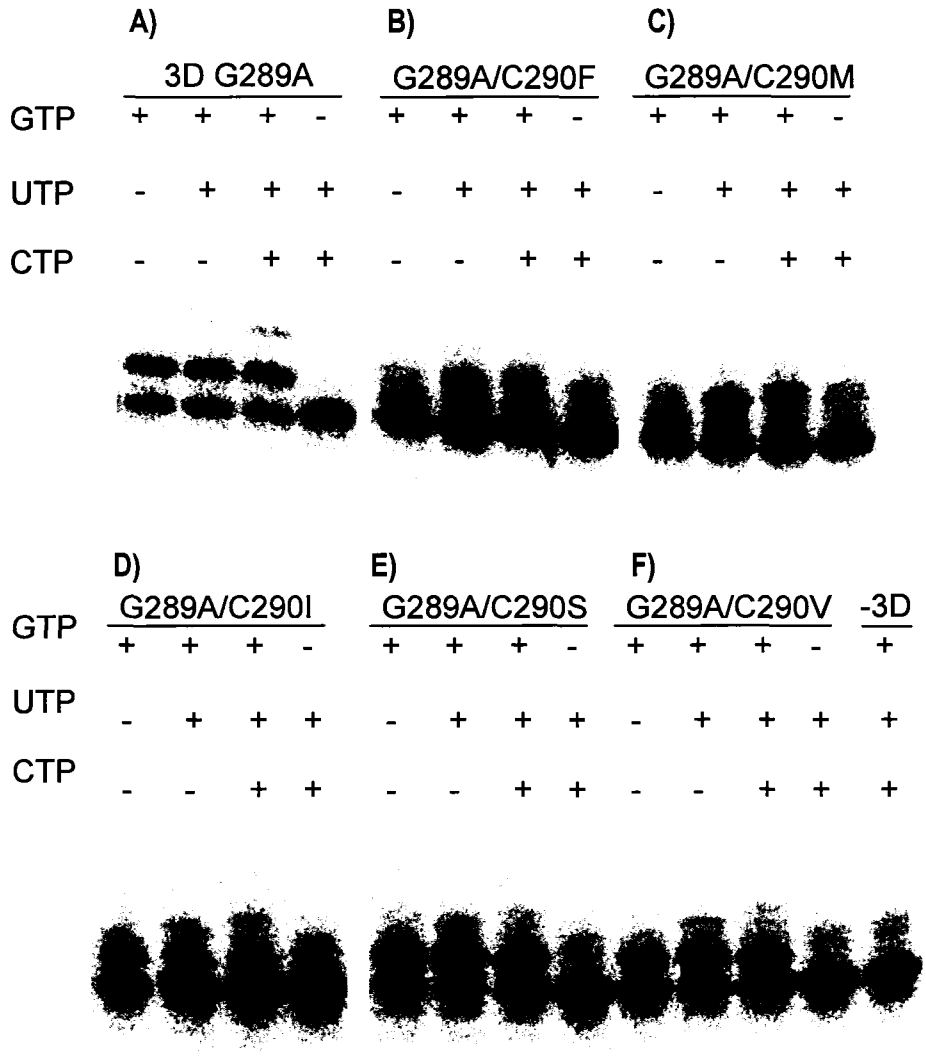


Figure 3.15: **Substitution of Gly289 with an alanine results in translocation defects**. Gels A)-F) demonstrate the inability of mutant polymerases with an alanine at 289 to translocate the product/template strand, resetting the active site for the next round of catalysis. The appearance of a template$_{n+1}$ band indicates that the active site is intact and the enzyme has the ability to position the nucleotide over the active site, albeit at a slower rate. 3D^{pol} G289A/C290M (C) does show faint plus 2 and 3 bands indicating that this mutant has the ability to translocate the nucleic acid following phosphoryl transfer. Assays were run as described in Figure 3.7 and Appendix 1.

in or *out* confirmations (Figure 3.16). The density of the loop is fairly weak and is consequently accompanied by relatively high B-factors as compared to the surrounding structural elements (Appendix 3 panel 7). To confirm whether the loop had been built in the correct conformation, an anomalous difference map was calculated, allowing us to specifically determine the position of the arsenic atom. While this map does indeed confirm the position of the arsenic atom, the electron density is quite weak, suggesting that this adduct adopts several conformations in the crystal lattice (Figure 3.17). The structure of G289A in the presence of GTP leads to a conformational change of the loop which is representative of the canonical *in* conformation and displays no anomalous density surrounding residue 290. Residue 290 now sits nicely in the hydrophobic pocket directly behind the loop and Tyr267 has not been moved to accommodate the large dimethyl-arsenic adduct as in the native $3D^{pol}$. These results, in combination with fairly strong electron density, suggest that the dimethyl-arsenic adduct was either never formed or has been removed due to the soaking conditions needed for GTP binding (Figure 3.18).

Mutations Designed to Prevent Loop Flexibility Abolish Activity

To further investigate the importance of loop flexibility, both Ser291 and Gly292 at the C-terminal end of the loop were mutated with proline (S291P) and alanine (G292A) respectively (Appendix 3 panels 14-15). The S291P mutation was originally made by the Kirkegaard lab in an effort to disrupt 2° structure and it turns out to be a dominant negative mutation in the context of the whole virus (Crowder and Kirkegaard, 2005). Interestingly, both mutations completely abolished activity in the oligo-dT/polyA extension assay, yet retained the ability to add 3 nucleotides to a defined template. The presence of a dark template band in each lane indicates that the addition of each nucleotide is significantly less efficient than native $3D^{pol}$, accounting for the drastic decrease in activity as measured by the extension assay (Figure 3.19 and Table 3.1).

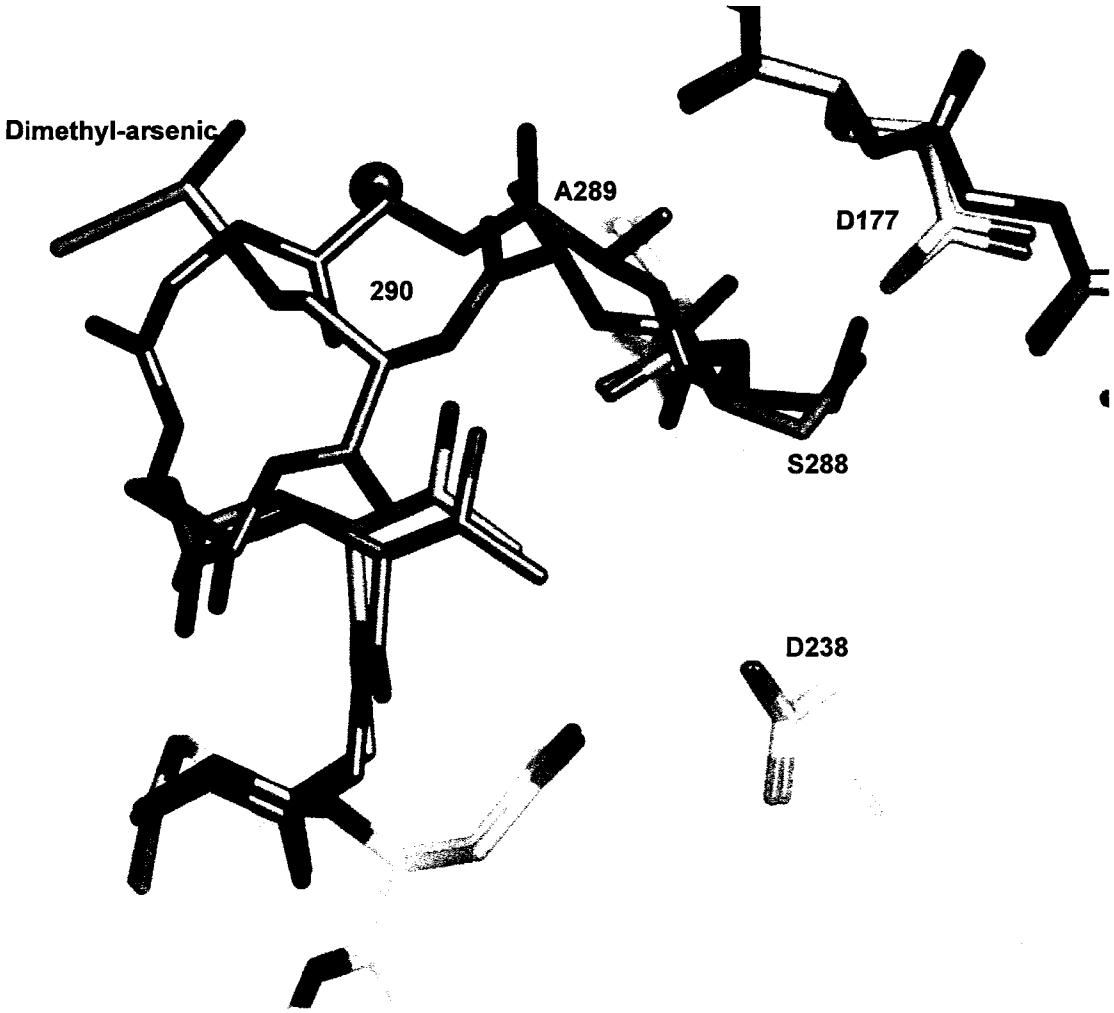


Figure 3.16: The dimethyl-arsenic adduct at Cys290 disrupts the backbone geometry of $3D^{pol}$ G289A. Structural alignment of the $3D^{pol}$ G289A (wheat) with other 289 alanine mutants (gray) demonstrates how the dimethyl-arsenic adduct at Cys290 disrupts the backbone geometry of the mutant. This loop does not adopt the *in* or the *out* conformation, and it appears that the adduct is preventing the adoption of the canonical *in* conformation, forcing residue 290-292 out. It is expected that removal of the modification would result in a canonical *in* conformation.

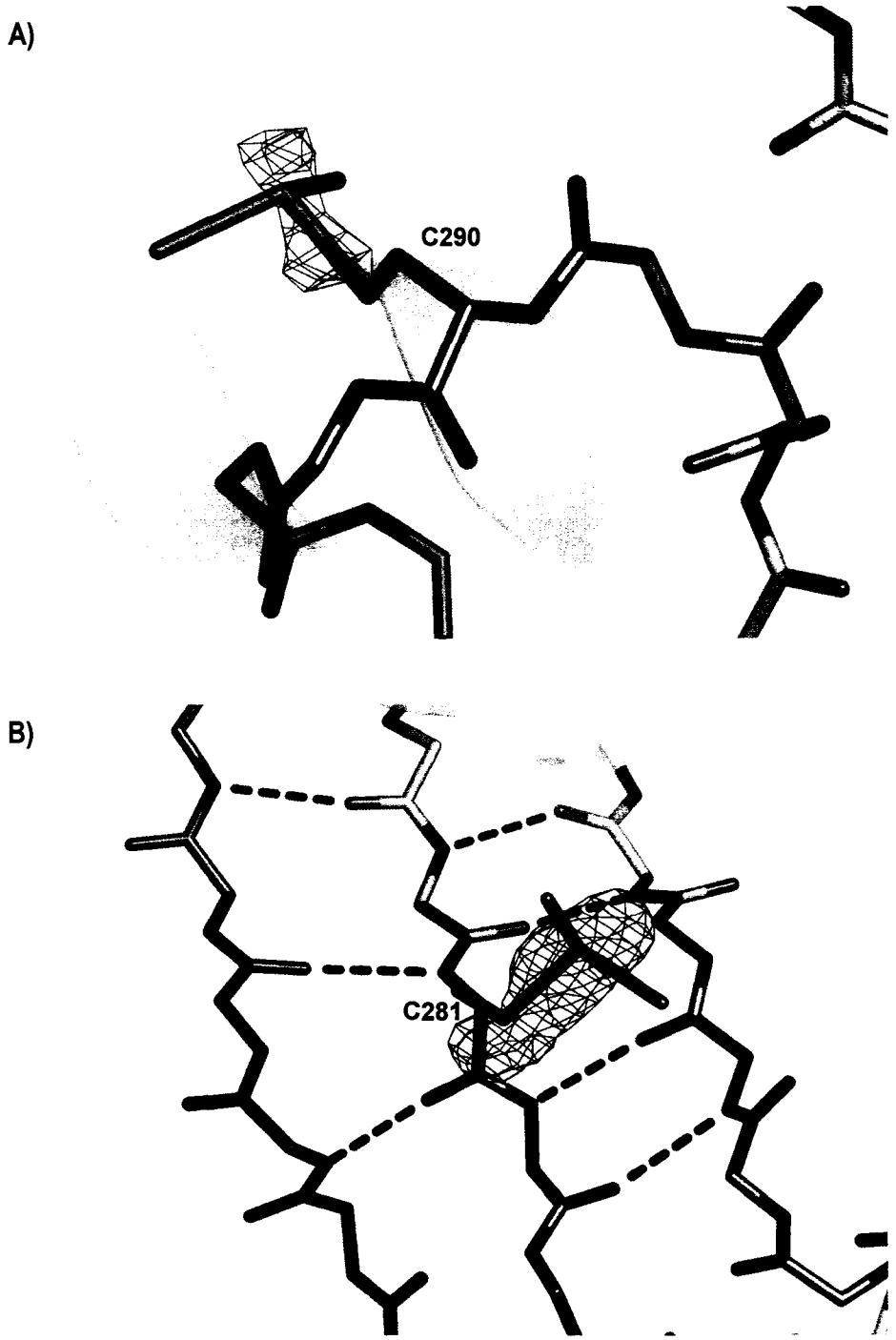


Figure 3.17: Anomalous difference maps of 3D^{pol} G289A contoured at 3.0 σ. The anomalous map describing the position of the arsenic atom at residue 290 (A) is significantly weaker than its counterpart at residue 281 (B) in 3D^{pol} G289A. It has been shown that the dimethyl-arsenic modification at residue 281 is crucial for crystallization (Thompson, 2006) and is therefore expected to have full occupancy. The lack of density for the arsenic atom at 290 indicates that this residue is likely adopting several conformations within the crystal lattice.

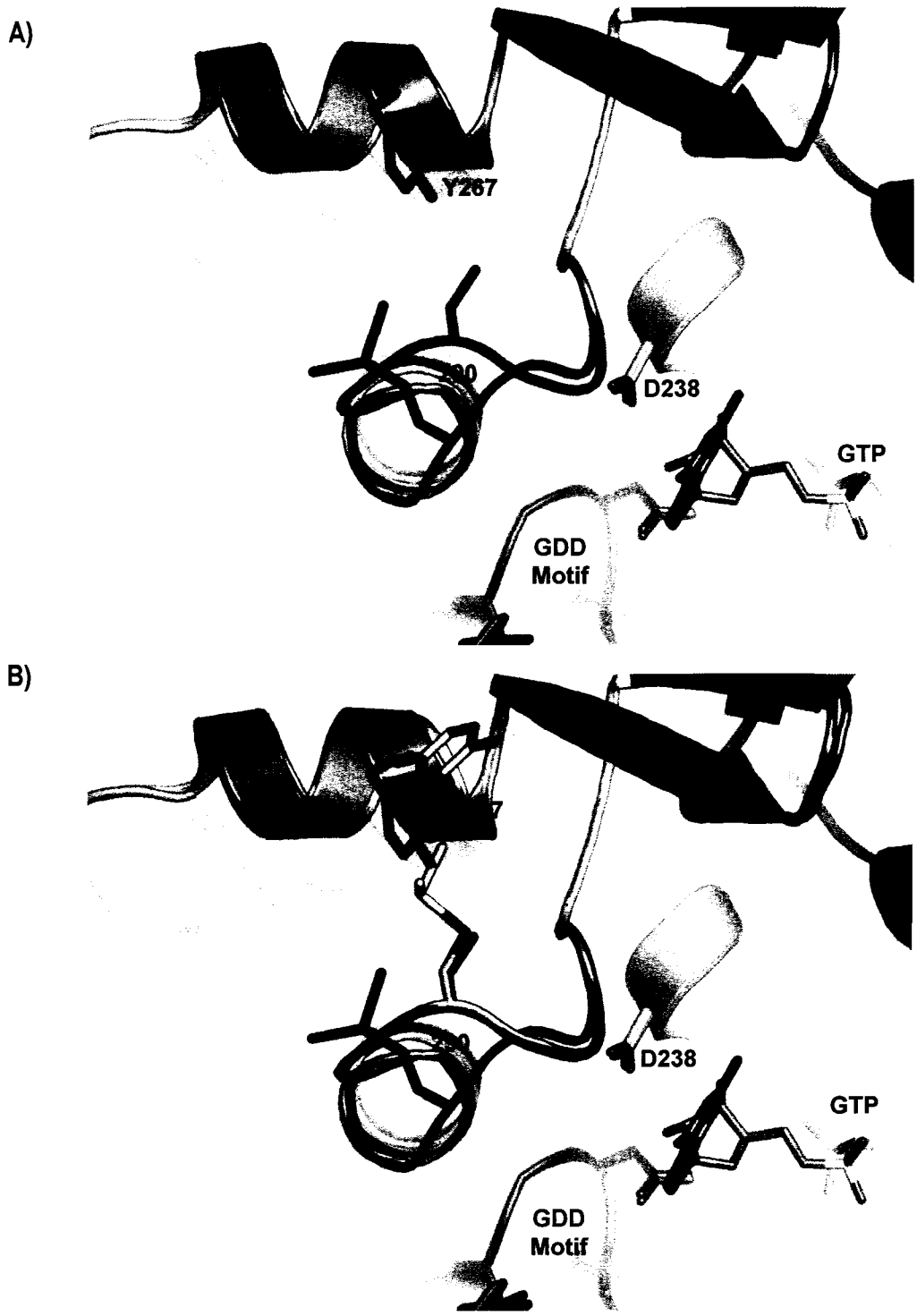


Figure 3.18: GTP binding appears to remove the dimethyl-arsenic adduct in $3D^{pol}$ G289A. A) GTP-bound $3D^{pol}$ G289A (cyan loop) adopts the canonical *in* conformation and lacks anomalous difference density for the arsenic atom on residue 290 found in apo $3D^{pol}$ G289A (black loop). This crystal was taken from the same tray as the apo $3D^{pol}$ G289A and soaked in a solution containing GTP as described in Appendix 1. **B)** Structural alignment of native GTP-bound $3D^{pol}$ (yellow loop) with the 289 mutant shows how Tyr267(color matched with loop) has not moved in the GTP-bound $3D^{pol}$ G289A structure, indicating that the adduct is absent.

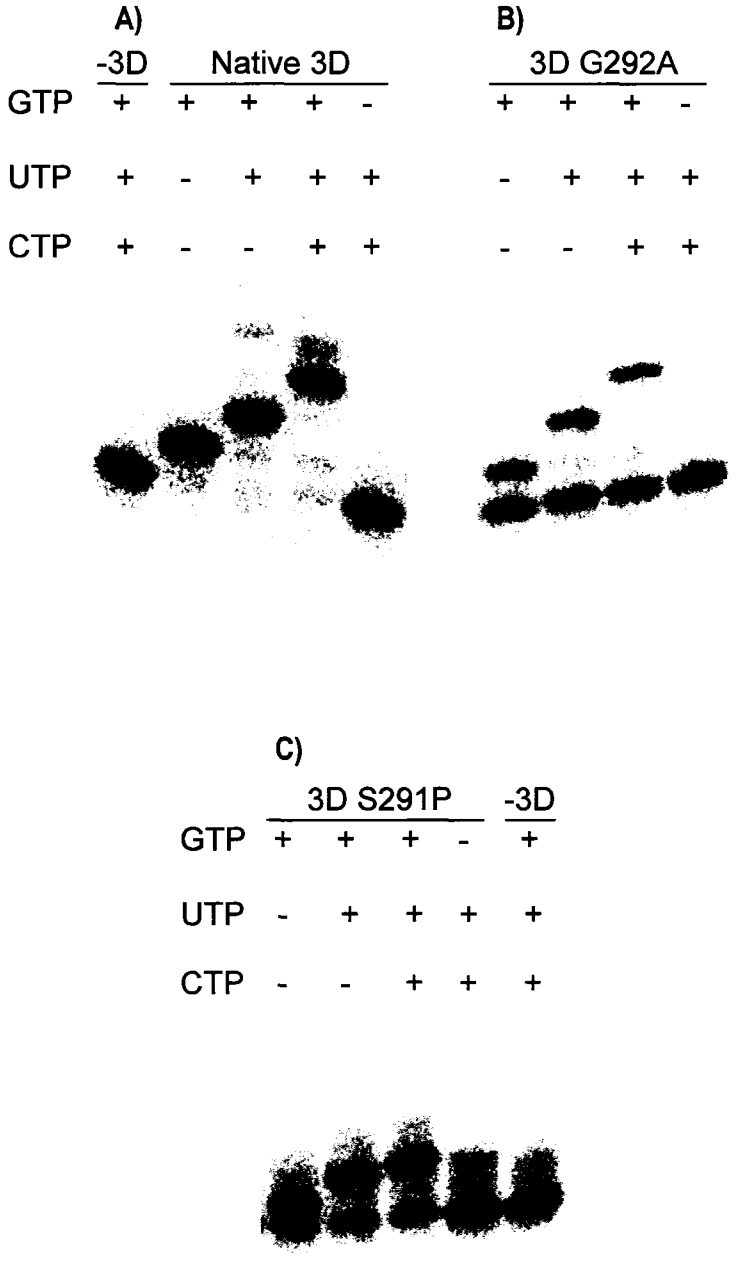


Figure 3.19: 3D^{pol} G292A and 3D^{pol} S291P are kinetically slower than native 3D^{pol}. Gels B) and C) exhibit the ability of 3D^{pol} G292A and 3D^{pol} S291P to add all three nucleotides to a defined template. The presence of leftover template strand in each lane indicates that these enzymes are significantly slower than native 3D^{pol}. Assays were run as described in Figure 3.7 and Appendix 1.

Structures of these polymerases demonstrate that these mutants, in contrast to the G289A mutants, have *out* loop conformations with Ser288 hydrogen-bonded to Asp238. It must be noted that the density describing each loop is quite weak, but improves slightly with the binding of GTP. The weak density can once again be accounted for by complications arising from the dimethyl-arsenic adduct at Cys290 that has α-carbon B-factors ranging from the low 50's to the upper 60's. Anomalous maps were again calculated for each structure and like G289A were quite weak, but in the G292A structures they demonstrated an alternative conformation for the dimethyl-arsenic upon GTP binding which allows the adduct to point into the hydrophobic pocket directly behind the loop (Figure 3.20). There are indications of this conformation in difference maps of the S291P structures. However, anomalous maps do not strongly support them and these conformations were not modeled.

Substitution of Ser288 with an Alanine

The conformation of Ser288 undergoes drastic changes in each of the various sets of loop mutations, allowing its side chain to form hydrogen bonds with either the amide nitrogen of Asp177 or the side chain of Asp238. To determine if the hydrogen-bonding potential of the Ser288 side chain plays an important role for catalysis, an alanine mutation of residue 288 was generated (S288A). This polymerase demonstrated a 10-fold reduction in activity and is capable of incorporating all three nucleotides, albeit at a slower rate (Figure 3.21). The electron density of the loop in both GTP-bound and apo S288A is very weak, making it difficult to determine what conformation the loop is adopting (Appendix 3 panel 13). Having said this, it appears that an *out* loop conformation with the side chain of Ala288 pointing down towards Asp238 can be modeled for the apo structure. GTP binding appears to lead to an *in* loop conformation with Ala288 pointing up towards the amide nitrogen of Asp177. Anomalous maps calculated for both structures make it

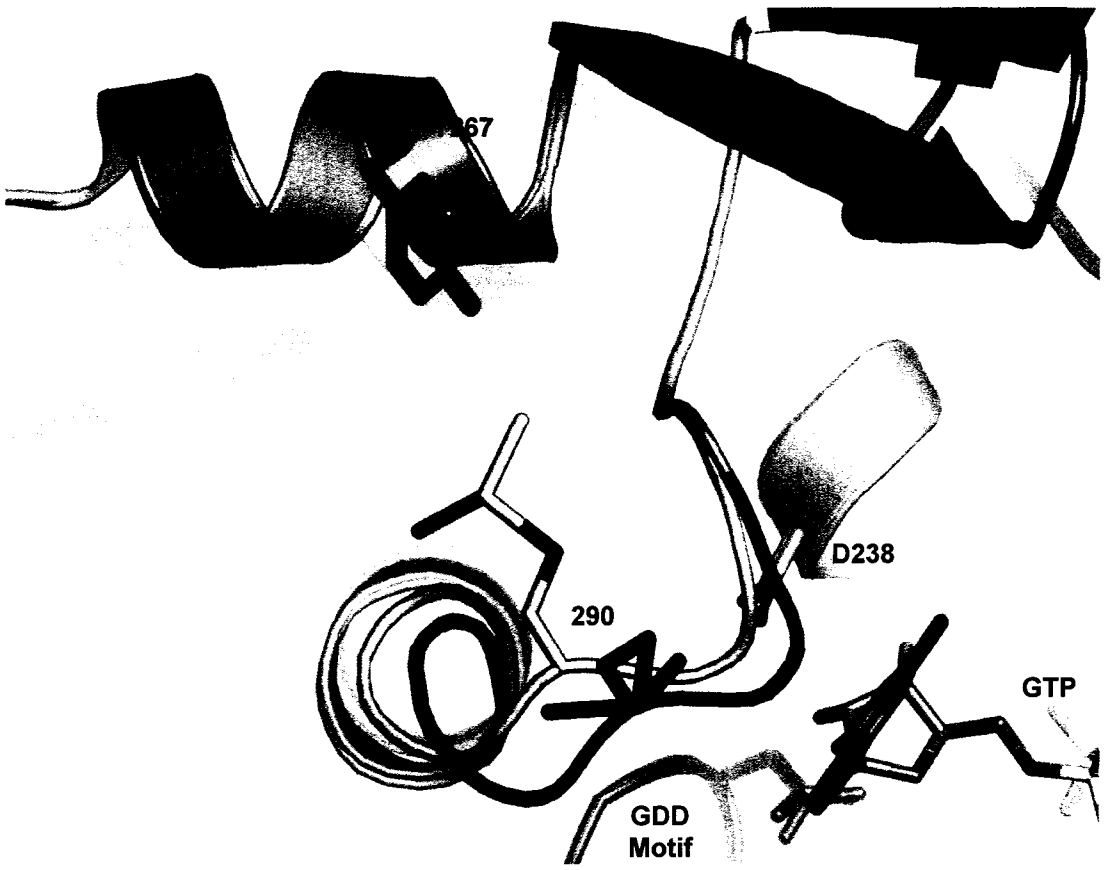


Figure 3.20: GTP binding to 3D^{pol} G292A results in a shift of the loop and the dimethyl-arsenic adduct. The apo 3D^{pol} G292A (cyan loop) adopts the canonical *out* conformation with the adduct sticking straight up, exposed to the surrounding solution. Anomalous and 2Fo-Fc difference maps suggest a slightly different conformation for the loop in the presence of GTP (yellow loop). This loop cannot be described by either the *in* or *out* conformations and is significantly affected by the presence of the adduct.

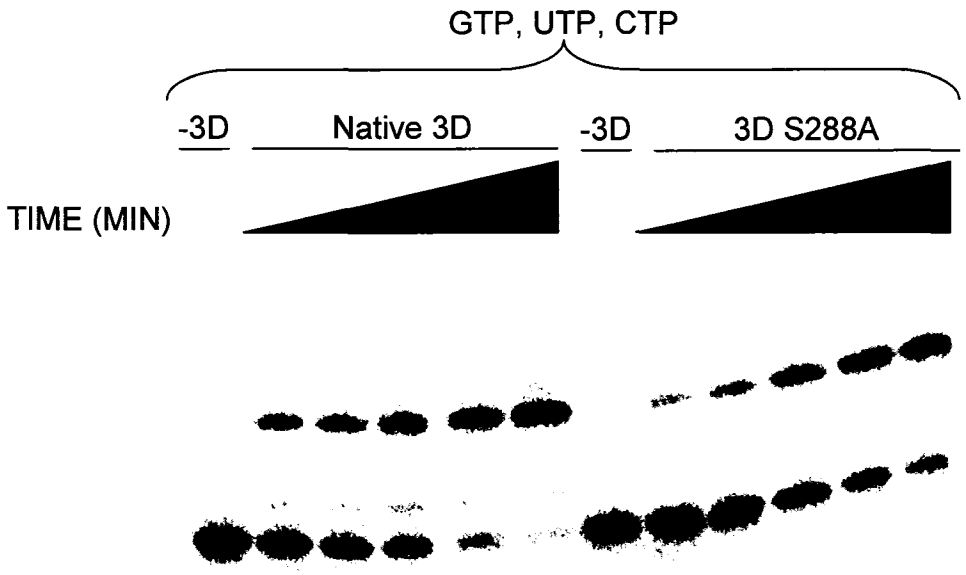


Figure 3.21: **3D^{pol} S288A is kinetically slow**. 2.5 μM 3D^{pol} or 3D^{pol} S288A was mixed with 500nM 5′-end labeled RNA template and 25μM GTP, UTP and CTP. The reaction is quenched at various time points up to 30 minutes, after which 2 μl of the each sample is run over a 20% polyacrylamide gel containing 7M urea. A qualitative assessment of this gel indicates that substitution of Ser288 with an alanine slows the catalytic cycle contributing to the observed effect on elongation. Time points were collected at 2, 5, 10, 20 and 30 minutes following complete mixing of all reactants at room temperature.

difficult to determine the position of the dimethyl-arsenic adduct, suggesting that it has either been removed or is in multiple conformations. The B-factors of α-carbons within these loops range from the lower 50's to the upper 60's, implying that these residues are quite flexible and exist in multiple conformations within the crystal lattice.

Discussion

The 288-292 loop exists in two stable conformations.

We report here the structural and biochemical analysis of 16 different mutants of the 288-292 loop that together suggest there are two conformational switches in this region that are important for $3D^{pol}$ activity. First, Ser288 can undergo a 180° flip down towards the active site, forming a hydrogen bond with the NTP-binding Asp238. This flip is accompanied by a large phi and psi angle change in residue 289 and is dependent on the flexibility of this residue. Second, the entire loop adopts two separate conformations described by the accessibility of Cys290 to a hydrophobic pocket directly behind the loop. This conformational change results in a 5.0 Å move of the α-carbon of residue 290 away from the hydrophobic pocket and towards the active site, resulting in the loss of several hydrogen bonds between the loop and the base of the ring finger. This movement, like the Ser288 flip, is dependent upon the flexibility of a Gly at residue 289 and mutations that restrict the mobility of this residue result in translocation defects. From our results we can safely say that $3D^{pol}$ catalysis is dependent on a Ser-Gly pair at residues 288 and 289 and the nature of the residue at 290.

The Importance of the 288-292 Loop

Our result become quite compelling when considering the structural and sequence homology that is observed for this loop in other RdRP's. Structural comparison of loops of polymerases from

Picornaviridae (Ferrer-Orta et al., 2004; Love et al., 2004), *Caliciviridae* (Ng et al., 2002; Ng et al., 2004), *Flavivirdae* (Ago et al., 1999; Bressanelli et al., 1999; Choi et al., 2004; Lesburg et al., 1999; O'Farrell et al., 2003) and two double-stranded RNA viruses (Butcher et al., 2001; Tao et al., 2002), demonstrate that the 288-289 Ser-Gly pair found in $3D^{pol}$ is a common structural motif within RdRP's. Furthermore, sequence alignments performed by Ann Palmenberg on *Picornaviridae* polymerases demonstrate that this Ser-Gly pair is found in 53 out of 54 of the sequences investigated (Palmenberg, A. and Sgro, JV., 2002). These sequence alignments, in combination with structural homology of RdRP's, indicate that the Ser-Gly pair is a fundamental structural element of RdRP's specifically designed for the proper functioning of the enzyme.

Structural alignments with the other picornaviral RdRP's, which lack the complication of a dimethyl-arsenic adduct at the Cys290 analog, reveal that each polymerase has a loop described by the *in* conformation with Ser288 analogs that can either point *up* and away from the active site or *down* towards the Asp238 analog. If the serine is flipped down, then the adjacent Gly has backbone geometry very similar to the distorted backbone seen in C290I and C290V mutants. This structural feature can be most notably seen when comparing the three structures of polymerase from different strains of human rhinovirus displaying different Ser conformations resulting in different Gly phi and psi angles (Figure 3.22). Interestingly, the binding of RNA to the FMDV RdRP causes the Ser288 analog to flip down, suggesting that the conformation of this Ser may be modulated by the presence or absence of RNA.

Because the RdRP's from poliovirus and FMDV are homologous enzymes, the structure of the FMDV polymerase in complex with an RNA molecule (Ferrer-Orta et al., 2004) can be used to model, by homology, a template-primer RNA molecule onto the active site of $3D^{pol}$ (Figure 3.23). This model, along with other structural and biochemical data, demonstrates the important location of the 288-292 loop with respect to the RNA template and the active site. Indeed, the backbone of

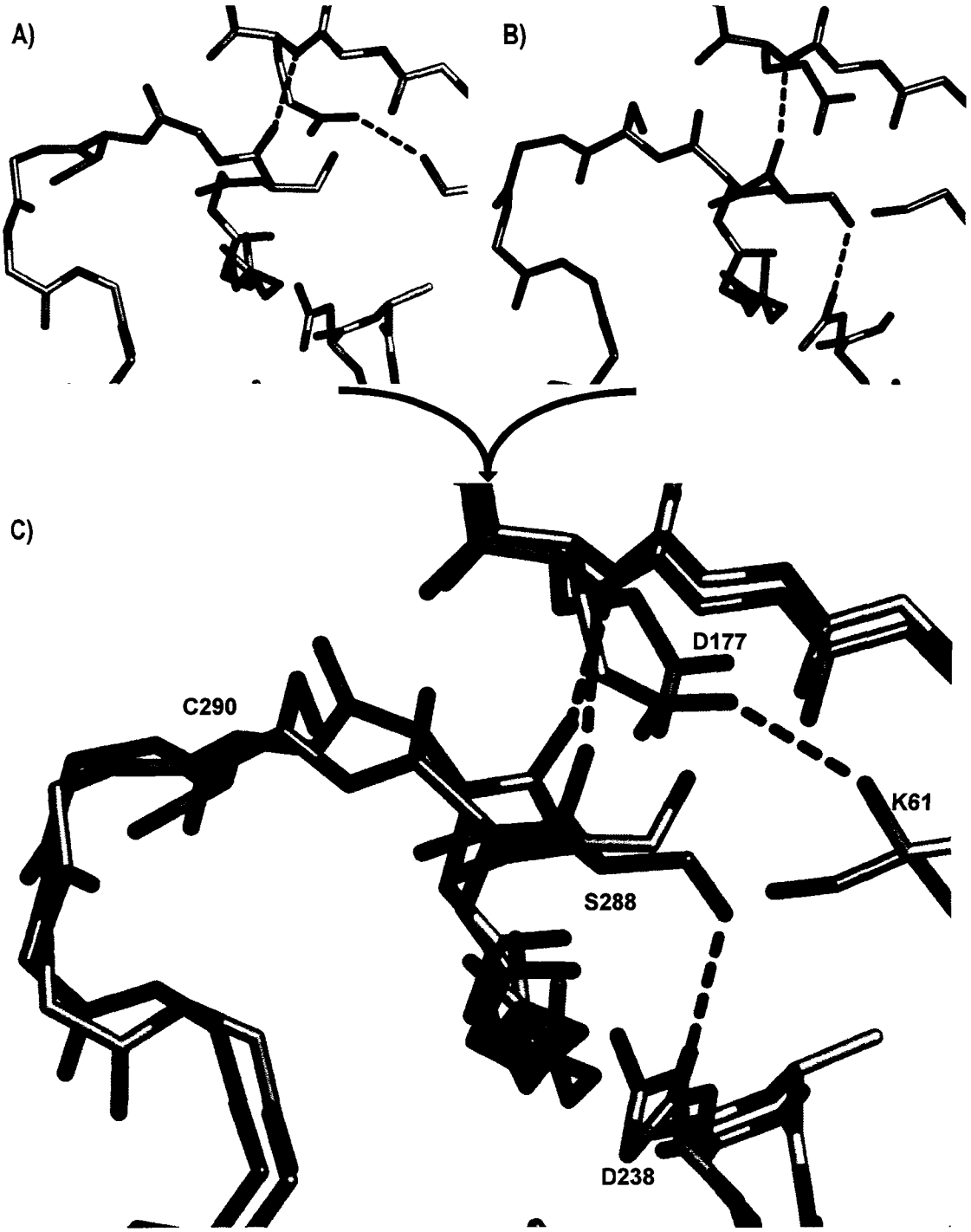


Figure 3.22: Structures of rhinovirus RdRP from different serotypes display an *up* or *down* Ser288 conformation. **A)** The loop of the RdRP from serotype 14 of rhinovirus adopts a conformation very similar to the *in* conformation with Ser288 pointing *up*. **B)** The loop of the RdRP from serotype 1B of rhinovirus adopts a loop conformation very similar to $3D^{pol}$ C290V with Ser288 pointing *down*, despite the burial of Cys290. **C)** Structural alignment of the two RdRP's from rhinovirus, all coloring schemes stayed the same as seen in A) and B). Interestingly, much like $3D^{pol}$ C290V or I, the salt bridge between Lys61 and Asp177 is absent when Ser288 is down and Cys290 remains buried in the hydrophobic pocket (Love et al., 2004 (1XR5 and 1XR6)).

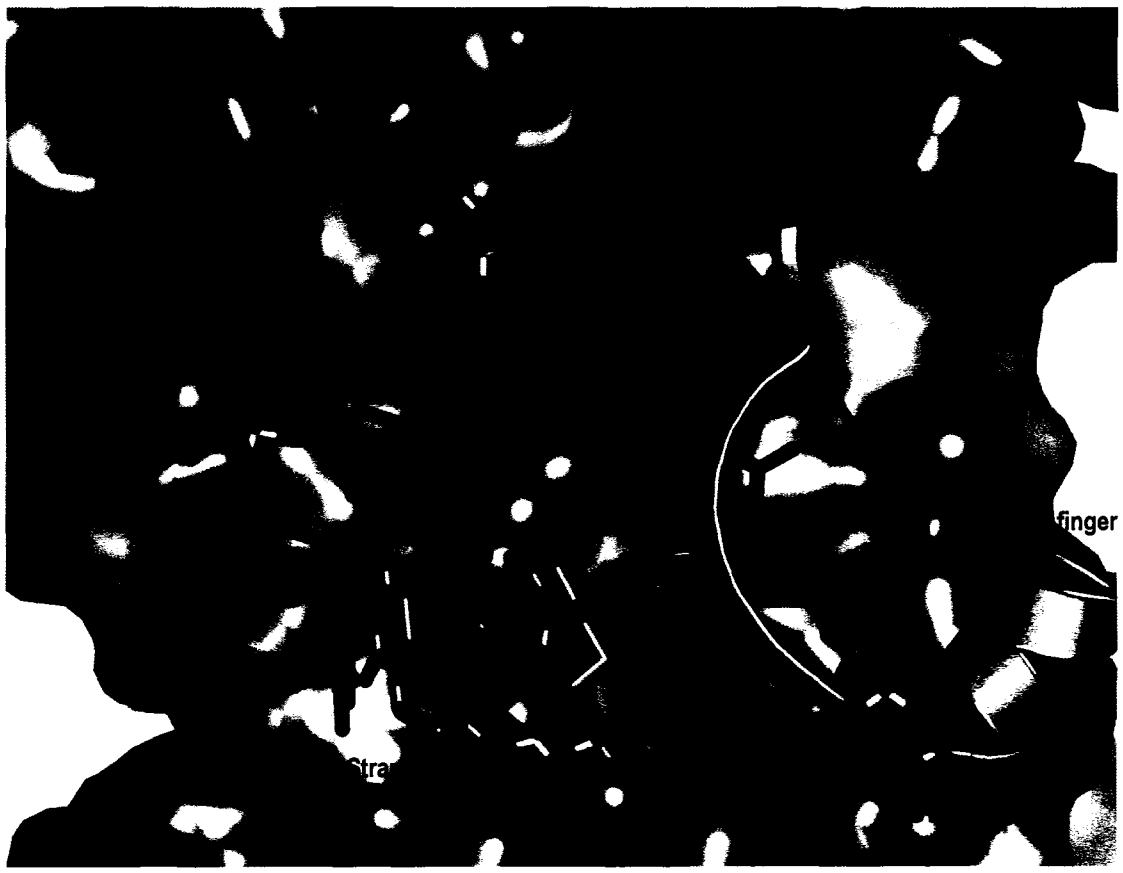


Figure 3.23: **The 288-292 loop is in direct proximity to the template strand**. Surface representation of $3D^{pol}$ with the active site colored in magenta and the 288-292 loop colored in purple. The ring finger (yellow) forms the roof of the NTP entry tunnel and GTP is bound in a pre-catalytic state (Thompson and Peersen, 2004). A duplex RNA molecule from FMDV has been modeled into the active site of $3D^{pol}$ with the template strand colored in green and the product strand colored in orange (Ferrer-Orta, 2004 (1WNE)). The loop is in direct proximity to the template strand where the templating base is colored to match the incoming NTP (tan). Based on location, significant movements in the loop would be expected to affect nucleic acid positioning. The index finger has been removed to reveal NTP positioning.

FMDV Gly302 (which is analogous to Gly292 of PV $3D^{pol}$) forms a hydrogen bond with the template strand at the 2′-OH of the fifth base. This is the first base to form a base pair with the product strand. Thus, the 288-292 loop is in very close proximity to the templating base for the incoming nucleotide.

A Model for Catalysis

As previously mentioned, kinetic studies of polymerases reveal that a single round of catalysis consists of five separate steps (Figure 3.1). Within these steps, two conformational changes flank phosphoryl transfer, the first of which positions the NTP over the active site and the second translocates the nucleic acid. Both of these conformational changes have been observed structurally in the RNA polymerase from bacteriophage T7 (Yin and Steitz, 2004). We feel that the alternate conformations of the 288-292 loop presented here can be correlated with the larger conformational changes needed for NTP positioning and translocation and as such we propose a molecular model for a single round of catalysis that is supported by our data.

We propose that a round of catalysis begins with a flip of Ser288 down towards the active site resulting in a distorted backbone for residues 288 and 289 and is dependent upon the flexibility of a glycine at residue 289. The distortion of the loop backbone causes the salt bridge between Asp177 and Lys61 to break, which we expect to be one in a series of events leading to a large movement in the ring finger that pushes the NTP into the active site (Figure 3.24). The formation of a hydrogen bond between Ser288 and Asp238, as a result of this flip, likely acts to stabilize both the strained backbone geometry of Gly289 and the position of Asp238 during large conformational changes occurring in the fingers domain.

Once phosphoryl transfer has occurred, the template/product strand must be translocated. At this point the loop will undergo a second conformational change that, like the Ser288 flip, is

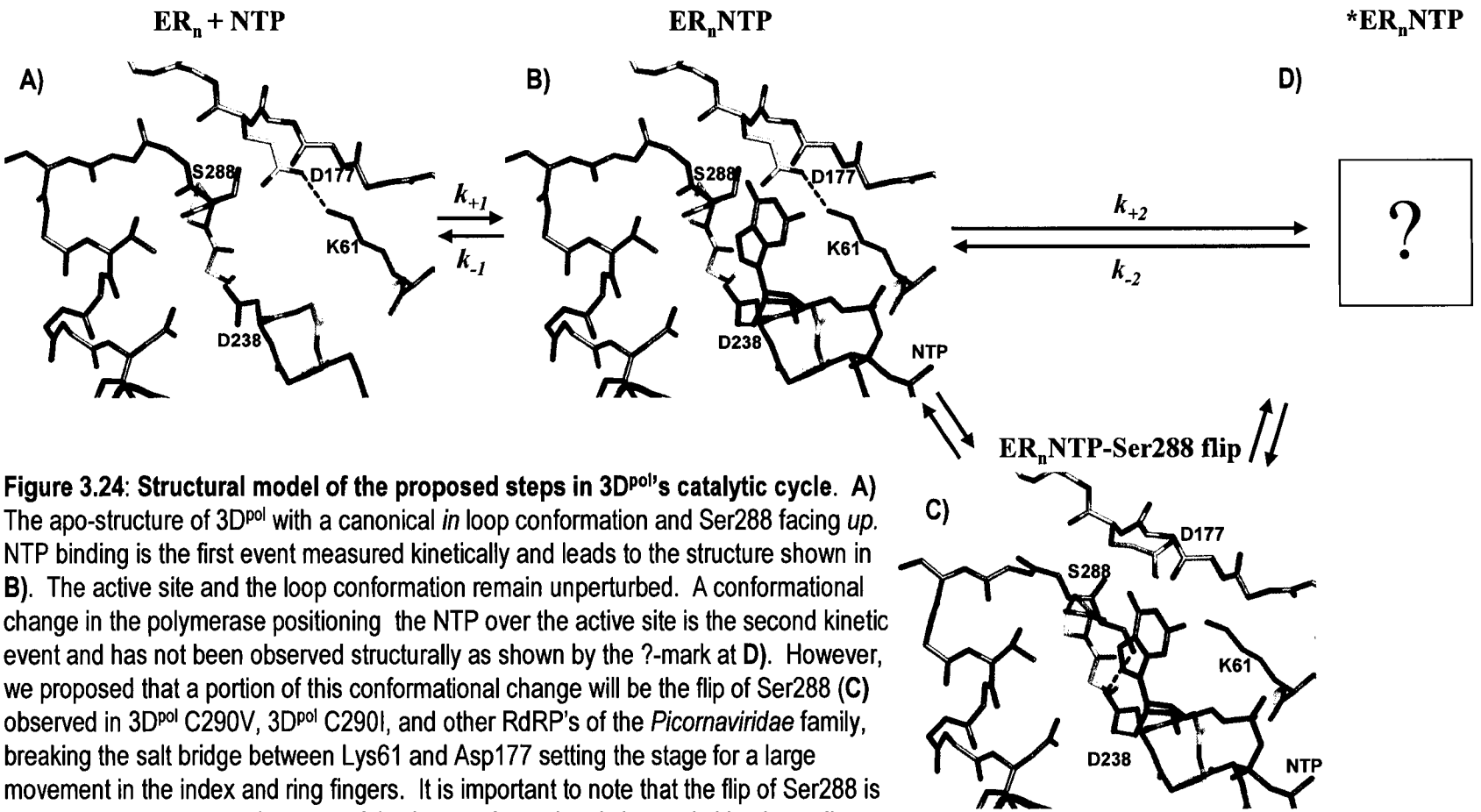


Figure 3.24: Structural model of the proposed steps in 3D^{pol}'s catalytic cycle. A) The apo-structure of 3D^{pol} with a canonical *in* loop conformation and Ser288 facing *up*. NTP binding is the first event measured kinetically and leads to the structure shown in **B)**. The active site and the loop conformation remain unperturbed. A conformational change in the polymerase positioning the NTP over the active site is the second kinetic event and has not been observed structurally as shown by the ?-mark at **D)**. However, we proposed that a portion of this conformational change will be the flip of Ser288 (**C**) observed in 3D^{pol} C290V, 3D^{pol} C290I, and other RdRP's of the *Picornaviridae* family, breaking the salt bridge between Lys61 and Asp177 setting the stage for a large movement in the index and ring fingers. It is important to note that the flip of Ser288 is probably being measured as part of the first conformational change in kinetic studies. The third step (see next page) is phosphoryl transfer **(E)** followed by translocation **(F-G)**. Neither of these steps have been observed structurally. However, we propose that the loop moves from the *in* (**F**) to the *out* **(G)** conformation, leading to translocation. The final step is PPi release, returning the enzyme to a state ready to bind the next NTP **(H)**.

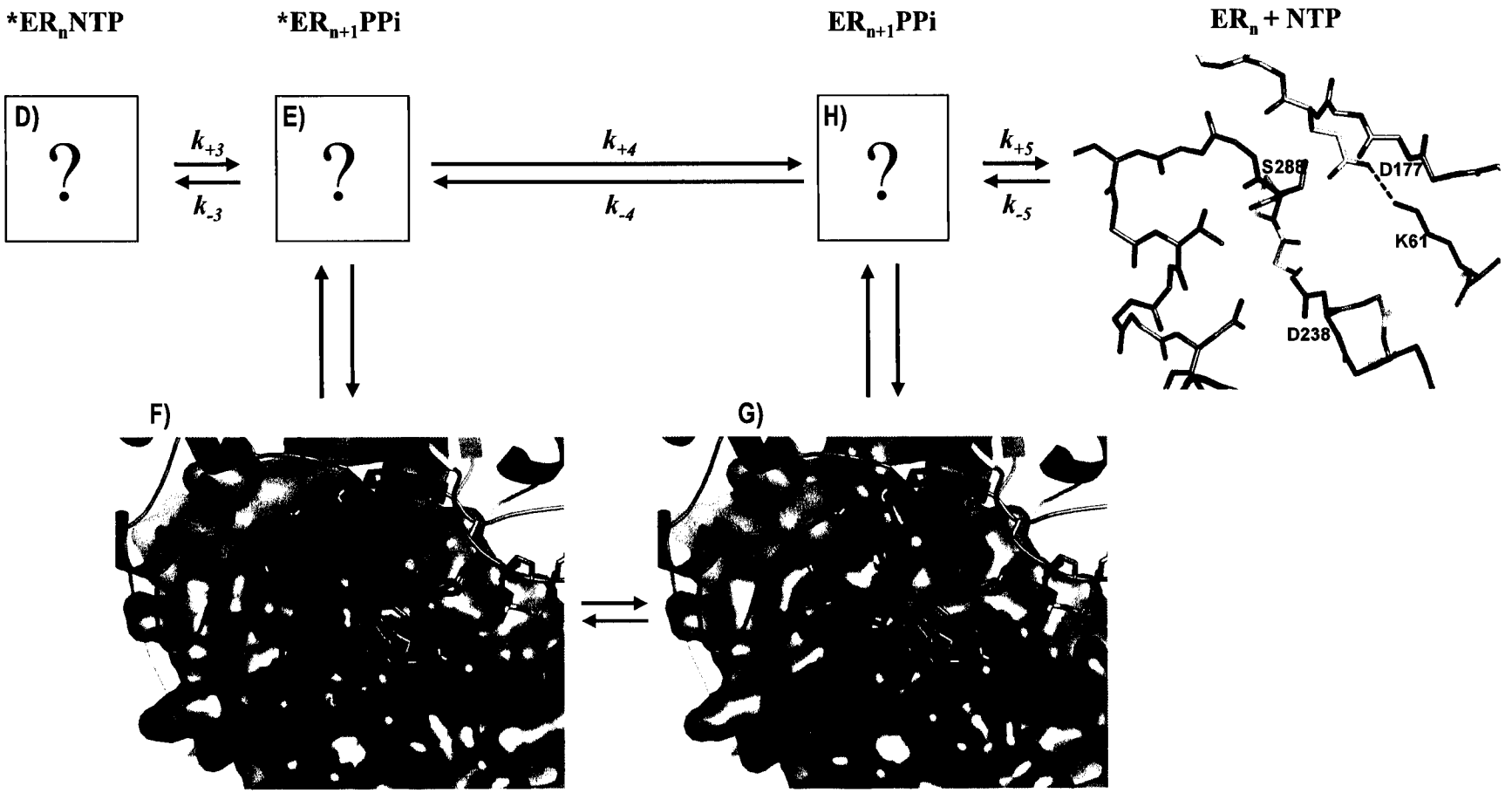


Figure 3.24: **Structural model of the proposed steps in 3D^{pol}'s catalytic cycle** See previous page for a full description of the kinetic scheme. **F)** Surface representation of 3D^{pol} with an *in* loop (purple) conformation. RNA has been modeled from FMDV (Ferrer-Orta, 2004) and the template strand is shown in green and the product in orange. The incoming NTP is colored to match the templating base in the template strand (tan). The active site has been colored magenta. **G)** Coloring scheme is identical to F) except for the loop (wheat), which is now in the *out* conformation and has sterically clashed with the RNA template.

dependent upon the flexibility of residue 289. The second conformational change of the loop will lead to the *out* conformation and translocation (Figure 3.24). Indeed, homology modeling of the RNA from the FMDV structure onto loop mutant structures with the *out* conformation shows how residue 291, when moving from the *in* to the *out* conformation, will run directly into the backbone of the template strand (Figure 3.25). Finally, the loop will move from the translocation-competent *out* conformation back to the *in* conformation, resetting the active site. The rate at which this happens depends on how readily the side chain of residue 290 can be tucked back into the hydrophobic pocket. Small hydrophobics, like Val, Ile, or Cys can readily be buried in a hydrophobic pocket, making this transition rapid and resulting in normal or hyperactive polymerases. A Ser will be less likely to be buried, slowing this transition and resulting in reduced activity. An acidic side chain will not be buried, thus completely abolishing activity.

288-292 Loop Mobility is Severely Restricted by G289A Mutants

We feel that several lines of reasoning support this model, bringing coherence to a rather large body of data. The model first predicts that mobility of the loop is essential for the translocation of the template/product strand following phosphoryl transfer and that restriction of this mobility will either completely abolish translocation or severely limit its rate. By placing an Ala at 289 we have severely limited the phi and psi angles this residue can adopt, preventing the ~180° phi angle flip needed for the conformational change of the loop. The effects of restricting the mobility of this residue is strongly demonstrated by the results obtained from the G289A mutant alone and in combination with C290I,V,F or M. In each case, the presence of an Ala289 results in the canonical *in* conformation with Ser288 pointing up away from the active site (Figure 3.14). In all cases, with the exception of G289A/C290M, the presence of an Ala at residue 289 completely abolishes activity and prevents the addition of more than one nucleotide to a defined RNA template

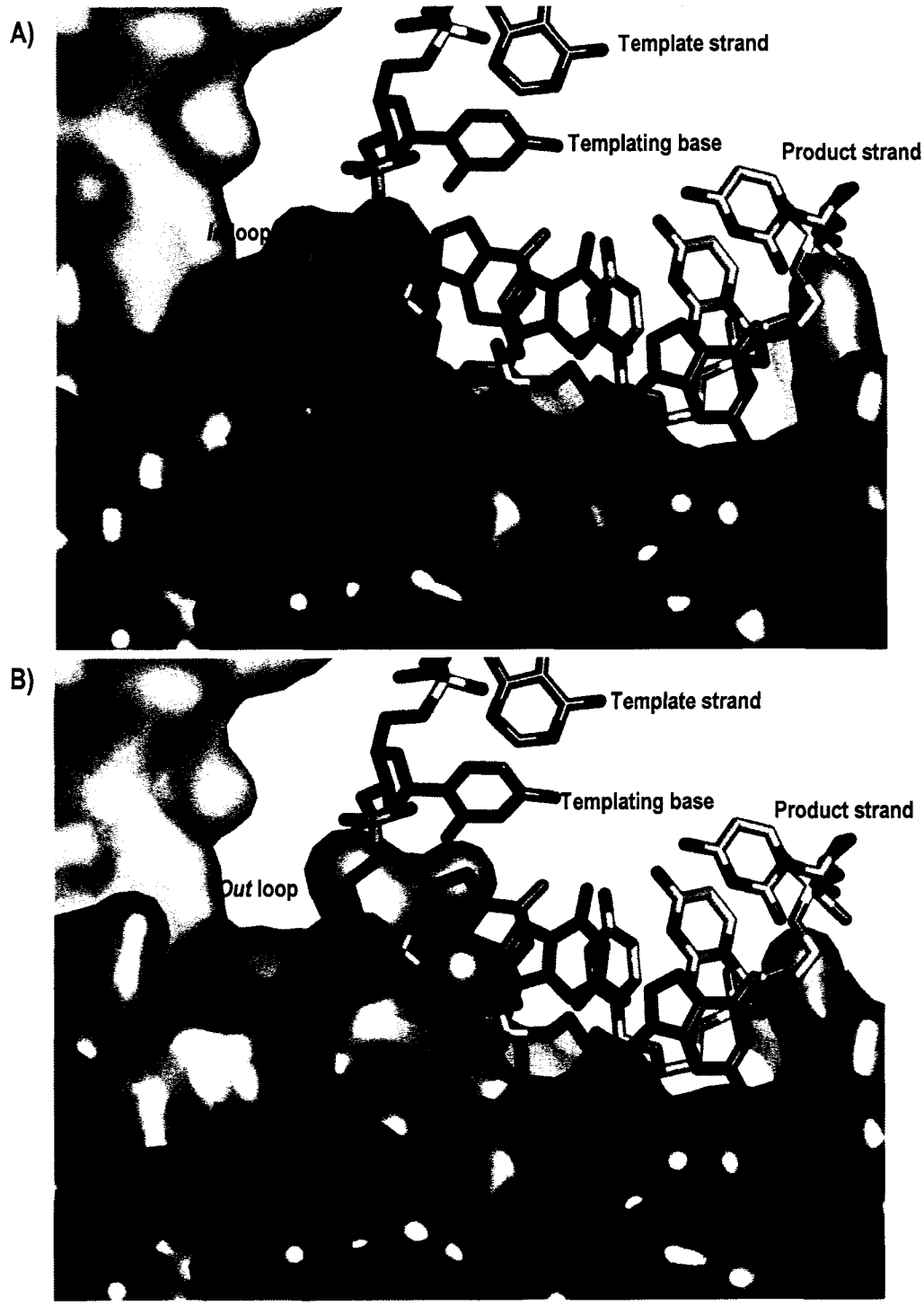


Figure 3.25: The 288-292 loop clashes with the RNA template when in the *out* conformation. A) Surface representation of $3D^{pol}$ with the active site colored in magenta and the 288-292 loop colored in purple in the *in* conformation. A duplex RNA molecule from FMDV has been modeled into the active site of $3D^{pol}$ with the template strand colored in green and the product strand colored in orange (Ferrer-Orta, 2004 (1WNE)). Large portions of the fingers domain have been removed to reveal the loop. **B)** Surface representation of the loop (wheat) in the *out* conformation clashing with the RNA template strand.

(Figure 3.15). These results directly demonstrate that the mobility of the loop is modulated by the flexibility of Gly289 and that this flexibility is essential for the translocation of the template/product strand.

Mutants Designed to Alter 288-292 Loop Mobility Affect Activity

Other mutations designed to restrict the mobility of the loop also demonstrate drastic effects on the activity of the polymerase. The substitution of Cys290 with charged residues is expected to prevent the *in* conformation due to the fact that the burial of a charged residue in a hydrophobic pocket is unfavorable and unlikely to occur. The C290E mutant supports this hypothesis in that it completely abolishes activity, prevents the addition of even a single nucleotide (Figure 3.13), results in a decrease in affinity of the polymerase for RNA, and adopts the canonical *out* conformation described by excellent density (Appendix 3 panel 1). These results can be best explained by a loop that never gains access to the *in* conformation, thereby disrupting RNA-binding and the geometry of the active site, preventing even a single round of phosphoryl transfer.

Biochemical results from the G292A and S291P mutations can also be explained in light of loop flexibility. The true conformations of these loops are not known due to the complication of the dimethyl-arsenic adduct, but it is likely that both mutations would considerably strain the dynamics of the loop, again leading to marked decreases in activity. The G292A mutant, like C290E, displays a decrease in affinity for RNA, implying that this loop prefers the *out* conformation. Each mutant does show the ability to add at least 3 nucleotides to a defined RNA template, but they do so at a significantly slowed rate (Figure 3.19). The slower overall rate of these enzymes is consistent with a decrease in the rate of translocation, which has become rate limiting due to the inflexibility of the loop.

Large Hydrophobics at Residue 290 Slow the Rate of Translocation

Activity effects for C290F and C290M are consistent with an overall decrease in the kinetics of nucleotide addition. A large residue, like Phe, may have a difficult time gaining access to the hydrophobic pocket in which it must sit when adopting the *in* conformation due to a restricted number of rotamers. The weakened density for these loops in the presence of GTP implies that these mutants may be in transit from the *out* to the *in* conformation and upon RNA binding would be found fully occupying the canonical *in* conformation. Once catalysis and translocation occurs, the loop is in the *out* conformation and has to return back to the canonical *in* conformation. This transition is expected to be rate-limiting, due to the size of the side chain, leading to the observed decreases in activity.

A Cysteine at Residue 290 Anchors the Loop for the Ser 288 Flip

The two separate conformational changes of the loop predicted by our model also explain why a conservative mutation at residue 290, such as Ser, will result in a fairly large change in activity. The first conformational change requires residue 290 to be anchored in the hydrophobic pocket while a significant change in the backbone geometry of Gly289 occurs. The slightly hydrophobic nature of a Cys may be perfectly tuned for this interaction, allowing the conformational change to occur without a major distortion of the loop. Indeed, the structures of other picornaviral RdRP suggest that this conformational change can occur without displacing the loop (Figure 3.23). From this argument it would be expected that because of its polar nature a serine side chain would not be able to hold the loop in place during the first conformational change, leading to a distortion of the loop that has implications for the subsequent steps of catalysis. Conversely, the Val and Ile mutant structures demonstrate the ability of the loop to make this transition while maintaining burial of residue 290 in the hydrophobic pocket.

Craig Cameron's lab has demonstrated that the rate of each round of nucleotide incorporation is dictated by both the first conformational change and phosphoryl transfer (Arnold and Cameron, 2004). The initial flip of Ser288 is expected to impact the rate of the first conformational change by allowing the Asp177-Lys61 salt bridge to break. It is our hypothesis that this salt bridge must break for the first conformational change to occur, enabling a yet undefined movement of the ring finger that results in the proper positioning of the NTP over the active site. If this process is slowed, then the overall rate of the enzyme will decrease, consistent with the effects of a Ser substitution at residue 290. Conversely, if this process happens efficiently, the overall rate of the enzyme could increase, consistent with Val or Ile substitution at residue 290.

Furthermore, these mutations, like the Phe and Met mutations, will affect the accessibility of the loop to the hydrophobic pocket. Once translocation has occurred, the loop must return from the *out* to the *in* conformation and a small hydrophobic side chain like Cys, Val or Ile will favor a rapid transition. Burial of a Ser in the hydrophobic pocket is disfavored, thus the *out*-to-*in* transition and the overall rate of the enzyme will be slower as compared to the native enzyme (Figure 3.21).

The Ser288 Flip is important but not Essential for Proper Loop Function

Our model predicts that the 180° flip of Ser288 plays an important role in the catalytic cycle by providing a means to break the Lys61-Asp177 salt bridge. The G289A mutant immobilizes Ser288, preventing the 180° flip observed in C290I, C290V and other *Picornaviridae* polymerases. The G289A mutants however, still display the ability to add one nucleotide, suggesting that the flip of Ser288 is not required for catalysis as the model suggests. The addition of the single nucleotide in the G289A mutants, however is very slow and in some cases almost non-existent (Figure 3.15). The slow kinetics could be explained by a slow transition of Ser288 from the *up* conformation to the *down* conformation, but this possibility is unlikely due to the torsional constraints of Ala289.

Rather, the slow kinetics are likely due to a complete lack of the Ser288 flip, presenting an energy barrier to the first conformational change because the Lys61-Asp177 salt bridge is unbroken. This energy barrier must be overcome by the enzyme, leading to a slow conformational change prior to phosphoryl transfer. Once the energy barrier has been overcome, phosphoryl transfer would occur as normal, but the G289A mutation completely prevents the second conformational change and translocation can not occur.

3D^{pol} G289A/C290M Translocation can be Described by a Distorted Loop Conformation

The double mutant G289A/C290M is ~4% active and displays the ability to incorporate all three nucleotides, albeit very slowly, indicating that this loop can adopt a conformation competent for translocation while being severely restricted. It is doubtful however, that this would be the canonical *out* conformation, but rather one can envision the C-terminal end of the loop moving *out* while residue 288 and 289 remain *in*. In support of this view residue, 288 can be modeled with a fair amount of certainty and is best described by the canonical *in* conformation. The C-terminal portion of the loop (290-292) however, exhibits very poor electron density with B-factors in the 60's and 70's and an accurate model could not be generated (Appendix 3 panel 10). The lack of density and high B-factors indicate that this loop likely adopts several conformations in crystal lattice and it is possible that the Met side chain acts similarly to the dimethyl-arsenic adduct of G289A, forcing residue 291 and 292 out (Figure 3.2 and 3.16) and resulting in translocation.

The Ser288 Flip is Stabilized by a Hydrogen Bond to Asp238

The altered backbone torsion angles resulting from a flipped Ser288 in combination with a buried Cys290 are likely stabilized by the hydrogen bond that forms between Ser288 and Asp238. This hydrogen bond will hold Ser288 in the *down* conformation, allowing the Lys61-Asp177 salt

bridge to break. Without this stabilization, one can envision a slow *up*-to-*down* transition of residue 288, subsequently slowing the processes leading to the first conformational change and resulting in marked decreases in measured activity. In support of this hypothesis, substituting Ser288 with an Ala, which removes the hydrogen bonding potential of the 288 side chain, results in a 10-fold decrease in the activity of the polymerase.

Conclusions

In conclusion, our results demonstrate the need for flexibility within the 288-292 loop for completion of the entire catalytic cycle of the enzyme, in particular translocation. Small changes within this region of the protein can have drastic affects on the activity of the polymerase, likely due to the disruption of a delicate equilibrium between loop conformations that must exists for proper functioning of the enzyme. These data have allowed us to propose a model of translocation whereby the loop adopts three separate conformations, leading to the completion of the catalytic cycle. The cycle begins with the loop in the canonical *in* conformation with Ser288 pointing *up*. Following NTP binding Ser288 flips to the *down* position setting the stage for NTP positioning and phosphoryl transfer. Following phosphoryl transfer the loop adopts the third conformation whereby the loop switches to the canonical *out* conformation, leading to translocation (Figure 3.24). This model is consistent with the data presented in this chapter. However, without further kinetic and structural investigation the exact molecular details of such a mechanism will remain unknown.

Chapter IV

Discussion and conclusions

Introduction

RNA-dependent RNA polymerases encompass a large class of enzymes responsible for the complete genomic replication of RNA viruses. As such, this class of enzymes is a natural target for antiviral therapies and has been the focus of intensive research for many years. Recently, Craig Cameron's lab reported the full kinetic mechanism for the RdRP from poliovirus demonstrating that a single round of nucleotide addition involved 5 steps: 1) NTP binding 2) a conformational change 3) phosphodiester bond formation 4) a second conformational change and 5) pyrophosphate release (Arnold and Cameron, 2004). This kinetic analysis is consistent with those reported for other classes of polymerases, including reverse-transcriptase from the human immunodeficiency virus (Hsieh et al., 1993; Kati et al., 1992), the Klenow fragment from *Escherichia coli* (Dahlberg and Benkovic, 1991; Kuchta et al., 1987), and the DNA polymerase from bacteriophage T7 (Patel et al., 1991).

Kinetic mechanisms in combination with X-ray crystallography and biochemistry are leading to a greater understanding of the structural-functional relationships of each class of polymerases and, in the case of reverse-transcriptase, allowing for the development and discoveries of new antiviral therapies for HIV. Perhaps the most influential work for the enzymology of RNA polymerases (RNAP) has come from Tom Steitz and coworkers who have published a number of papers describing the structure of the RNAP from bacteriophage T7 trapped in various stages of its

catalytic cycle (Cheetham and Steitz, 1999; Yin and Steitz, 2002; Yin and Steitz, 2004). This work has given a structural basis for the conformational changes described by the kinetic mechanisms and has influenced the conclusions presented in this thesis. In this chapter I will discuss the Steitz lab findings and how they have influenced my interpretation of the data presented in chapters 2 and 3.

Discussion

Overview of the T7 RNAP

The RNAP of bacteriophage T7 is a DNA-dependent RNA polymerase responsible for mRNA production at all stages of the phage's life cycle. To carry out its function, the enzyme must specifically recognize and bind a DNA promoter, forming a transcription "bubble" of approximately 11 nucleotides. Once the open promoter complex has been formed, the enzyme initiates RNA synthesis *de novo*, making short 2-6 nucleotide RNA fragments before switching to the elongation phase (Carpousis and Gralla, 1980; Martin et al., 1988). Once a 10-12 nucleotide RNA transcript has been made, the polymerase undergoes a significant structural rearrangement that creates a channel for the RNA transcript to exit the enzyme (Yin and Steitz, 2002). After this structural rearrangement, the enzyme carries out successive rounds of catalysis until the polymerase dissociates from the DNA upon termination. These functional aspects of the T7 RNAP differ from those of RdRP's in that the latter enzymes carry out the replication of viral genomes, do not specifically recognize a promoter sequence, and in many cases are primed by either an oligonucleotide or peptide. Having said this, the catalytic mechanism employed by the enzymes are not expected to differ greatly in that they are both magnesium-dependent, synthesize ribo-nucleic acids, and display structural homology and similar kinetic properties.

The structure of the T7 RNAP can be described by the right-hand analogy consisting of fingers, palm and thumb domains. The core of the palm domain is comprised of several α-helices and

β-sheets making up motifs A,B,C and D. These motifs are a defining structural characteristic of RNA polymerases and exhibit sequence and structural homology to elements found in the palm of RdRP's (Poch et al., 1989) (Figure 4.1). In addition to the palm, finger and thumb domains, the T7 RNAP also has a 324-amino acid N-terminal domain that interacts with 13-bases of the T7 promoter in a non-specific fashion. Specific binding of the T7 promoter is accomplished by an anti-parallel β-loop consisting of residues 739-770 in the fingers domain (Cheetham et al., 1999). The additional N-terminal domain found in T7 RNAP has no structural analog in PV $3D^{pol}$ and other RdRP's, consistent with the fact that these polymerases do not specifically recognize a promoter sequence (van Dijk et al., 2004).

PV $3D^{pol}$/NTP Complex Demonstrates a "pre-insertion" Conformation

The active site of T7 RNAP consists of two highly homologous aspartate residues (D537 and D812) that are responsible for the coordination of catalytic magnesium cations required for phosphodiester bond formation (Sousa et al., 1993). The incoming NTP interacts with basic residues found in a small α-helix adjacent to the active-site called the O-helix (Cheetham and Steitz, 1999). In 1999 and 2002 the structures of initiation and elongation complexes of T7 RNAP were solved. Both of these complexes exhibit an active site conformation consistent with a "post-translocation" model in which the incoming NTP is expected to bind, setting the stage for the subsequent round of catalysis (Cheetham and Steitz, 1999; Yin and Steitz, 2002). The Steitz lab has demonstrated that a non-hydrolyzable ATP analog bound to either of these complexes will display density for only the phosphate moiety and is said to be in the "pre-insertion" site (Figure 4.2d) (Cheetham and Steitz, 1999). The pre-insertion site is characterized by an NTP orientation where the phosphate moiety is not positioned over the active site and base pairing between the

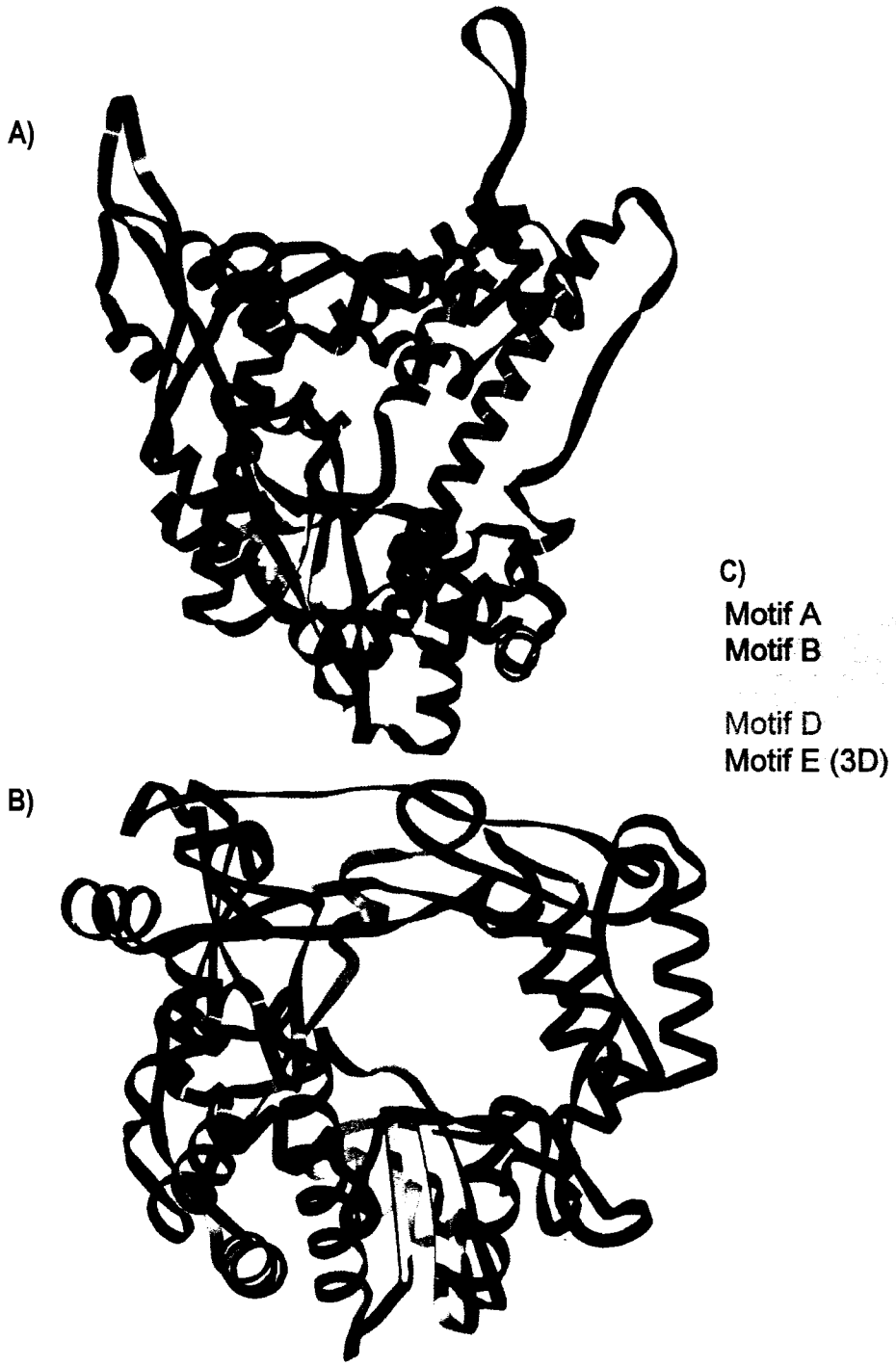


Figure 4.1: Structural homology between T7 RNA polymerase and Poliovirus RNA-dependent RNA polymerase. A) Crystal structure of T7 RNA polymerase with common structural motifs in the palm colored as shown in panel C) (Cheetham et al, 1999 (1QLN)). **B)** Crystal structure of Poliovirus RdRP with common structural motifs in the palm colored as shown in panel C) (Thompson and Peersen, 2004 (1RA6)). The structure of each motif displays a great deal of homology between each enzyme, specifically with respect to motif C that contains aspartic acid residues that coordinate catalytic magnesium cations. RdRP's contain an additional structural element, motif E, that is not found in T7 RNA polymerase.

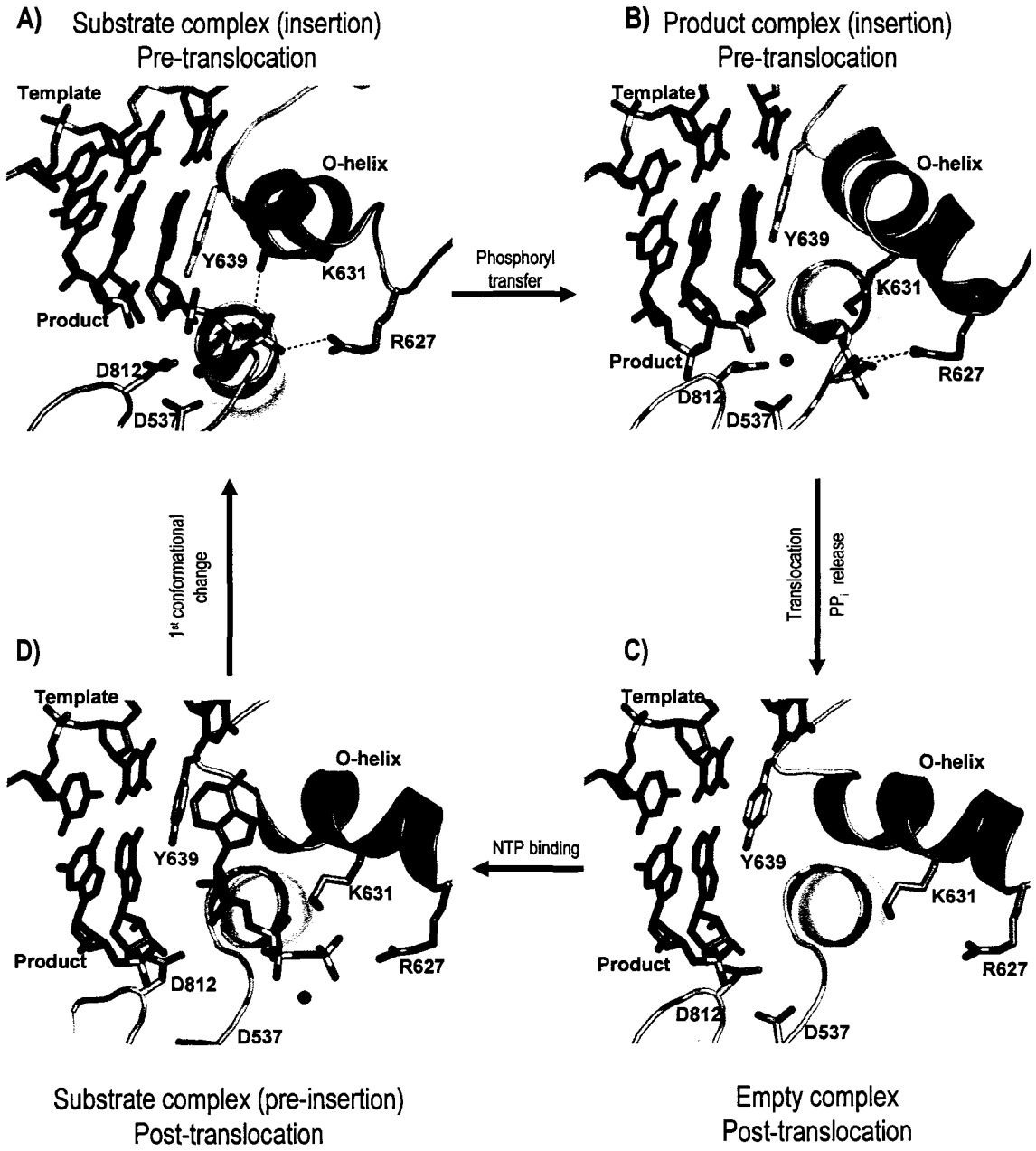


Figure 4.2: Structral basis for the catalytic cycle of T7 RNA polymerase. **A)** NTP (tan) bound in the insertion site with it's phosphate moiety positioned over D812 and D537 that coordinate two Mg cations (Yin and Steitz, 2004 (1S77)). The incoming NTP is base pairing with the correct base (tan) in the template strand. **B)** Following phosphoryl transfer pyrophosphate remains bound to the O-helix and the product strand has grown by 1 nucleotide (Yin and Steitz, 2004 (1S77)). **C)** Upon PPi release the O-helix makes a 22.5° degree rotation resulting in translocation of the nucleic acid. This event displaces the newly incorporated NTP due to a steric clash with the nucleic acid and Tyr639 moving the entire nucleic acid 3.4Å. (Yin and Steitz, 2004 (1S77)). The resulting structure is now ready to bind the next incoming NTP (Cheetham and Steitz, 1999 (1QLN)). **D)** The NTP initially binds in the pre-insertion site with a phosphate moiety not positioned over the active site and a base conformation unsuitable for base pairing (Cheetham and Steitz, 1999). Upon binding the O-helix makes a second 22.5° rotation placing the NTP into insertion site A).

template and the incoming nucleotide is obstructed by Tyr639. The conformation of Tyr639 in these complexes causes the templating base to flip out of the base-binding site and offers an explanation for the lack of density for the ribose and base of the ATP analog (Figure 4.2d) (Cheetham and Steitz, 1999).

By analogy to T7 RNAP, the conformation of the NTP bound to PV $3D^{pol}$ described by Thompson and Peersen in 2004 is in the pre-insertion site (Thompson and Peersen, 2004). The phosphates of this NTP are bound by basic residues found in the ring finger and, like the pre-insertion site of T7 RNAP, are not positioned over the active site. Unlike T7 RNAP, the base of the NTP in $3D^{pol}$ is well-ordered and a structural analog for Tyr639 does not exist. However, the base does exist in a conformation unsuitable for base pairing because the base-pairing edge faces into the protein forming a hydrogen bond with the backbone carbonyl of residue 175 (Figure 4.3).

The Ring Finger of $3D^{pol}$ is Analogous to the O-helix of T7 RNAP

For phosphoryl transfer to occur in both T7 RNAP and PV $3D^{pol}$, the NTP must be moved into an orientation that places its phosphates directly over the aspartate side chains of the active site. This movement corresponds to the first conformational change described by the respective kinetic mechanisms, and in the case of $3D^{pol}$ is partially rate-limiting (Arnold and Cameron, 2004; Arnold et al., 2005). In 2004 the Steitz lab published a paper describing the structural basis for this movement (Yin and Steitz, 2004). In this paper, two structures are presented, revealing that prior to translocation and pyrophosphate release the enzyme exists in a "pre-translocation" conformation. This conformation is characterized by a 22.5° rotation of five α-helices in the fingers domain, including the NTP binding O-helix, as compared to the post-translocation complex. This rotation moves the phosphate moiety of the NTP over the active site and into an orientation suitable for phosphoryl transfer. In addition, Tyr639 moves out of the base-binding site, allowing

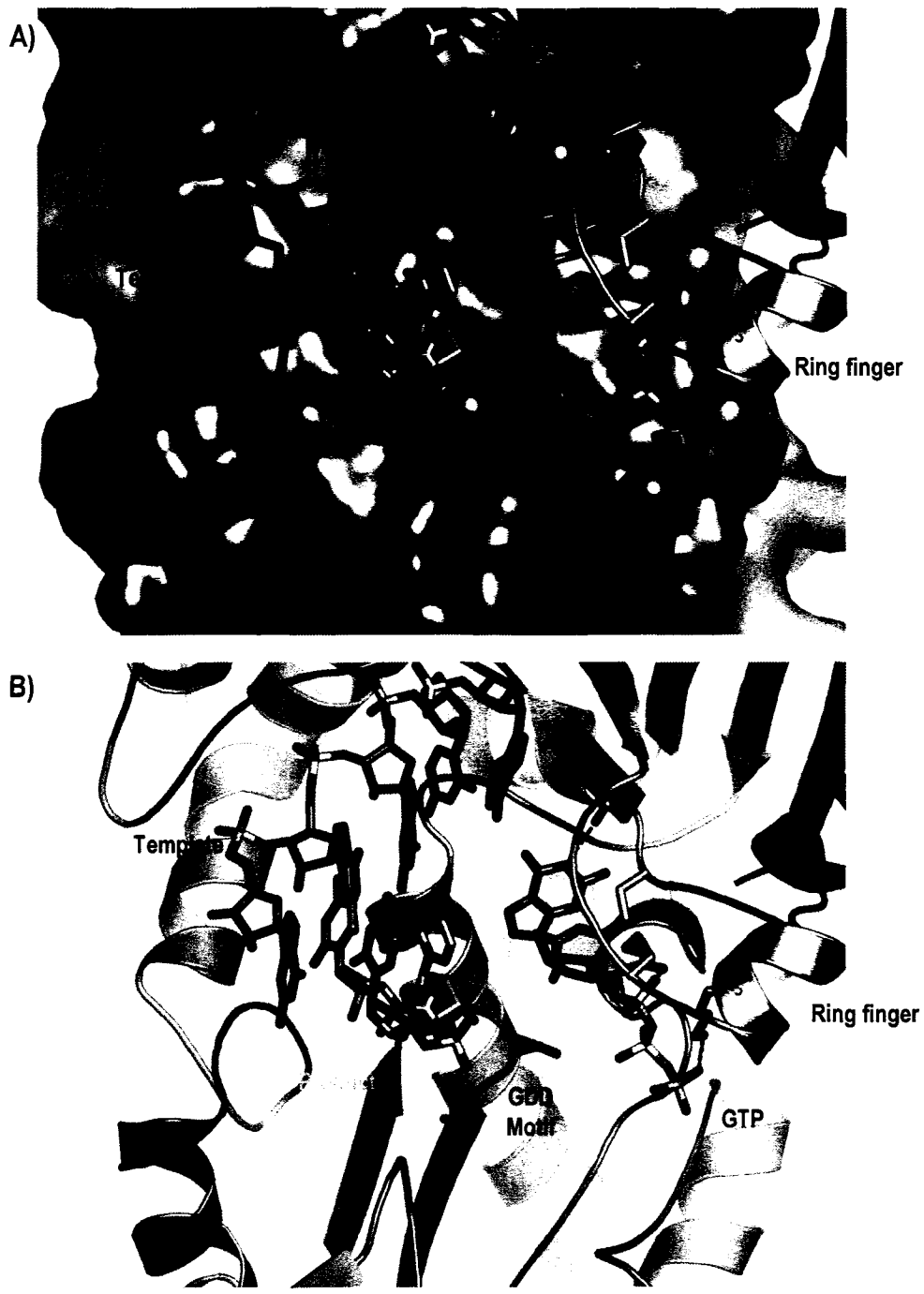


Figure 4.3: **The NTP bound to poliovirus 3D^{pol} is in the pre-insertion site.** **A)** Surface representation of the active site (magenta) of 3D^{pol} with GTP bound and a modeled duplex RNA molecule from FMDV RdRP (Ferrer-Orta et al., 2004 (1WNE)). The template strand is colored in green and the product strand is colored in orange. The NTP is bound by basic residues found in the ring finger (yellow). **B)** Cartoon representation of the RdRP with an identical view and coloring scheme as seen in A). Both figures demonstrate how the NTP is bound in the pre-insertion site where the phosphate moiety is not positioned over the catalytic aspartic acids and the NTP is in a conformation unsuitable for base-pairing with the templating base (colored to match the incoming NTP) of the template strand (Thompson and Peersen, 2004 (1RA7)).

for a canonical Watson-Crick pair between the template and the incoming NTP (Figure 4.2 a and b) (Yin and Steitz, 2004). Following this movement, phosphoryl transfer occurs and the enzyme remains in the pre-translocation state until release of pyrophosphate (Figure 4.2 b and c).

Homology modeling of PV $3D^{pol}$ with the RNA complex of the RdRP from foot-and-mouth disease virus (FMDV) (Ferrer-Orta et al., 2004) demonstrates how both the NTP and the templating base will have to undergo a fairly significant conformational change for catalysis to occur (Figure 4.3). By analogy to T7 RNAP, we expect that the ring finger of $3D^{pol}$, like the O-helix, will push the NTP into the active site forming a catalytically competent conformation.

Molecular Events Leading to the 1st Conformational Change in $3D^{pol}$

Craig Cameron's lab has demonstrated that the conformational change prior to phosphoryl transfer is rate-limiting for the $3D^{pol}$ G64S mutant (Arnold et al., 2005). I reported in Chapter 2 the structure of this mutant, along with a variety of other structures, all of which demonstrate movements in a small α-helix comprised of amino acids 52-62. This movement is less exaggerated in the presence of an NTP, indicating that structural elements in the fingers are flexible and responsive to events taking place in the active site. This is likely due to a yet-to-be understood relationship between the N-terminus, Asp238, and the ring and index fingers. It is also important to note that the index finger appears to be a structural platform maintaining the integrity of the fingers domain by an interaction with the top of the thumb (Figure 2.14). This suspicion has been confirmed by Aaron Thompson in our lab who showed that mutations specifically designed to disrupt the index/thumb interaction will destabilize the entire fingers domain (Thompson, 2006). I propose that a large movement in the 52-62 α-helix, dependent on flexibility within the structural elements surrounding the N-terminus-binding pocket, will result in a clockwise rotation of the small

α-helix at the tip of the finger and result in the subsequent movement of the NTP into the active site (Figure 2.14).

A second conformational change in a separate portion of the enzyme that could lead to a movement in the ring finger, analogous to the O-helix, was discussed in chapter 3. A flip of Ser288 down towards Asp238, such as that seen for the hyperactive mutants C290V and C290I, allows the salt-bridge between Asp177 and Lys61 to break. This salt-bridge anchors the ring finger to the index finger and when broken would set the stage for a large movement of the 52-62 α-helix and reorientation of the ring finger analogous to the O-helix.

Translocation of the Template/product Strand

Following phosphoryl transfer, kinetic schemes predict a second conformational change that corresponds to the translocation of the product/template strand, resetting the active site for the next round of catalysis. Again, the molecular details of this movement have been described by the Steitz lab for the T7 RNAP. The conformational change that results in translocation of the T7 RNAP template/product strand is the same movement required to position the NTP over the active site, albeit in the reverse direction. The five α-helices that moved forming the pre-translocation conformation rotate 22.5° back to their original position described by the post-translocation conformation. This movement brings Tyr639 back into the nucleotide base-binding site, pushing the template/product strand 3.4 Å (Figure 4.2 b and c). The polymerase now sits in the pre-insertion state ready to bind the next NTP, and start the next round of catalysis.

As mentioned previously, PV $3D^{pol}$ does not have a structural analog of Tyr639, indicating that the molecular mechanism for translocation may be quite different from that seen for T7 RNAP. In chapter 3 I describe two stable conformations of a loop containing residue 288-292 that, based on homology modeling with FMDV RdRP, would be in direct contact with the template strand of a

bound RNA molecule. Biochemical evidence demonstrates how inflexibility within this loop leads to translocation defects, most notably seen with the $3D^{pol}$ G289A mutants. I have proposed a model in which transition between the two conformations of this loop results in translocation (Figure 3.1 and 3.24). Prior to translocation the loop sits in the *in* conformation in which residue 290 resides in a hydrophobic pocket directly behind the loop. Following phosphoryl transfer the loop undergoes a drastic conformational change, moving residue 290 ~5.0 Å towards the active site resulting in the *out* conformation. We hypothesize that this conformational change may act alone or in concert with other structural rearrangements to translocate the template product strand, resetting the active site for the next round of catalysis.

Conclusion

My data demonstrates a number of small structural changes within the RdRP of poliovirus that each result in rather large effects on the elongation activity of the enzyme. We hypothesize that the changes observed offer a glimpse of larger structural rearrangements expected to occur for the polymerase to carry out its catalytic cycle. Full support of this hypothesis awaits the publication of RdRP elongation complex structures with NTP's bound in both a pre and post-translocation conformation.

Chapter V

Oxytricha nova Telomere End Binding Protein

Purpose of this Chapter

My first two years in the Peersen lab were spent on an entirely separate project from that which is entailed in Chapters 1-4 of this thesis. Upon entrance into the lab it was decided that the focus of my research would be thermodynamic, kinetic and structural investigation of the Telomere End Binding Protein from *Oxytricha nova* (*On*TEBP). To accomplish this, I set out to develop several biophysical techniques and in the process ran into what, at the time, we considered a significant experimental difficulty that ultimately lead to the cessation of the project. The purpose of this chapter is to describe these endeavors, detailing the difficulties encountered along the way and interpreting them in light of recent findings in telomere biology.

Introduction

Telomeres are specialized nucleoprotein structures that protect the ends of eukaryotic chromosomes, ensuring proper maintenance and chromosomal stability, by acting as a cap for the end of the chromosome. This "cap" enables cells to differentiate between chromosomal ends and an unnatural double-strand DNA break, thereby preventing improper invocation of the DNA damage-repair system (Maser and DePinho, 2002). Activation of these mechanisms, at telomere ends would result in either degradation or non-homologous end-joining, eventually leading to apoptosis. Due to the inability of the replicative machinery to completely synthesize the entire strand of DNA, telomeres get successively shorter. The shortening of telomeres within somatic

cells places a limit (Hayflick limit) on the proliferative life of the cell. Once this limit has been reached, p53 and retinoblastoma pathways induce cellular senescence (Wright and Shay, 1992). The inactivation of these pathways will lead to cellular crisis and apoptosis, but this can be overcome by abnormal activity and expression of telomerase, the enzyme responsible for maintenance of telomere length (Counter et al., 1992). Aberrant telomerase expression is seen in ~95% of all known cancers and consequently it is thought that improper maintenance of these protective ends plays a direct role in oncogenesis (Maser and DePinho, 2002).

Telomeres have a unique structure characterized by short repeating TG-rich sequences that terminate in a 3'-ssDNA overhang. The telomere repeat sequence provides multiple binding sites for a variety of proteins, and proteins that are specifically targeted to the 3'-ssDNA have been identified in several organisms, including the telomere end-binding protein (TEBP) in *Oxytricha nova* (*O. nova*) (Gottschling and Zakian, 1986; Price and Cech, 1987), Cdc 13 in *Saccharomyces cerevisiae* (Lin and Zakian, 1996; Nugent et al., 1996) and Pot1 (a homologue of the *On*TEBP α-subunit) in a host of higher eukaryotes (Baumann and Cech, 2001). These ssDNA-binding proteins are an essential component of the nucleoprotein complex found at the very ends of telomeres that prevent the natural ends of the chromosome from being rapidly degraded or treated as a double-stranded DNA break (Baumann and Cech, 2001).

Oxytricha nova Telomeres and Telomere End-Binding Protein

In 1980 Swanton *et al.* were able to demonstrate that the entire genome of the macronuclear ciliate, *Oxytricha nova*, is split into ~24 million small gene-sized fragments entirely contained within the macronucleus (Swanton et al., 1980). Tom Cech, and others in the telomere field, immediately recognized that the macronuclear ciliates represented enormous potential as model systems for telomere biology due to the fact that they were predicted to have up to ~48 million chromosomal

ends in a single cell. To this day the macronuclear ciliates continue to be used as a model system for investigation of the chromosomal ends.

Oxytricha nova has a telomere consisting of the short repeat sequence, T_4G_4 (Klobutcher et al., 1981). The final 16 bases ($T_4G_4T_4G_4$) of the telomere are single-stranded and bound by a DNA-dependent heterodimer consisting of two subunits, α and β (Fang and Cech, 1993; Gottschling and Zakian, 1986; Price and Cech, 1989). The β-subunit is a 41-kD protein which does not specifically interact with ssDNA and during purification it is readily proteolyzed, resulting in a fully functional 28-kD core. The α-subunit is a 56-kD protein consisting of N- and C-terminal domains, and unlike the β-subunit, is capable of specifically interacting with the telomeric ssDNA (Gray et al., 1991). When α, β and ssDNA are mixed, they form a very stable α:β:ssDNA ternary complex which has been shown to have an *in vitro* half-life exceeding 100 hours (Fang and Cech, 1993).

In 1998 the Schultz lab published the high-resolution structure of the α and β-subunits in complex with a ssDNA 5′-$G_4T_4G_4$-3′ molecule (Figure 5.1)(Horvath et al., 1998). This structure showed that the ssDNA molecule adopts a highly irregular conformation where each base is partially or completely buried in a deep cleft between the two proteins. The last four bases of the ssDNA form a loop and the very 3′-end of the ssDNA is buried up against the N-terminal domain of the α-subunit, demonstrating the ability of this complex to act as a cap for the telomeric end of *O. nova* (Horvath and Schultz, 2001; Horvath et al., 1998). The α and β-subunit consist of a total of four-different OB-folds imparting, great sequence specificity to the *On*TEBP whereby bases have multiple interactions with the protein, often forming stacking interactions with aromatic residues (Theobald and Schultz, 2003). This fold has been described for a variety of different proteins with vast biological functions and has been specifically designed to participate in ligand recognition of

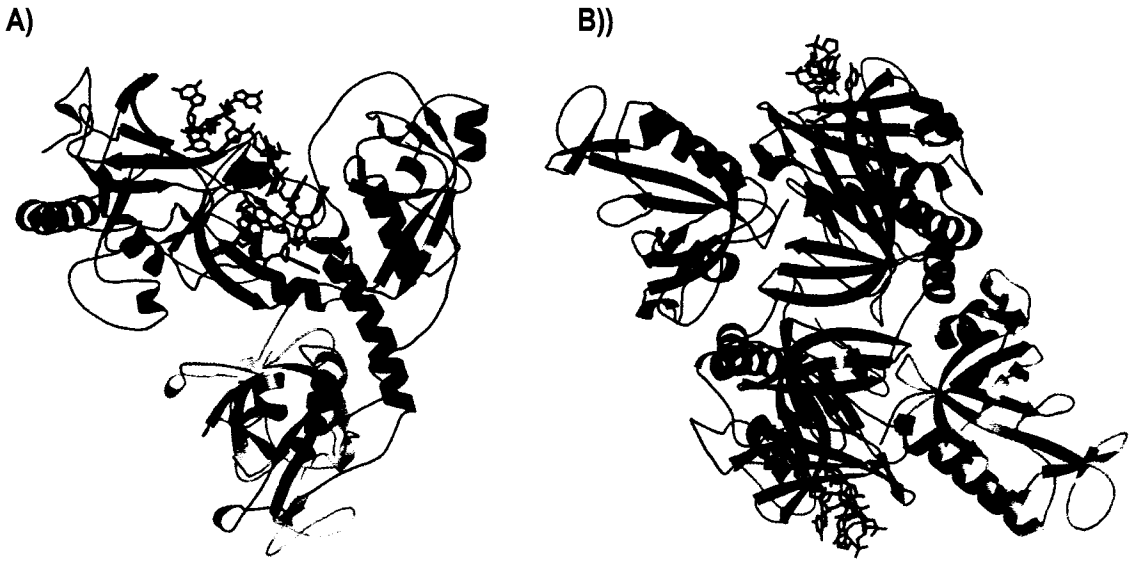


Figure 5.1: **X-ray crystal structures of the *On*TEBP**. **A)** The crystal structure of the α:β:ssDNA ternary complex (Horvath et al., 1998 (1OTC)). The N-terminal domain (green) of the α-subunit binds specifically to the ssDNA that has it's very 3′-end buried against the α-subunit. The β-subunit (blue) interacts with both the N-terminal and C-terminal (orange) domains of α as well as the ssDNA molecule. **B)** The crystal structure of the α-subunit in complex with ssDNA (Peersen et al., 2002 (1KIX)). The structure demonstrates how α can form a homodimer with one α-subunit colored as seen in A) and the other colored blue. The dimer is stabilized by extensive contacts between the N- and C-terminal domains of the separate α-subunits, including an extended β-sheet seen in the middle of the complex. Two ssDNA molecule are found bound to the separate N-terminal domains of the two subunits making an $(\alpha:ssDNA)_2$ complex.

oligonucleotides, oligosaccharides and oligopeptides. Since publication of *On*TEBP structure, the structures of the ssDNA-binding proteins Cdc13 (Mitton-Fry et al., 2002) and Pot1 (Lei et al., 2003; Lei et al., 2004) have been solved and, despite very weak sequence homology, each protein contains at least one OB-fold.

In addition to the structure of the α:β:ssDNA ternary complex, the Schultz lab has also published the structure of the α-subunit in complex with a ssDNA 5′-T_4G_4-3′ molecule (Peersen et al., 2002). This structure demonstrates that the α-subunit is capable of forming a homodimer within the crystal lattice that buries a total of ~7000 $Å^2$ of solvent accessible surface area (Peersen et al., 2002). The interaction between the two α-subunits is mediated by an extensive interface, including a continuous six-stranded anti-parallel β-sheet between the two N-terminal domains of the separate α-monomers. This interface overlaps with the ssDNA-binding pocket in the α:β:ssDNA complex, resulting in a binding register shift of the ssDNA in the α:T_4G_4 complex. The elucidation of the α-homodimer structure and separate biochemical data lead to the proposal that the α-subunit played a role in regulating the assembly and disassembly of higher-order telomeric complex observed by electron microscopy (Murti and Prescott, 2002). Having said this, the question still remains whether or not α is indeed capable of forming a dimer in solution and if so what role does this dimer play in the biology involving the *O. nova* telomere ends? To answer these questions I set out to obtain an in-solution measurement of the α-dimer using analytical ultracentrifugation.

My results indicate that α does form a higher order complex, but this complex is likely dependent on the formation of an alternate ssDNA structure, the G-quartet. G-rich single-stranded DNA sequences are known to form several alternate conformations, all stabilized by hydrogen-bonded G-quartets whereby four guanine bases participate in Hoogstein base-pairing stabilized by

a central cation (Figure 5.2). These alternate ssDNA structures form readily in the presence of sodium or potassium cations and it is unclear to us how they may be mediating the formation of α-subunit higher order complexes and whether the α-subunit can catalyze their formation.

Results

Velocity sedimentation analysis is an assay performed by subjecting the protein in question to a large centrifugal force while monitoring the rate (or velocity) at which the protein sediments. This information is then transformed into a sedimentation coefficient, a value dependent on both the size and shape of the molecule that can be used to determine if a single protein is capable of self-association (van Holde and Weishcet, 1975). All sedimentation velocity data was collected at the absorption maximum of oregon-green, a fluorescent tag used to selectively label either the DNA or the protein. Although we did not use oregon green's fluorescence, its absorption at 490nm allowed us to avoid complication arising due to the overlapping absorption spectrums of unlabeled DNA and proteins.

Initial velocity sedimentation analysis of the α-subunit of *On*TEBP indicated that α, in the absence of ssDNA, has a sedimentation coefficient value of 3.5S. This value is consistent with a monomeric form of α based on its molecular weight and the shape it adopts in the crystal lattice. Upon the addition of a 5′-T_4G_4-3′ ssDNA molecule, α exhibits a range of sedimentation coefficients (3.5-7.5S), suggesting that α is indeed capable of forming ssDNA-dependent higher order complexes (Figure 5.3). In support of these results, sedimentation equilibrium data of α in the presence of ssDNA fit to a single species with a molecular weight of 120kDa, the approximate mass of an α-homodimer (Figure 5.4).

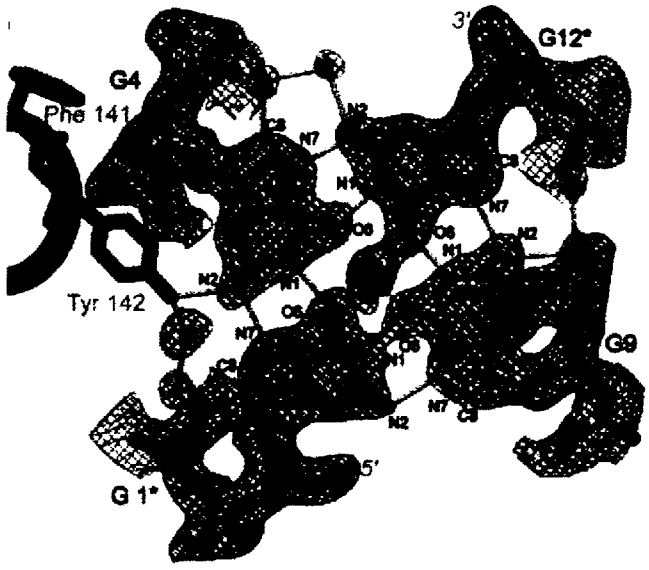


Figure 5.2: **Hoogstein base pairing of a G-quartet**. The G-quartet is stabilized by Hoogstein base pairing where four guanine bases form two hydrogen bonds between O6 and N7 of one ring and N1 and the exocyclic amine, N2, of another ring. The structure is further stabilized by a sodium or potassium cation that is situated in the middle of the four bases making favorable electrostatic interactions with O6 of each base (Figure from Horvath and Schultz, 2001 (1JB7)).

● - 10μM T_4G_4

■ - 10μM α

▲ - 10μM T_4G_4 and 10μM α

Figure 5.3: Sedimentation velocity of the α-subunit in sodium cation-contaminated solutions. Sedimentation velocity experiments of samples containing 10μM T_4G_4, 10μM α and a 1:1 mix of T_4G_4 and α (both at 10μM) were run in a buffer containing 150mM LiCl and 25mM HEPES pH 7.5 (pH of HEPES was adjusted with NaOH). The sedimentation coefficient of oregon-green labeled T_4G_4 is ~1S, and that of the α-subunit is ~3-3.5S. When the labeled ssDNA is mixed with α a range of sedimentation coefficients from ~4-7.5S is obtained indicating that the α-subunit, in the presence of ssDNA is capable of forming higher-order complexes. This result however, was difficult to reproduce, likely due to formation of G-quartets stabilized by the sodium cations derived from NaOH.

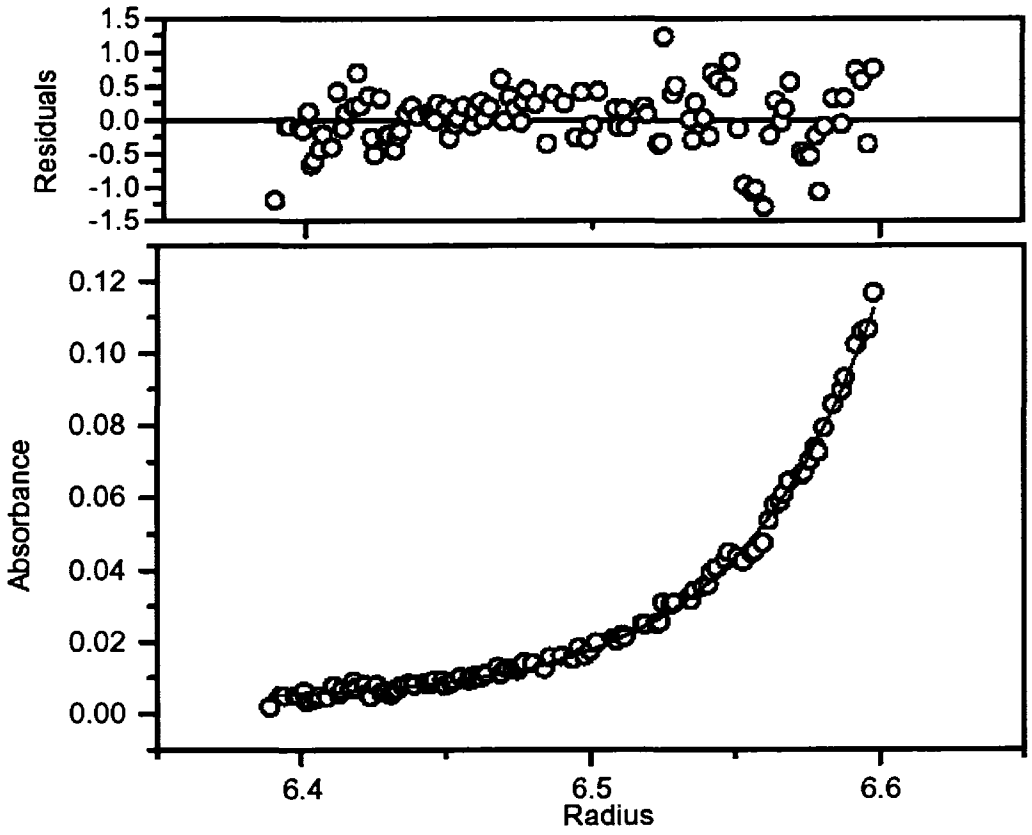


Figure 5.4: Sedimentation equilibrium analysis of the α-subunit in the presence of ssDNA. A sedimentation equilibrium experiment was run on a sample consisting of 0.6μM α and 0.5μM T_4G_4 ssDNA in a buffer containing 150mM LiCl and 25mM HEPES pH7.5 (again the HEPES pH was adjusted with NaOH). Data obtained for the α-subunit fits a single-species with a molecular weight of 120kD, the mass of an α-dimer.

These results were difficult to reproduce with any certainty, and the fact that the120kD species observed in the sedimentation equilibrium experiment would not fit to curves representing a monomer-dimer equilibrium suggested that α was forming a kinetically-trapped higher order complex. G-quadruplexes are often thought of as a kinetic trap for ssDNA, making us wonder whether these results were complicated by the presence of G-quartets. Circular dichroism has proven to be a useful tool in the detection of G-quartets, which give a distinct large positive band at 260nm followed by a negative band at 240nm (Henderson et al., 1990). To detect whether G-quartets were in our samples we ran them out on a gel-filtration column that separated the samples into three separate peaks. Using the CD spectrophotometer we were able to determine that the first peak was primarily protein, the second peak was G-quartet DNA and that the third peak was ssDNA.

The presence of G-quartets in our samples was likely due to sodium or potassium cation contamination and further measures were taken to prevent G-quartet formation, including the use of LiOH to adjust the pH of our buffers. Gel-filtration runs of ssDNA equilibrated in these new conditions demonstrated that even after 24 hours of incubation, an appreciable amount of G-quartets would not form (Figure 5.5). Using these new conditions, the sedimentation velocity experiments were repeated. Unfortunately, the data now demonstrated that the α-subunit adopted a single sedimentation coefficient of 3.5S both in the presence and absence of ssDNA (Figure 5.6).

To confirm whether these results were due to the absence of sodium or potassium and likely G-quartet formation, solutions containing either sodium chloride or potassium chloride were run side-by-side with the sodium-deficient samples. Both the presence of sodium and potassium lead to the originally observed increase in sedimentation coefficients, indicating that α is capable of forming sodium/potassium-dependent higher order complexes (Figure 5.6).

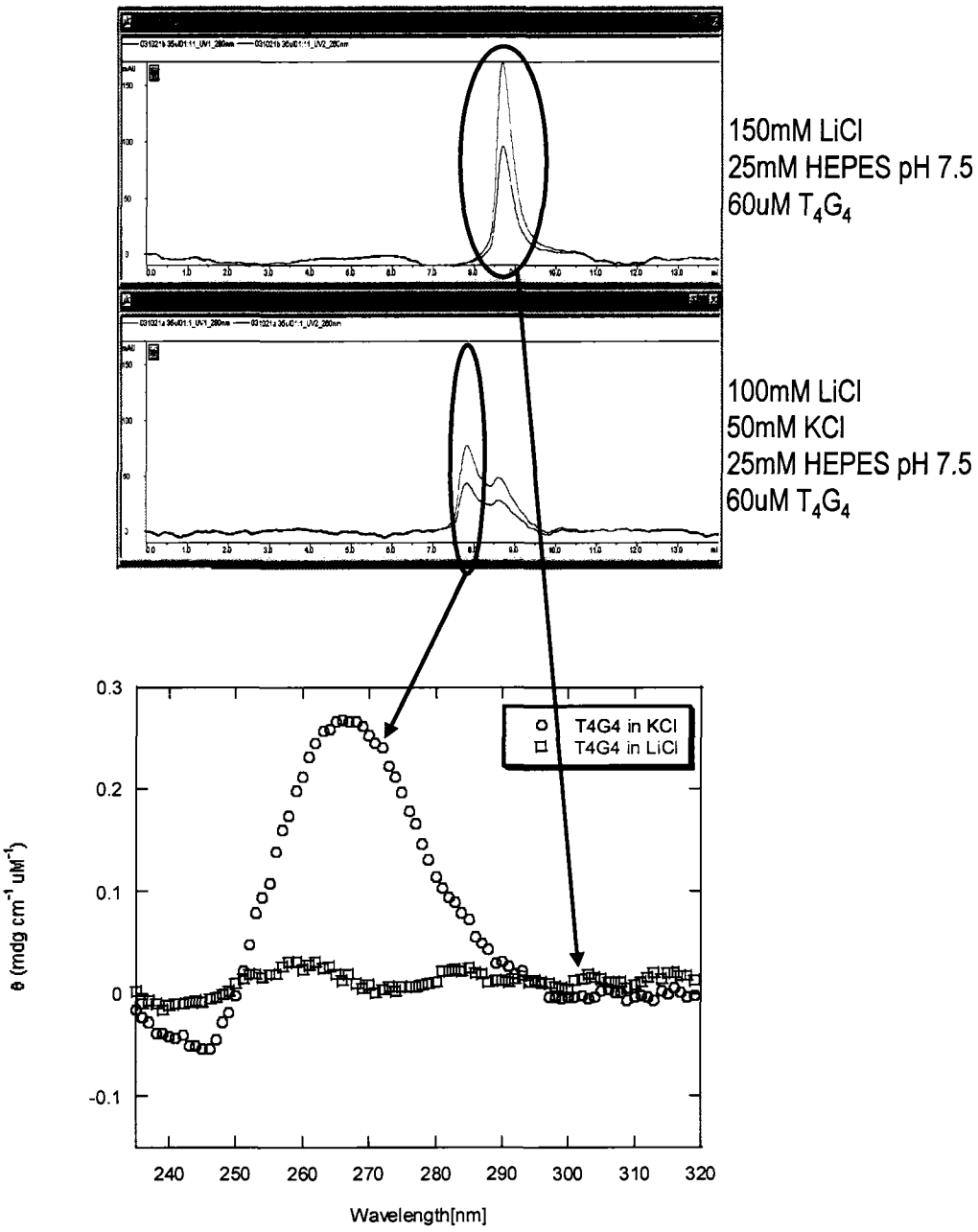
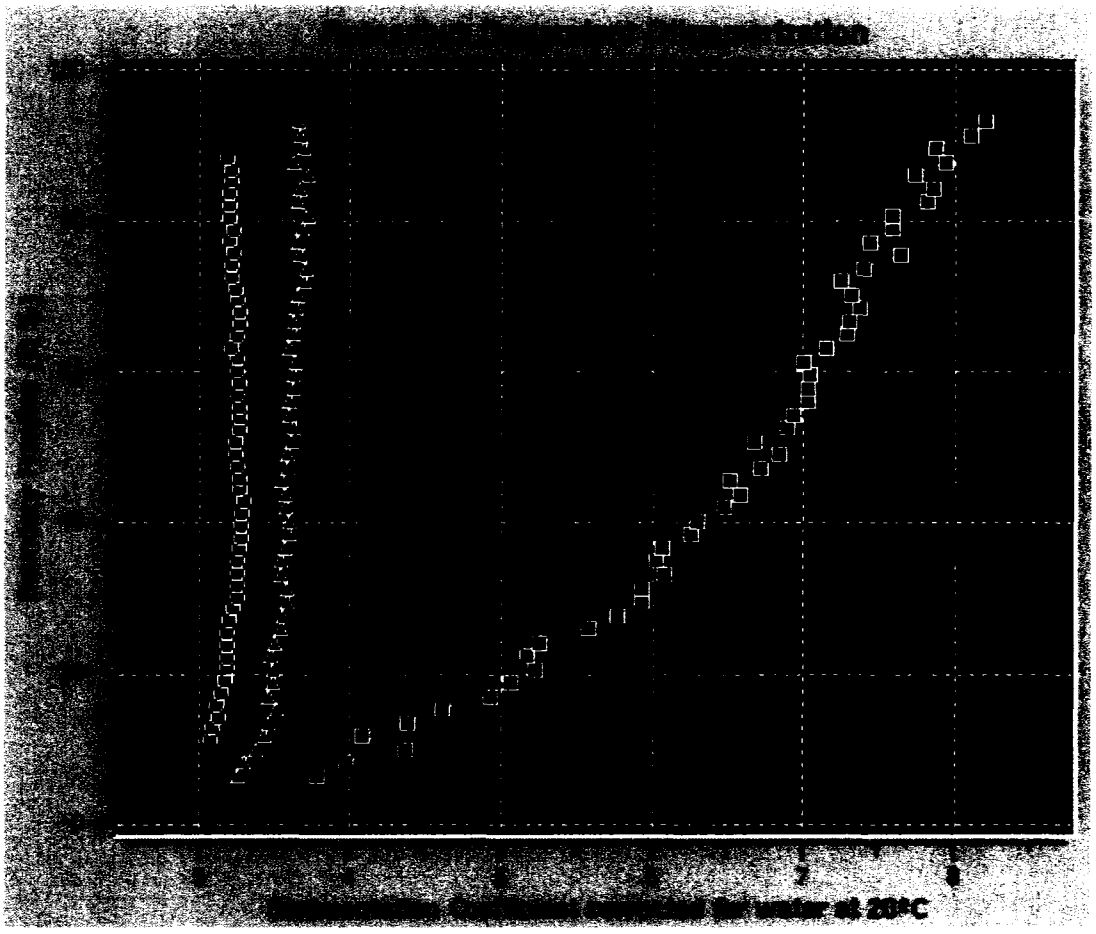


Figure 5.5: Adjusting the pH of buffers with LiOH instead of NaOH prevents G-quartet formation. 60μM T_4G_4 was equilibrated in buffer containing either 150 mM LiCl and 25mM HEPES pH 7.5 or 100mM LiCl, 50mM KCl and 25mM HEPES pH 7.5 (pH of HEPES was adjusted with LiOH in both samples). Samples were allowed to sit for 24 hours after which they were run over a gel-filtration column (A). The sample containing LiCl only produced a single peak consisting entirely of ssDNA as demonstrated by CD spectrophotometry (B). The sample containing KCl split into two separate peaks, the first consisting of G-quartets (confirmed by CD (B)) and the second, corresponding with the LiCl sample, consisting of ssDNA.



● - 10μM α-subunit

▲ - 10μM α-subunit and 20 μM T_4G_4

■ - 10μM α-subunit and 20 μM T_4G_4 + KCl

Figure 5.6: **The α-subunit forms potassium-dependent higher-order complexes.** Sedimentation velocity experiments were set up using buffers adjusted with LiOH. In the absence of ssDNA the α-subunit, as previously shown, adopts a single sedimentation coefficient of ~3.0S. A 1:2 mixture of the α-subunit with ssDNA in the absence of sodium or potassium, adopts a single sedimentation coefficient of ~3.5S. Finally, the addition of KCl to the α:ssDNA mixture results in the previously seen range of sedimentation coefficients, from ~3.5S-8.0S. The combination of these results demonstrates that higher-order complex formation of the α-subunit is dependent on potassium or sodium.

The above data strongly suggests that G-quartets are playing a role in the formation of α:T_4G_4 higher-order complexes and we set out to confirm the absence of this ssDNA species in samples containing only lithium cations. Samples were prepared that contained either 150 mM LiCl or 100mM LiCl and 50mM KCl, the α-subunit and ssDNA. These samples were run on the gel filtration column at different time points and to our surprise a significant amount of G-quartets was found in both the LiCl and KCl samples. Furthermore, as time progressed the amount of G-quartet DNA in the LiCl sample decreased, but remained constant in the KCl sample (Figure 5.7). These results suggest that the α-subunit has the ability to catalyze rapid G-quartet formation, but that this ssDNA conformation is unstable if a sodium or potassium cation is not present.

Discussion

Despite the abundant data supporting the fact that telomeric sequences will readily adopt G-quartet conformations *in vitro,* there has been a significant debate regarding their biological relevance. Until recently, the only direct evidence supporting a role for G-quartets *in vivo* is that an antibody raised against G-quartets labels the macronuclei of macrociliates (Schaffitzel et al., 2001). These results are complicated by the suspicion that the antibody actually induces G-quartet formation. Recently Paeschke *et al.*. were able to demonstrate, using the same antibody, that the RNAi knockdown of the α and β-subunit in *O. nova* lead to a decrease in epitope labeling of the macronuclei, suggesting that these proteins play a role in maintenance of the G-quartets *in vivo* and that in their absence the G-quartet is rapidly disassociated (Paeschke et al., 2005). Furthermore, these results indicate that the antibody interaction with G-quartets is likely not due to G-quartet induction by the antibody, but rather a valid labeling of pre-existing G-quartets, demonstrating that these structures do indeed play a relevant biological role.

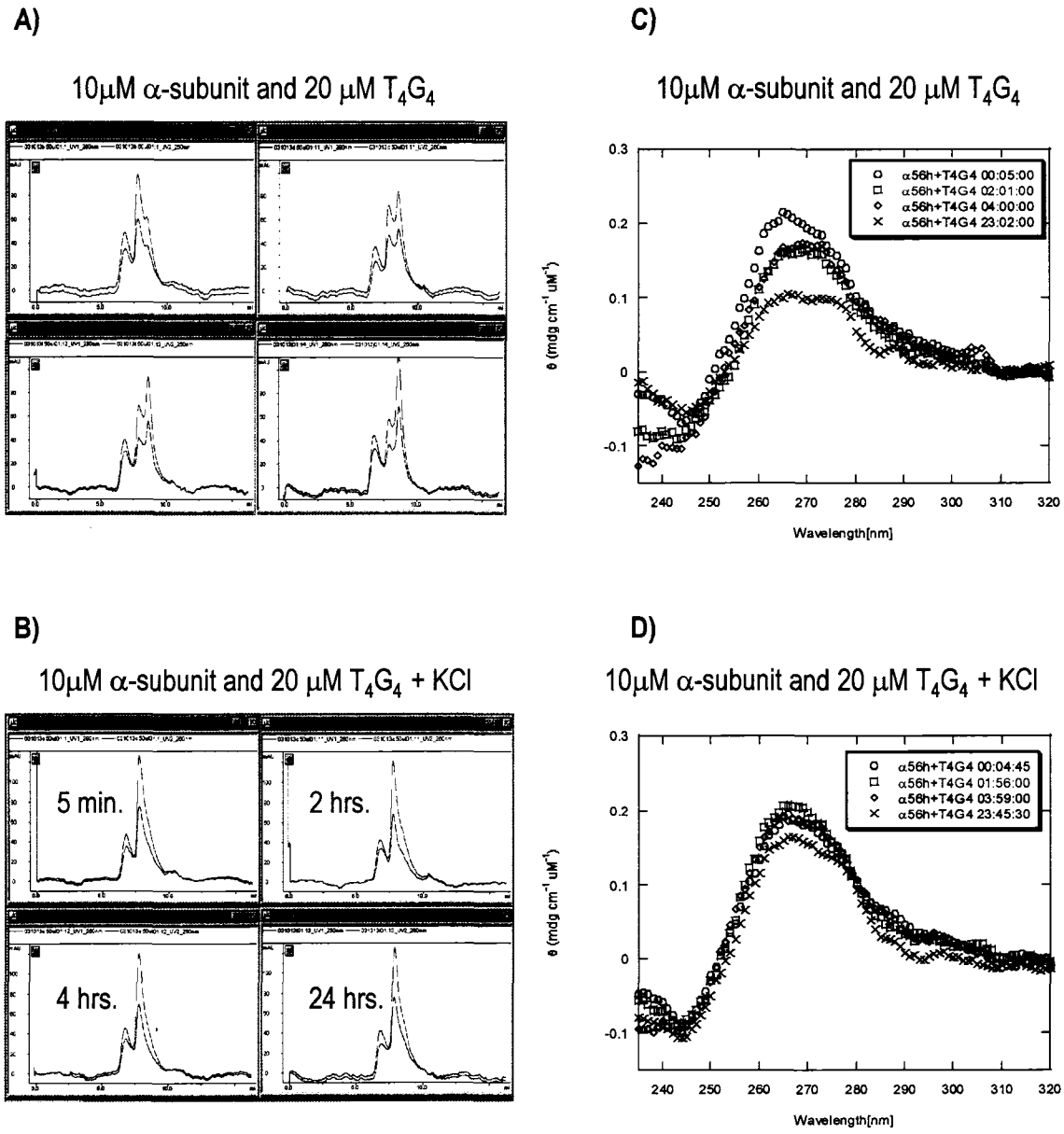


Figure 5.7: **G-quartets form in the presence of the α-subunit regardless of buffering conditions**. Samples containing 10μM α and 20 μM T_4G_4 in the absence (A) or the presence (B) of potassium were run over a gel-filtration column at various time-points. To our surprise a significant G-quartet peak (the second peak) appeared in both buffering conditions. In the absence of KCl this peak decreased significantly over a 24 hour time period, favoring the formation of a ssDNA peak (3rd peak). To confirm these results new samples were made and G-quartet formation and dissociation was monitored using the CD spectrophotometry (C) and (D). These results indicate that the α-subunit may lead to the formation G-quartets.

In 1993 Fang and Cech published a paper in which they demonstrated that the C-terminal portion of the β-subunit had the ability to mediate G-quartet formation, impelling the authors to suggest a biological function for the G-quartet (Fang and Cech, 1993). Since the publication of this paper the lack of direct evidence for the G-quartet have made those in the field of telomere biology increasingly wary of their formation and the ability of the β-subunit C-terminal domain to catalyze them has largely been ignored. The result presented by Paeschke *et al.* have renewed interest in this aspect of telomere biology, especially the role of the β-subunit's C-terminal domain. Interestingly, the results presented here suggest that, like the β-subunit, α may also play a role in mediating the formation of G-quartets. Indeed, this would be consistent with Paeschke *et al.* results in which they showed that the knockdown of α alone lead to a decrease in G-quartet epitope labeling in the macronucleus of *O. nova* (Paeschke et al., 2005).

Having said this, it is important to note that our results have proven to be difficult to reproduce and have many unexplained parameters. For instance, G-quartets do indeed appear in LiCl buffer only if the α-subunit is present, suggesting that α is in some way catalyzing their formation. However, the fact that these structures dissociate after ~24 hours of incubation with α suggest that they are indeed unstable and that the α-subunit has lost its ability to catalyze their formation. The question then becomes, why does the α-subunit lose this ability? The α-subunit has been shown to adopt higher order structures that may be incapable of catalyzing G-quartet formation. However, I demonstrated that these structures are completely absent in the LiCl samples. At this point I am unsure of how to explain the rapid appearance of G-quartets in the presence of the α-subunit and their subsequent disappearance and I feel that additional experimentation would have to be done before any conclusive arguments could be made for α's role in G-quartet formation.

Another unexplained parameter is the sodium/potassium dependence of higher order α-subunit complex formation. We have shown, by CD and gel-filtration, that a significant amount of G-quartets exists in a solution containing α regardless of buffering conditions. If then, G-quartets are the explanation for the appearance of higher-order complexes, as measured by analytical ultracentrifugation, one would expect to see them in all buffering conditions including those that lack sodium or potassium. This however is not the case, and higher-order complexes form in a potassium- or sodium-dependent manner. A possible explanation for this discrepancy may be kinetic in that the amount of G-quartet in a Li^+- only buffer decreases and is almost gone after 24 hours. However, this possibility is unlikely due to the fact that analytical ultracentrifugation data is collected over a 4-5 hour window and, based on our CD and gel-filtration data, a significant amount of G-quartet should still be present within this time span. Again, I have no plausible explanation for these discrepancies and I feel that additional experimentation would have to be done before any conclusive arguments could be made.

At the time these results were obtained, G-quartets were thought by the majority of the telomere field to be strictly an experimental complication that was to be avoided. I clearly demonstrated a great deal of difficulty in avoiding this complication, placing doubt on our ability to publish my results in any scientific journal. While we did consider the possible implications of α mediated G-quartet formation, we felt that, despite our every effort to eliminate experimental variables a great deal of our data could not be accounted for. These discrepancies, in combination with the lack of evidence for a biological role of G-quartets, lead to the cessation of the project. However, in light of the Paeschke *et al.* paper a revaluation of the role α may play in mediating G-quartets may indeed be a valuable endeavor.

Appendix I

Materials and Methods

Table of Contents

Poliovirus RNA-dependent RNA polymerase project

3D^{pol} Mutagenesis and Purification

Clones of the Mahoney type 1 strain of poliovirus polymerase were obtained from the Schultz Lab at C.U. Boulder. Mutations were introduced into the 3D^{pol} sequence by site-directed PCR mutagenesis and verified by DNA sequencing. All polymerases were expressed in 2xYT media with *E. coli* BL-21 (DE3) pLysS cells using inducible T7-based expression plasmids. Cells

were grown at 37°C to an $O.D._{600}$ of ~0.4, cooled to room temperature to an $O.D._{600}$ of ~0.6, and then induced by addition of IPTG to 0.5 mM. Growth was allowed to continue between 5 - 18 hours at room temperature before the cells were centrifuged into a pellet and frozen at -20°C. These cells were resuspended in 100 mM KPO_4 (pH 7.5), 0.5 mM EDTA, 0.02% NaN_3, 1:1000 β-mercaptoethanol and lysed by sonication. The cell debris was removed by centrifugation and the polymerase was precipitated with ammonium sulfate at 40% of saturation. The protein was resuspended in ~20 mL of 10 mM HEPES (pH 7.5), 0.01% NaN_3, 50 mM NaCl, 0.1 mM EDTA and 1:1000 β-mercaptoethanol, and dialyzed against two liters of the same solution overnight at 4°C. The protein was centrifuged, filtered through a 0.22 μM pore size membrane, loaded onto a S-sepharose column, and eluted with a linear gradient from 50 mM to 1 M NaCl in 25 mM HEPES (pH 8.5), 15 % glycerol, 0.1 mM EDTA, 0.02 % NaN_3, 1:1000 β-mercaptoethanol. Fractions containing the polymerase were pooled and diluted to reduce the NaCl concentration to ~0.15 M. This was then loaded onto a Q-sepharose column and eluted with a linear gradient to 1 M NaCl with 25 mM Tris (pH 8.5), 15 % glycerol, 0.1 mM EDTA, 0.02 % (w/v) NaN_3, and 1:1000 β-mercaptoethanol. The polymerase was concentrated in Centricon YM-30s (Millipore) and run over a Superdex 200 column equilibrated in 200 mM NaCl, 5 mM Tris (pH 7.5), 0.02% (w/v) sodium azide, and 2 mM DTT. Upon elution from this column, the proteins were >99 % pure as estimated from SDS-PAGE.

The protein pooled from the GF column was typically concentrated to approximately 10 mg/ml using Centricon YM-30s (Millipore). Concentrations were determined by the absorbance at 280 nm using extinction coefficients obtained from the Expasy web site (http://www.expasy.org/tools/protparam.html).

3Dpol Cystallization

Crystals were grown by hanging drop vapor diffusion at 16 °C using 8-16 mg/ml protein. Each mutant polymerase contained two separate mutations, L446D-R455D, that disrupts a persistent crystal lattice observed in the original structure that has since been shown to lead to disruption of the fingers domain(Hansen et al., 1997; Thompson and Peersen, 2004). Mutant polymerase crystals grew in four days with a precipitant/well solution containing 2 M sodium acetate, 0.1 M cacodylic acid (pH 6.9 - 7.1), 2 mM DTT and 0.02 % (w/v) sodium azide. All crystals were grown by mixing 2 μL of protein with 2 μL of precipitant/well solution on a siliconized cover slip and then placed over 1 mL of mother liquor in a 24-well tissue culture plate and sealed with Vaseline.

All crystals were harvested by first transferring them into micro-bridges (Hampton Research) containing 20 μL of precipitant/well solution and then transported into the 4°C cold room. Each crystal was then slowly equilibrated (25-30 minutes) in a final solution containing 250mM sodium acetate, 0.1 M cacodylic acid (pH 7.0), 2 mM DTT, 0.02 % (w/v) sodium azide and 30% (w/v) PEG-400. Ribonucleotide-divalent metal ion crystal "soaks" were accomplished by transfer of the crystals into a solution containing 250mM sodium acetate, 0.1 M cacodylic acid (pH 7.0), 2 mM DTT, 0.02 % (w/v) sodium azide, 30% (w/v) PEG-400, 10 mM NTP and 10mM Mg^{2+} for ~1 hour. All crystals were then flash frozen with liquid nitrogen.

$3D^{pol}$ Activity Assay

Twenty μL Poly(A)/oligo(dT) polymerase extension reaction mixtures containing 0.01 μg/μL poly(A) template (average length 300 nt), 0.005 μg/μL oligo(dT_{15}), 50 mM HEPES pH 7.5, 25 μM UTP, 0.5 mM each GTP, CTP, ATP, 4 mM DTT, 1.5 mM $MgAc_2$, 60 μM $ZnCl_2$, 0.1 % NP40, 0.01 Ci/μL [α-^{32}P]UTP, and 300 nM $3D^{pol}$ were assembled on ice in a 96-well microtiter plate. Each protein sample occupies an entire column (8-wells) or half-row (6-wells) so that time points

(between 10 - 30 minutes) can be quenched by addition of 30 μL of 0.5 mM EDTA to each well. Incorporation of radiolabel was evaluated by filtering ~35 μL of this reaction through Hybond N(+) nylon membrane supported on Whatman paper in a Schleicher & Schuell 96-well dotblot filter apparatus. The membrane is washed both before and after addition of the reaction mixtures with 25 mM MES (pH 6.5), 2.5 mM Mg $(OAC)_2$, 40 μM $ZnCl_2$, 10% glycerol (v/v), 2 mM DTT. Soaking the nylon membrane twice for 15 minutes in the above wash buffer further enhances the signal/noise. The amount of radio-labeled RNA bound to the membrane was quantitated using a phosphoimager. The relative activities were determined by graphing the amount of bound radio-labeled oligonucleotide retained by the Hybond N+ membrane vs. the various time points. Comparison of these slopes yielded the relative activities (Figure A1.1).

Fluorescence polarization assay

Measurements for fluorescence polarization were conducted in 384-well, black, flat-bottom, polystyrene plates (Corning). The 50 μL reaction mixtures contained 50 mM HEPES (pH 7.0), 4 mM DTT, 1.5 mM $MgAc_2$, 60 μM $ZnCl_2$ 0.1 % NP40, 30 mM NaCl (50 mM final in reaction including protein), 10 mM fluorescein-labeled RNA construct (Figure A1.2), and polymerase at various concentrations from 20 μM to 1 nM. $3D^{pol}$ was typically 1.2 - 1.5 fold serially diluted with 200 mM NaCl gel-filtration buffer. To ensure adequate mixing, small volumes of the protein dilutions (5 μL) were first transferred to the microtiter plate followed by larger volumes (45 μL) of a mixture containing the remaining components. Reaction mixtures were set up at 4°C and allowed to equilibrate for 30 minutes. The measurements were made in a Perkin Elmer/Wallac model 1420 Multilabel Counter using excitation and emission wavelengths of 480 and 535 nm (10 nm bandpass), respectively. Polarization was calculated as $mP = (I_{\parallel} - I_{\perp}) / (I_{\parallel} + I_{\perp}) \times 1000$, where $I_{\parallel}$

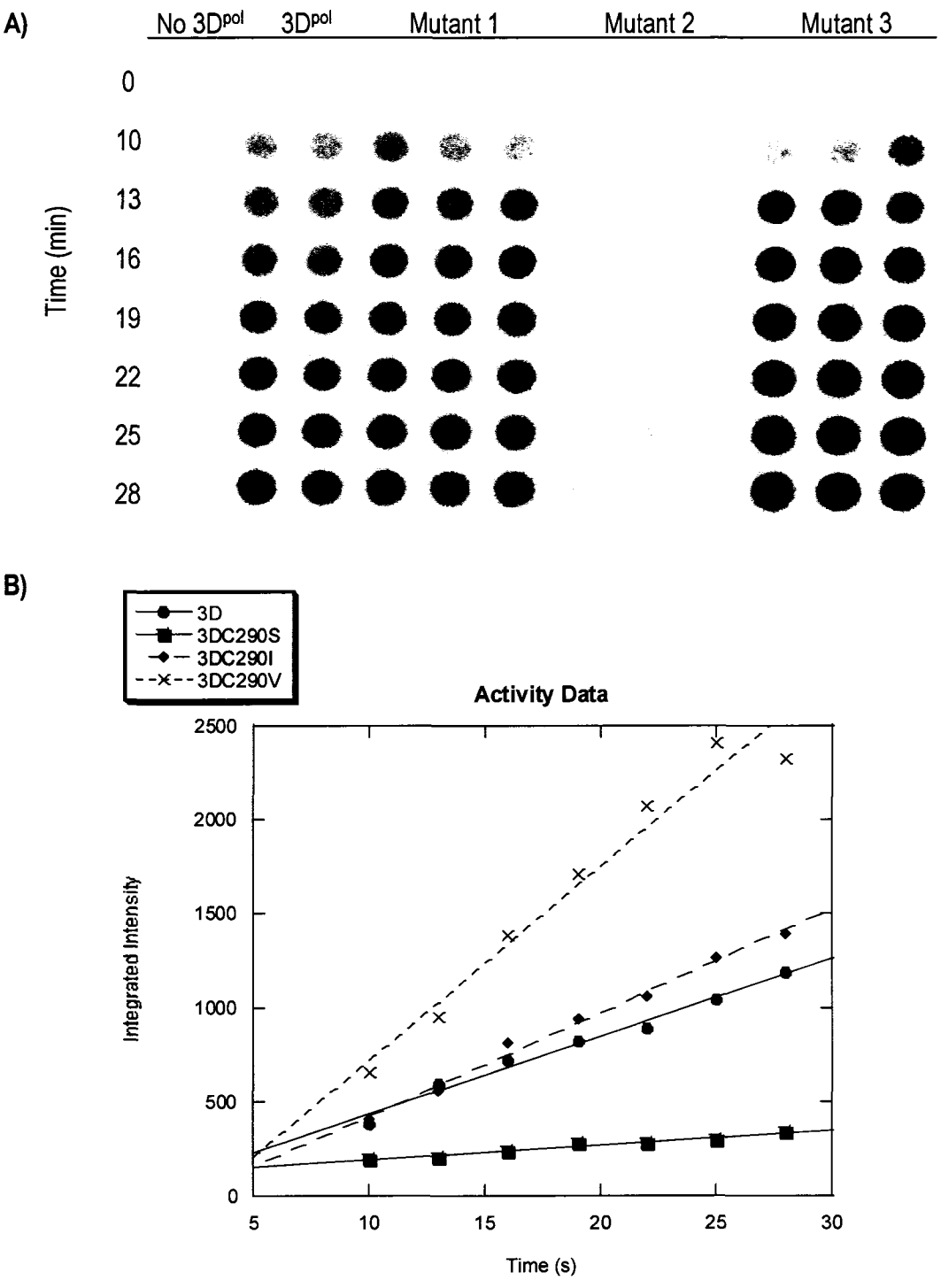


Figure A1.1: Activity assay. (**A**) Example of a Hybond N(+) filter bound with labeled nucleic acid after polymerase extension reaction mixtures are "pulled" through a 96-well dotblot apparatus. (**B**) Sample plot of integrated intensity from panel (A) versus time. The slopes relative to $3D^{pol}$ represent relative polymerase activity.

A)

B)

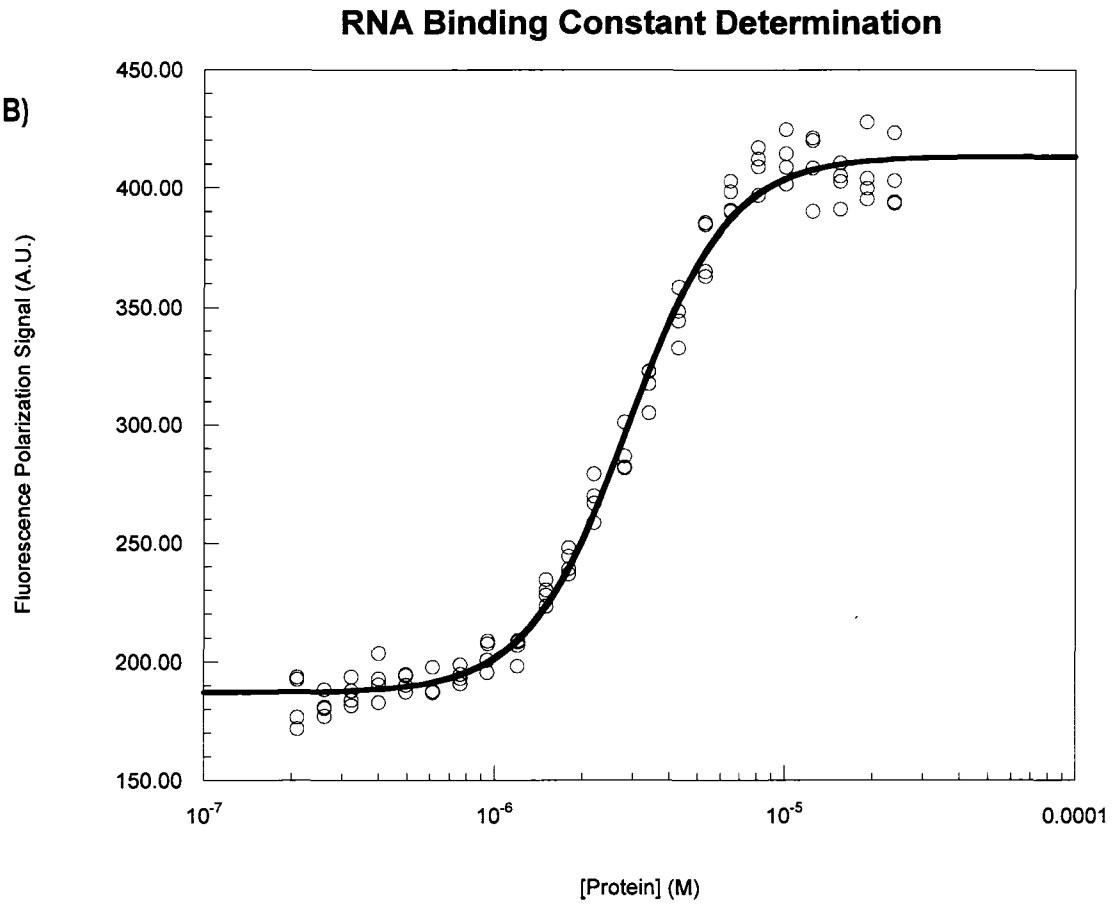


Figure A1.2: RNA-Binding Assay (**A**) RNA-binding affinities were determined using a fluorescence polarization assay employing a 5′-fluorescently labeled 26-base RNA hairpin that has been specifically designed to create a stable self-priming RNA template called Polymerase Elongation Template Element or PETE. The sequence of PETE leads to the formation of an 8-bp duplex region mediated by a 4-base GNRA tetra-loop, leaving a 6-base overhang acting as the template strand (**B**) Plot of milli-polarization versus log of protein concentration for determination of RNA binding constants to 3Dpol. Points were fit to the equation on page 127.

and $I_{\perp}$ are the fluorescence intensities polarized parallel and perpendicular, respectively, with reference to the plane-polarized excitation light. Equilibrium binding constants (K_ds) were determined by graphing the polarization values vs. the log of the protein concentration (Figure A1.2). Data were fit to a Hill coefficient model with baseline offset and scaling factor adjustments as follows:

$$\mathbf{mP} = \mathbf{S}\left(\frac{[\text{protein}]_{\text{total}}^{\ H}}{[\text{protein}]_{\text{total}}^{\ H} + K_d^{\ H}}\right) + \mathbf{B}$$

where **mP** is the milli-polarization which represents fractional binding, **S** is the scaling factor, **B** is the baseline correction, $\mathbf{[protein]_{total}}$ is the total protein concentration, and **H** is the Hill coefficient .

Translocation Assay

Translocation assays were performed using a 5′-P^{32} end-labeled PETE RNA identical to that used in the RNA binding assay minus the fluorescent label (Figure A1.2). Reaction samples contained 50mM HEPES pH 7.5, 4.0mM DTT, 1.5mM $MgCl_2$, 60uM $ZnCl_2$. 0.10% NP40, 5.5mM NaCl (transferred with protein), 147.5nM 5′-^{32}P end-labeled PETE and 2.5uM $3D^{pol}$. The correct NTP(s) were added up to a concentration of 25uM and samples were incubated for 30 minutes on ice. Each sample was than transferred to a room temperature water bath and quenched with EDTA at 5 minute intervals for up to 45minutes. 2 μl of each sample was then run on a denaturing 20% polyacrylamide gel containing 7M urea and 1xTBE. Bands were visualized using the phosphoimager.

Oxtricha nova Telomere end-binding protein project methods

Reagent Preparation

α and β were expressed in *E. coli* strain BL-21 pLysS using a T7-based expression plasmid. 750-ml cell cultures were grown at 37° C to an OD_{600} of ~0.4-0.6. Cells were then cooled to room temperature and expression was induced with the addition of IPTG to a final concentration of 0.4mM. Cells were allowed to express protein for ~15-16 hours, at which point they were collected in a centrifuge and frozen as pellets at -20° C. The pellets were then resuspended and lysed by sonication in a solution containing 50mM sodium phosphate (pH 8.0), 300mM NaCl, 10mM Imidazole, 1mM β-mercaptoethanol and 0.02% sodium azide. Each protein solution was then purified using a Ni-affinity column followed by S-Sepharose column and finally by a Superdex-200 gel filtration column (all columns were purchased from Pharmacia).

The single-strand DNA molecules were purchased from IDT ® and purified by HPLC using a Vydac reverse-phase C4 column. Each sample was then subjected to three rounds of lyophilization and resuspension in water to generate a highly soluble material.

Labeling reactions

Proteins were equilibrated in a solution containing 50mM LiCl, 25mM HEPES (pH 7.5), 0.1mM EDTA and 1mM TCEP. Labels were purchased from Molecular Probes® and brought up in 100% DMF. Labeling reactions were done at 50μM protein with a 2.5-3.5 fold excess of label for 4 hours. Each reaction was quenched with β-mercaptoethanol at a final concentration of ~1mM. The labeled protein was then purified away from excess unreacted label by ion exchange chromatography on an S-Sepharose column.

Circular Dichroism measurements

All CD spectra were collected on a Jasco J-720 CD spectrophotometer. The CD cell had a pathlength of 0.1 cm and the temperature was kept at 25° C by circulating water. T_4G_4 was equilibrated in various buffers with a final concentration of 60μM and α had a final concentration of 30μM. Spectra were collected with a scan rate of 100nm/min, 1-nm bandwidth, and a 4-sec response time.

Analytical Ultracentrifuge measurements

Analytical ultracentrifuge experiments were done with a Beckman XL-I. Sedimentation velocity experiments were run at a speed of 45000 rpm, a wavelength of 494nm and a temperature of 20°C.

Appendix II

Data Collection and Refinement Statistics

Content

$3D^{pol}$ G289A/C290V	A2.8	138
$3D^{pol}$ S288A	A2.9	139
$3D^{pol}$ S291P	A2.9	139
$3D^{pol}$ G292A	A2.10	140

Table A2.1-Data collection and refinement statistics

Mutant	$3D^{pol}$ G64S	$3D^{pol}$ G64S RTP bound	$3D^{pol}$ RTP bound	$3D^{pol}$ G64A
Space group	P6(5)	P6(5)	P6(5)	P6(5)
Unit cell	*a*=*b*=128.2 *c*=112.0	*a*=*b*=128.3 *c*=111.9	*a*=*b*=128.7 *c*=111.8	*a*=*b*=128.1 *c*=111.9
X-ray source	ALS 4.2.2-1.0 Å	Cu-Kα	Cu-Kα	ALS 4.2.2-1.0 Å
Resolution limits*	30.0-2.20 (2.27-2.20)	30.0-2.30 (2.38-2.30)	30.0-2.25 (2.32 -2.25)	30.0-2.10 (2.18-2.10)
Reflections				
Total collected	202677	266511	282531	229907
Unique	52204	46380	49268	60323
Redundancy*	3.88 (3.55)	5.75 (5.29)	5.79 (5.19)	3.81 (3.76)
I/σ*	12.2 (3.8)	5.6 (2.3)	8.6 (2.9)	10.7 (2.6)
Completeness (%)*	98.3 (97.4)	99.8 (99.4)	98.8 (97.8)	99.3 (99.9)
R_{merge} (%)*	6.3 (26.1)	12.5 (25.7)	9.9 (29.9)	6.5 (42.9)
Refinement				
Resolution range	30.0-2.20	30.0-2.30	30.0-2.25	30.0-2.10
R	22.8	24.1	21.1	24.0
R_{free}	24.8	26.7	23.3	25.8
Average *B*-factor	39.1	34.3	31.9	39.3

*Data in parentheses are for the highest resolution shell

Table A2.2-Data collection and refinement statistics

Mutant	$3D^{pol}$ G64D	$3D^{pol}$ G64N
Space group	P6(5)	P6(5)
Unit cell	*a*=*b*=128.7 *c*=112.2	*a*=*b*=128.7 *c*=111.9
X-ray source	Cu-Kα	Cu-Kα
Resolution limits*	30.0-2.30 (2.38-2.30)	30.0-2.60 (2.68-2.60)
Reflections		
Total collected	344968	241521
Unique	46029	32332
Redundancy*	7.49 (7.23)	7.47 (7.37)
I/σ*	17.9 (5.4)	13.9 (5.0)
Completeness (%)*	98.4 (97.9)	99.8 (100.0)
R_{merge} (%)*	6.2 (35.6)	9.4 (38.8)
Refinement		
Resolution range	30.0-2.30	30.0-2.60
R	23.0	21.9
R_{free}	25.9	24.2
Average *B*-factor	40.9	37.1

*Data in parentheses are for the highest resolution shell

Table A2.3-Data collection and refinement statistics

Mutant	$3D^{pol}$ C290E	$3D^{pol}$ C290E GTP bound	$3D^{pol}$ C290F	$3D^{pol}$ C290F GTP bound
Space group	P6(5)	P6(5)	P6(5)	P6(5)
Unit cell	*a*=*b*=128.8 *c*=111.7	*a*=*b*=129.0 *c*=111.5	*a*=*b*=128.7 *c*=111.9	*a*=*b*=128.7 *c*=111.4
X-ray source	Cu-Kα	Cu-Kα	Cu-Kα	Cu-Kα
Resolution limits*	30.0-2.30 (2.38-2.30)	30.0-2.20 (2.28-2.20)	30.0-2.30 (2.38-2.30)	30.0-2.20 (2.28-2.20)
Reflections				
Total collected	328004	372502	345165	360790
Unique	46381	52853	46552	52654
Redundancy*	7.07 (6.90)	7.05 (6.82)	7.41 (7.13)	6.85 (6.66)
I/σ*	14.6 (5.0)	11.6 (4.6)	12.8 (4.3)	13.5 (4.4)
Completeness (%)*	99.4 (100.0)	99.0 (97.9)	99.6 (100.0)	99.0 (100.0)
R_{merge} (%)*	7.4 (37.0)	8.6 (37.4)	7.0 (34.8)	7.5 (42.1)
Refinement				
Resolution range	30.0-2.30	30.0-2.20	30.0-2.30	30.0-2.20
R	21.5	22.0	22.2	22.8
R_{free}	23.9	24.3	24.5	25.3
Average *B*-Factor				
Residues 283-287	27.6	23.9	29.2	28.0
Residues 288-292	43.2	31.2	42.7	49.2
Residues 293-297	29.1	25.3	28.3	28.3

*Data in parentheses are for the highest resolution shell

Table A2.4-Data collection and refinement statistics

Mutant	$3D^{pol}$ C290I	$3D^{pol}$ C290I GTP bound	$3D^{pol}$ C290M	$3D^{pol}$ C290M GTP bound
Space group	P6(5)	P6(5)	P6(5)	P6(5)
Unit cell	*a*=*b*=129.0 *c*=11.9	*a*=*b*=128.9 *c*=111.8	*a*=*b*=128.3 *c*=111.6	*a*=*b*=129.1 *c*=111.6
X-ray source	Cu-Kα	Cu-Kα	Cu-Kα	Cu-Kα
Resolution limits*	30.0-2.20 (2.28-2.20)	30.0-2.40 (2.48-2.40)	30.0-2.35 (2.43-2.35)	30.0-2.30 (2.38-230)
Reflections				
Total collected	388942	303346	317205	333798
Unique	52853	41229	43350	46300
Redundancy*	7.36 (6.93)	7.36 (7.10)	7.32 (7.01)	7.20 (6.71)
I/σ*	13.2 (4.4)	8.5 (3.1)	7.90 (3.2)	11.2 (4.4)
Completeness (%)*	98.6 (97.2)	99.9 (100.0)	99.8 (99.4)	98.7 (98.1)
R_{merge} (%)*	9.4 (44.3)	15.4 (36.7)	15.2 (39.4)	9.1 (37.8)
Refinement				
Resolution range	30.0-2.20	30.0-2.40	30.0-2.35	30.0-2.30
R	22.6	22.0	22.2	21.9
R_{free}	25.1	25.0	24.9	24.6
Average *B*-Factor				
Residues 283-287	25.3	30.1	31.2	30.1
Residues 288-292	42.3	51.2	52.2	51.2
Residues 293-297	25.2	28.8	27.2	28.8

*Data in parentheses are for the highest resolution shell

Table A2.5-Data collection and refinement statistics

Mutant	$3D^{pol}$ C290S	$3D^{pol}$ C290S GTP bound	$3D^{pol}$ C290V	$3D^{pol}$ C290V GTP bound
Space group	P6(5)	P6(5)	P6(5)	P6(5)
Unit cell	*a*=*b*=129.5 *c*=112.4	*a*=*b*=129.3 *c*=111.8	*a*=*b*=128.7 *c*=111.8	*a*=*b*=128.8 *c*=111.5
X-ray source	ALS 4.2.2-1.0 Å	Cu-Kα	Cu-Kα	Cu-Kα
Resolution limits*	45.0-1.85 (1.92-1.85)	30.0-2.30 (2.38-2.30)	30.0-2.20 (2.28-2.20)	30.0-2.40 (2.48-240)
Reflections				
Total collected	615519	345609	360872	307527
Unique	91000	47090	53309	40894
Redundancy*	6.76 (6.62)	7.34 (7.06)	6.77 (6.55)	7.52 (7.42)
I/σ*	9.8 (2.3)	8.2 (4.2)	15.3 (4.5)	14.6 (5.1)
Completeness (%)*	100.0 (100.0)	99.8 (100.0)	100.0 (100.0)	99.6 (100.0)
R_{merge} (%)*	7.4 (48.2)	14.6 (41.7)	7.2 (39.2)	8.5 (38.4)
Refinement				
Resolution range	30.0-2.00	30.0-2.30	30.0-2.20	30.0-2.40
R	22.7	21.9	22.2	21.3
R_{free}	24.0	24.2	24.8	23.3
Average *B*-Factor				
Residues 283-287	24.0	28.2	23.5	26.3
Residues 288-292	31.1	37.6	32.5	33.6
Residues 293-297	22.2	25.8	21.7	25.3

*Data in parentheses are for the highest resolution shell

Table A2.6-*Data collection and refinement statistics*

Mutant	$3D^{pol}$ G289A	$3D^{pol}$ G289A GTP bound	$3D^{pol}$ G289A/C290F	$3D^{pol}$ G289A/C290F GTP bound
Space group	P6(5)	P6(5)	P6(5)	P6(5)
Unit cell	*a*=*b*= 129.1 *c*= 112.5	*a*=*b*= 128.7 *c*= 111.3	*a*=*b*= 128.5 *c*= 111.6	*a*=*b*= 128.8 *c*= 111.8
X-ray source	ALS 4.2.2-1.0 Å	Cu-Kα	Cu-Kα	Cu-Kα
Resolution limits*	30.0-1.95 (2.02-1.95)	30.0-2.30 (2.38-2.30)	30.0-2.30 (2.38-2.30)	30.0-2.30 (2.38-2.30)
Reflections				
Total collected	526621	641561	304190	347934
Unique	77401	46335	46277	46492
Redundancy*	6.80 (6.68)	7.37 (7.13)	6.57 (6.44)	7.48 (7.33)
I/σ*	11.0 (2.0)	14.4 (4.3)	11.9 (3.7)	14.8 (5.6)
Completeness (%)*	100.0 (100.0)	99.7 (100.0)	99.6 (100.0)	99.5 (99.5)
R_{merge} (%)*	6.1 (49.7)	7.0 (38.5)	7.5 (35.8)	8.0 (36.0)
Refinement				
Resolution range	30.0- 2.1	30.0-2.3	30.0- 2.3	30.0-2.3
R	22.5	21.7	21.7	21.4
R_{free}	24.6	24.4	24.0	23.8
Average *B*-Factor				
Residues 283-287	27.2	29.6	26.7	24.1
Residues 288-292	42.8	38.2	35.6	32.9
Residues 293-297	26.6	27.7	25.3	23.1

*Data in parentheses are for the highest resolution shell

Table A2.7-Data collection and refinement statistics

Mutant	$3D^{pol}$ G289A/C290I	$3D^{pol}$ G289A/C290I GTP bound	$3D^{pol}$ G289A/C290M	$3D^{pol}$ G289A/C290M GTP bound
Space group	P6(5)	P6(5)	P6(5)	P6(5)
Unit cell	*a*=*b*=128.8 *c*= 111.7	*a*=*b*= 129.0 *c*= 111.8	*a*=*b*= 1128.8 *c*= 111.7	*a*=*b*= 128.8 *c*= 111.5
X-ray source	Cu-Kα	Cu-Kα	Cu-Kα	Cu-Kα
Resolution limits*	30.0-2.10 (2.18-2.10)	30.0-2.20 (2.28-2.20)	30.0-2.20 (2.28-2.20)	30.0-2.30 (2.38-2.30)
Reflections				
Total collected	455274	391453	397793	336399
Unique	61261	53436	53150	46044
Redundancy*	7.43 (7.23)	7.33 (7.00)	7.48 (7.33)	7.31 (6.89)
I/σ*	13.9 (4.9)	13.7 (5.1)	14.1 (4.0)	173 (5.1)
Completeness (%)*	99.8 (100)	99.7 (99.8)	99.7 (100.0)	98.8 (98.7)
R_{merge} (%)*	7.1 (36.1)	7.3 (33.8)	6.4 (38.3)	6.5 (36.9)
Refinement				
Resolution range	30.0-2.10	30.0-2.20	30.0- 2.20	30.0-2.30
R	22.2	21.8	22.1	21.5
R_{free}	24.2	23.9	23.7	23.9
Average *B*-Factor				
Residues 283-287	24.0	24.7	31.6	31.8
Residues 288-292	29.3	29.5	62.1	57.2
Residues 293-297	23.0	23.1	28.9	30.4

*Data in parentheses are for the highest resolution shell

Table A2.8-Data collection and refinement statistics

Mutant	$3D^{pol}$ G289A/C290S	$3D^{pol}$ G289A/C290S GTP bound	$3D^{pol}$ G289A/C290V	$3D^{pol}$ G289A/C290V GTP bound
Space group	P6(5)	P6(5)	P6(5)	P6(5)
Unit cell	*a*=*b*=129.4 *c*= 111.4	*a*=*b*=128.8 *c*=111.3	*a*=*b*=129.2 *c*=111.6	*a*=*b*=128.6 *c*=111.5
X-ray source	Cu-Kα	Cu-Kα	Cu-Kα	Cu-Kα
Resolution limits*	30.0-2.40 (2.48-2.40)	30.0-2.80 (2.88-2.80)	30.0-2.60 (2.68-2.60)	30.0-2.50 (2.58-2.50)
Reflections				
Total collected	306907	191395	243888	271053
Unique	40979	25489	32293	35805
Redundancy*	7.49 (7.29)	7.51 (7.48)	7.55 (7.45)	7.57 (7.49)
I/σ*	11.4 (4.5)	10.3 (4.3)	14.0 (5.5)	12.4 (5.3)
Completeness (%)*	99.0 (98.5)	98.6 (98.0)	99.2 (99.8)	98.7 (98.9)
R_{merge} (%)*	11.4 (42.0)	16.8 (44.4)	9.3 (31.1)	9.2 (32.7)
Refinement				
Resolution range	30.0-2.40	30.0-2.80	30.0-2.60	30.0-2.50
R	21.3	20.7	20.8	21.2
R_{free}	24.2	24.5	23.6	24.0
Average *B*-Factor				
Residues 283-287	23.0	16.4	21.8	23.9
Residues 288-292	22.9	15.8	26.6	25.1
Residues 293-297	20.3	12.3	20.0	21.1

*Data in parentheses are for the highest resolution shell

Table A2.9-Data collection and refinement statistics

Mutant	$3D^{pol}$ S288A	$3D^{pol}$ Ser288 GTP	$3D^{pol}$ S291P	$3D^{pol}$ S291P GTP bound
Space group	P6(5)	P6(5)	P6(5)	P6(5)
Unit cell	*a*=*b*= 129.3 *c*=112.5	*a*=*b*= 129.0 *c*= 111.4	*a*=*b*=129.4 *c*=111.9	*a*=*b*=128.9 *c*=111.3
X-ray source	ALS 4.2.2-1.0 Å	Cu-Kα	Cu-Kα	Cu-Kα
Resolution limits*	40.0-2.0 (2.07-2.00)	30.0-2.20 (2.28-2.20)	30.0-2.50 (2.58-2.50)	30.0-2.20 (2.28-2.20)
Reflections				
Total collected	488012	393731	264980	394628
Unique	71990	52036	36788	52927
Redundancy*	6.8 (6.7)	7.57 (7.37)	7.20 (6.76)	7.46 (7.22)
I/σ*	10.7 (2.4)	20.3 (5.5)	7.3 (2.7)	14.7 (4.8)
Completeness (%)*	100 (99.9)	97.4 (95.7)	99.0 (100.0)	99.5 (100.0)
R_{merge} (%)*	6.3 (43.1)	6.1 (35.4)	18.0 (35.8)	6.5 (33.5)
Refinement				
Resolution range	30.0- 2.00	30.0-2.20	30.0-2.50	30.0-2.20
R	23.4	22.0	22.2	22.5
R_{free}	25.6	24.6	25.5	25.5
Average *B*-Factor				
Residues 283-287	30.3	28.6	27.8	27.0
Residues 288-292	55.8	53.7	50.3	42.0
Residues 293-297	28.2	26.3	28.0	27.2

*Data in parentheses are for the highest resolution shell

Table A2.10-Data collection and refinement statistics

Mutant	$3D^{pol}$ G292A	$3D^{pol}$ G292A GTP bound
Space group	P6(5)	P6(5)
Unit cell	*a*=*b*= 129.0 *c*=112.3	*a*=*b*= 129.2 *c*= 111.6
X-ray source	ALS 4.2.2-1.0 Å	Cu-Kα
Resolution limits*	30.0-1.95 (2.02-1.95)	30.0-2.4 (2.49-2.40)
Reflections		
Total collected	519719	304662
Unique	76329	40956
Redundancy*	6.8 (6.7)	7.4 (7.2)
I/σ*	9.1 (3.2)	12.3 (3.8)
Completeness (%)*	99.0 (98.0)	99.0 (98.3)
R_{merge} (%)*	8.8 (40.6)	8.5 (37.2)
Refinement		
Resolution range	30.0-2.10	30.0-2.40
R	22.9	21.6
R_{free}	25.3	24.9
Average *B*-Factor		
Residues 283-287	26.7	27.7
Residues 288-292	46.2	54.7
Residues 293-297	26.0	30.0

*Data in parentheses are for the highest resolution shell

Appendix III

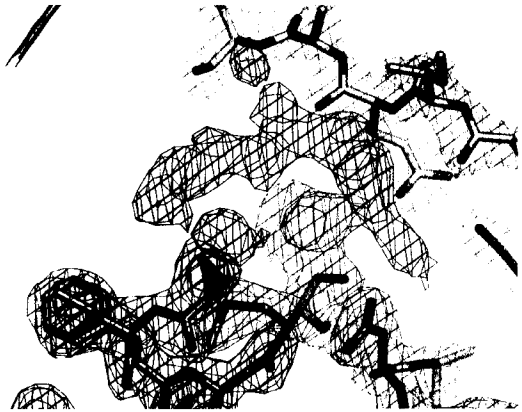
Electron Density for the 288-292 Loop

Purpose of this Appendix

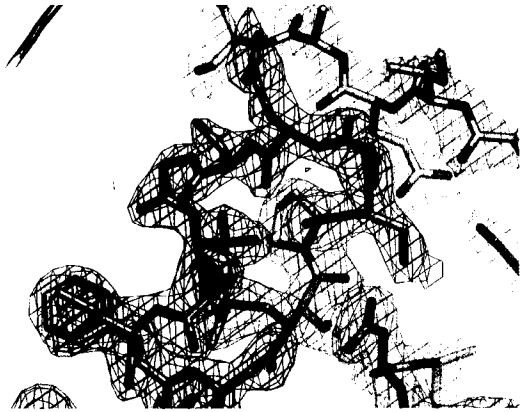
This appendix contains 15 panels each consisting of 6 individual representations of the electron density surrounding a single mutation in the 288-292 loop. The top three figures represent the electron density surrounding the 288-292 loop mutant in its apo form. The bottom three figures represent the electron density surrounding the 288-292 loop mutant in its GTP bound form. The three representations of electron density for either the apo or GTP bound form differ in how each map was calculated. The first representation is the initial 2FoFc difference map that was calculated using models that had residues 287-293 missing so as to avoid model bias. As such this figure is missing a model for the loop. The second representation is the final 2FoFc difference map that was calculated using the model from the last round of refinement. The third representation is a 1000K composite simulated anneal omit map. All three maps have been contoured at 1.6 σ.

The loop in all representation is orange, surrounded by a blue map and starts at the right of the figure with Ser288. Directly above and behind the loop is the ring finger (yellow) that is surrounded by a yellow map and shows a stick representation of each residue including Asp177 that forms a salt bridge with Lys61 (green, surrounded by a yellow map) in the majority of the structures solved. Finally, Asp238 (gray) is directly below the loop surrounded by a red map.

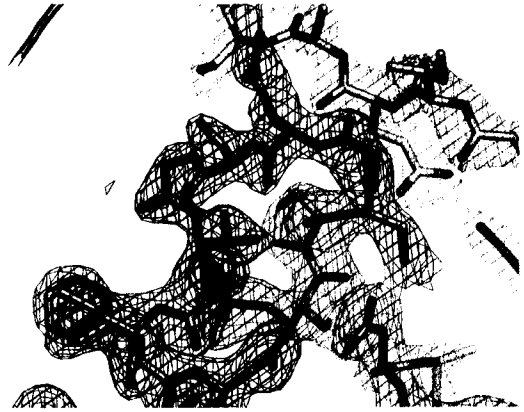
Contents



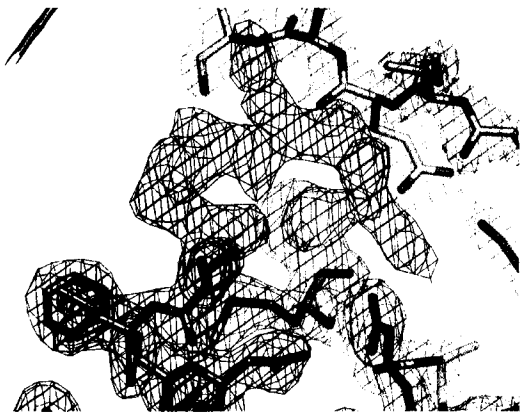
3DDD C290E Initial 2FoFc map.
2.3 Å map contoured at 1.6σ



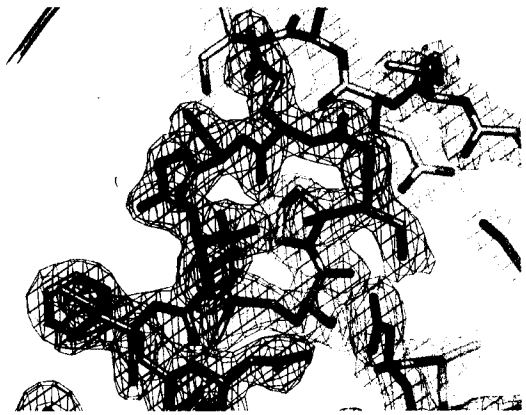
3DDD C290E Final 2FoFc map.
2.3 Å map contoured at 1.6σ



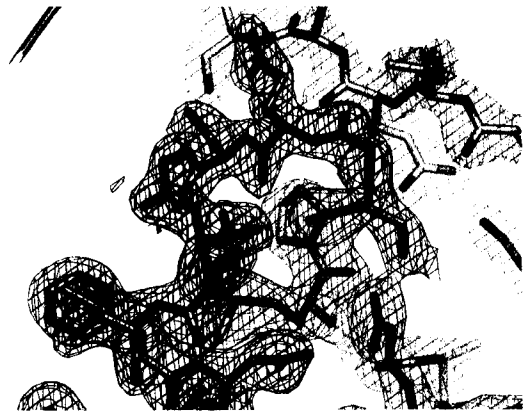
3DDD C290E composite simulated anneal omit map. 2.3 Å map contoured at 1.6σ



GTP bound 3DDD C290E Initial 2FoFc map.
2.2 Å map contoured at 1.6σ

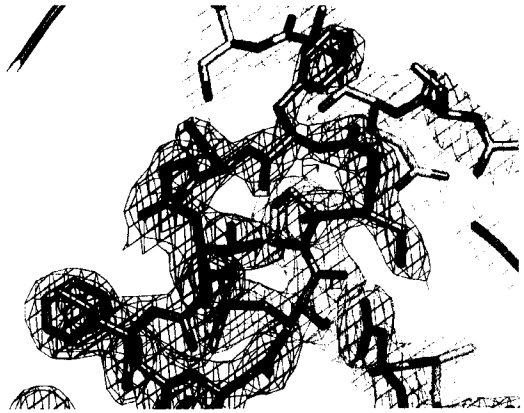


GTP bound 3DDD C290E Final 2FoFc map.
2.2 Å map contoured at 1.6σ



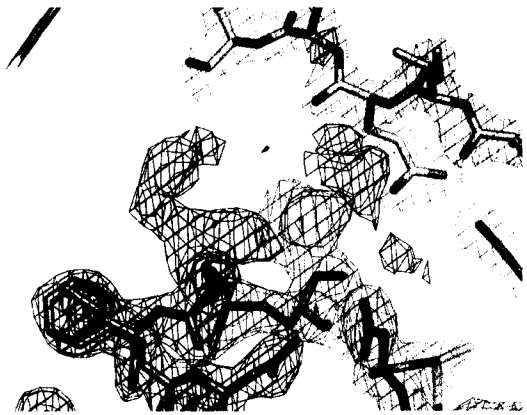
GTP bound 3DDD C290E composite simulated anneal omit map.
2.2 Å map contoured at 1.6σ

3DDD C290F Initial 2FoFc map.
2.3 Å map contoured at 1.6σ

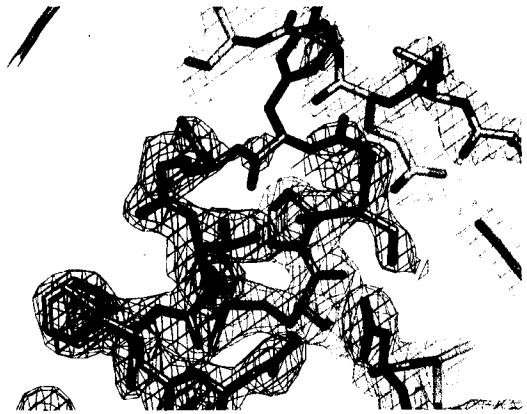


3DDD C290F Final 2FoFc map.
2.3 Å map contoured at 1.6σ

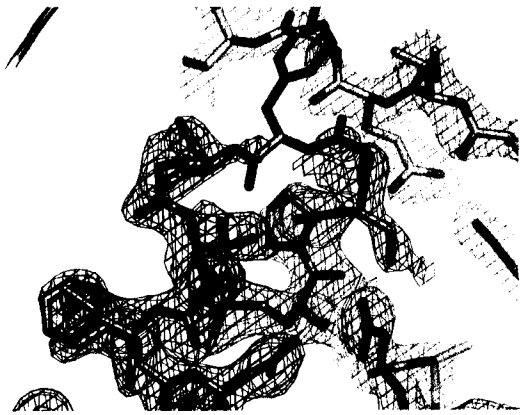
3DDD C290F composite simulated anneal omit map. 2.3 Å map contoured at 1.6σ



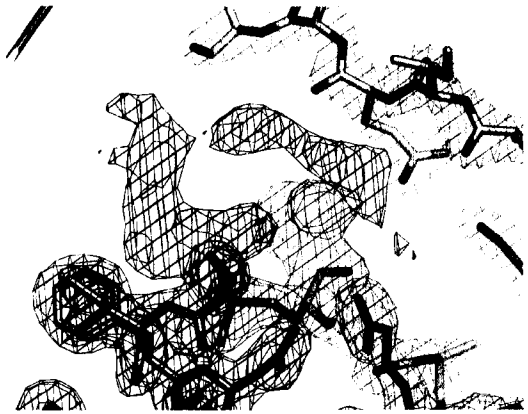
GTP bound 3DDD C290F Initial 2FoFc map.
2.2 Å map contoured at 1.6σ



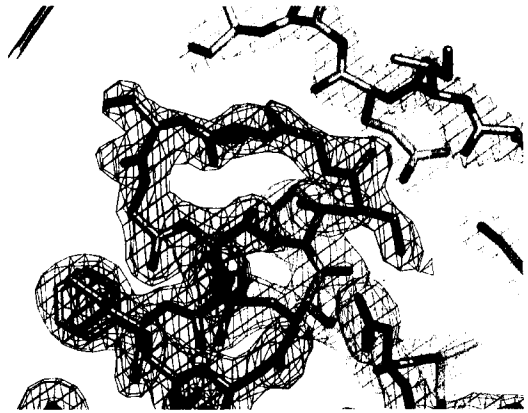
GTP bound 3DDD C290F Final 2FoFc map.
2.2 Å map contoured at 1.6σ



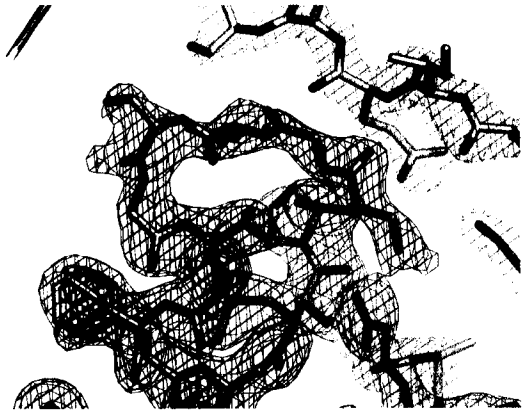
GTP bound 3DDD C290F composite simulated anneal omit map.
2.2 Å map contoured at 1.6σ



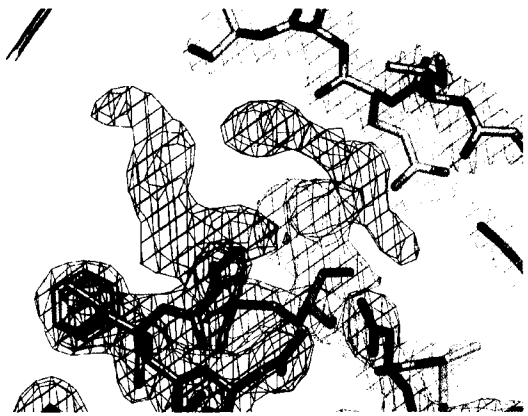
3DDD C290I Initial 2FoFc map.
2.2 Å map contoured at 1.6σ



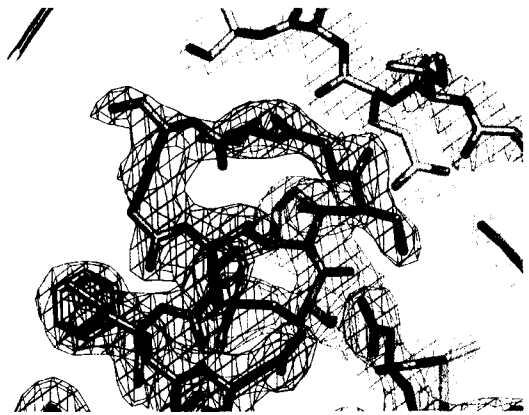
3DDD C290I Final 2FoFc map.
2.2 Å map contoured at 1.6σ



3DDD C290I composite simulated anneal omit map. 2.2 Å map contoured at 1.6σ

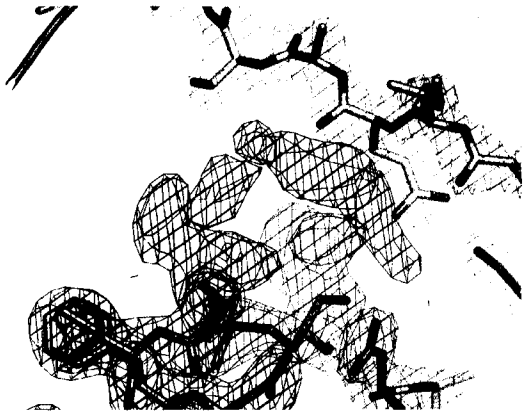


GTP bound 3DDD C290I Initial 2FoFc map.
2.4 Å map contoured at 1.6σ



GTP bound 3DDD C290I Final 2FoFc map.
2.4 Å map contoured at 1.6σ

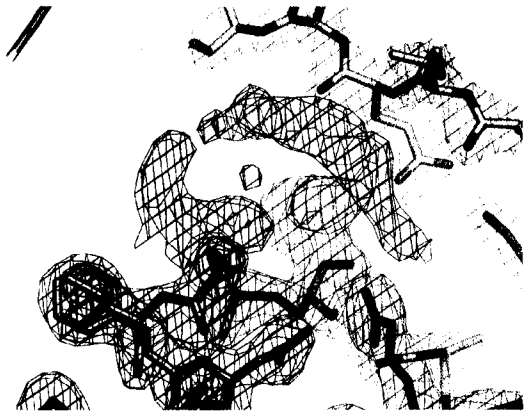
GTP bound 3DDD C290I composite simulated anneal omit map.
2.4 Å map contoured at 1.6σ



3DDD C290M Initial 2FoFc map.
2.35 Å map contoured at 1.6σ

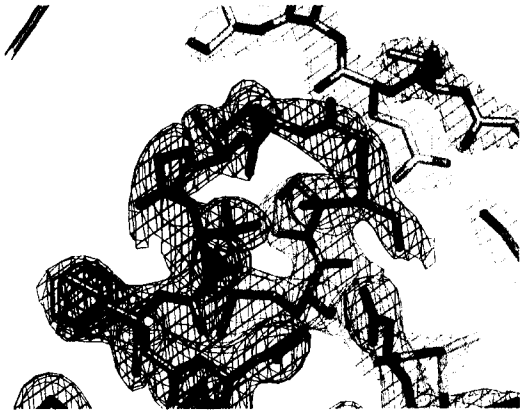
3DDD C290M Final 2FoFc map.
2.35 Å map contoured at 1.6σ

3DDD C290M composite simulated anneal omit map. 2.35 Å map contoured at 1.6σ

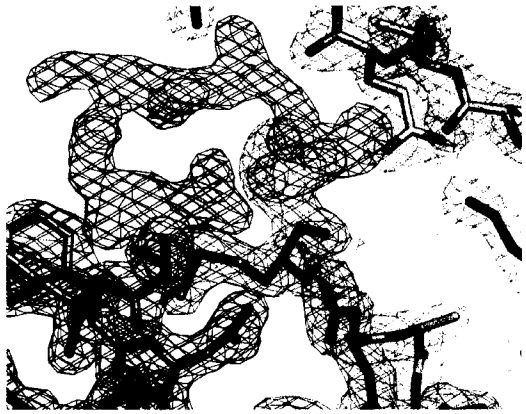


GTP bound 3DDD C290M Initial 2FoFc map.
2.3 Å map contoured at 1.6σ

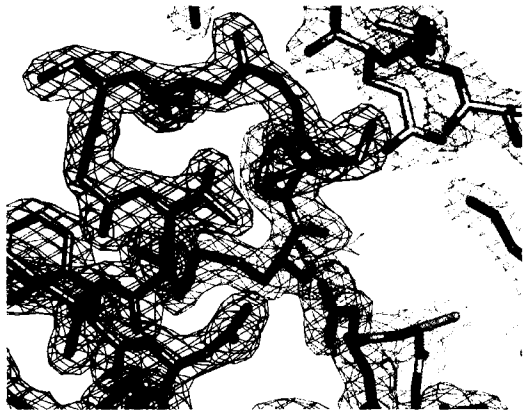
GTP bound 3DDD C290M Final 2FoFc map.
2.3 Å map contoured at 1.6σ



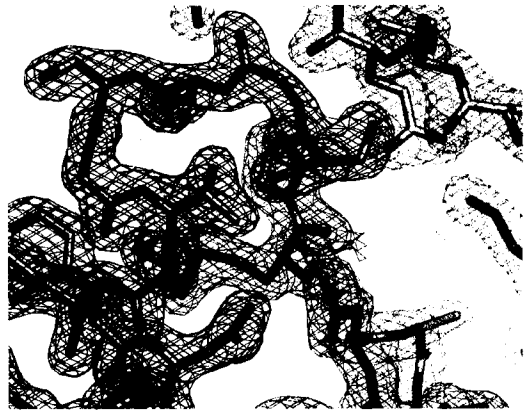
GTP bound 3DDD C290M composite simulated anneal omit map.
2.3 Å map contoured at 1.6σ



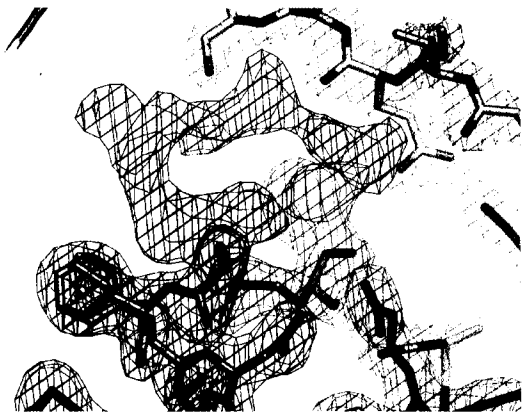
3DDD C290S Initial 2FoFc map.
2.0 Å map contoured at 1.6σ



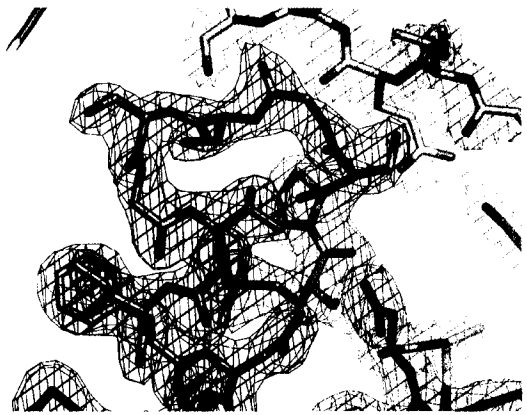
3DDD C290S Final 2FoFc map.
2.0 Å map contoured at 1.6σ



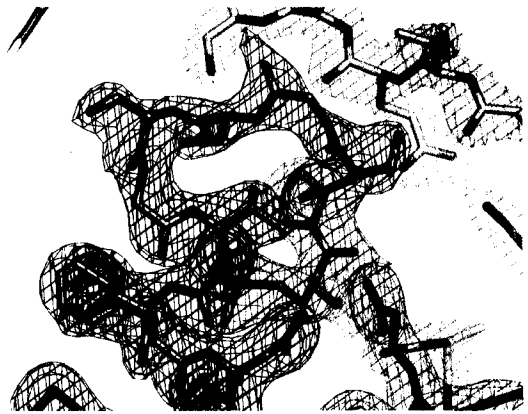
3DDD C290S composite simulated anneal omit map. 2.0 Å map contoured at 1.6σ



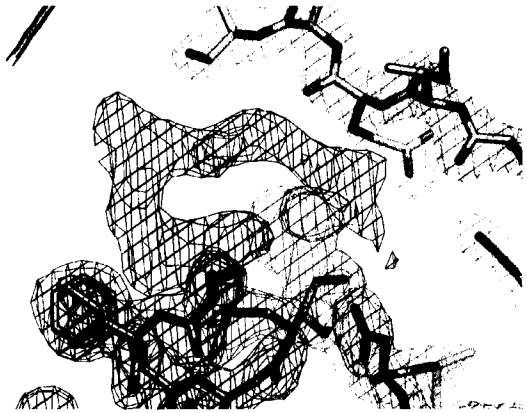
GTP bound 3DDD C290S Initial 2FoFc map.
2.3 Å map contoured at 1.6σ



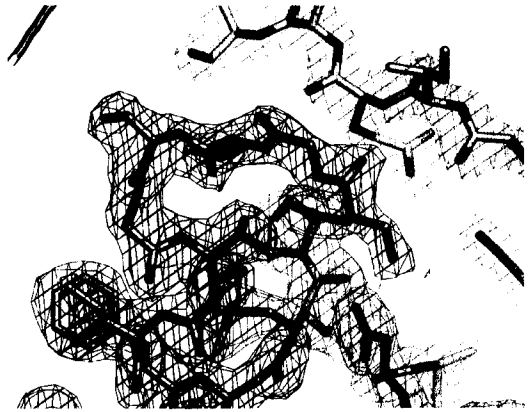
GTP bound 3DDD C290S Final 2FoFc map.
2.3 Å map contoured at 1.6σ



GTP bound 3DDD C290S composite simulated anneal omit map.
2.3 Å map contoured at 1.6σ

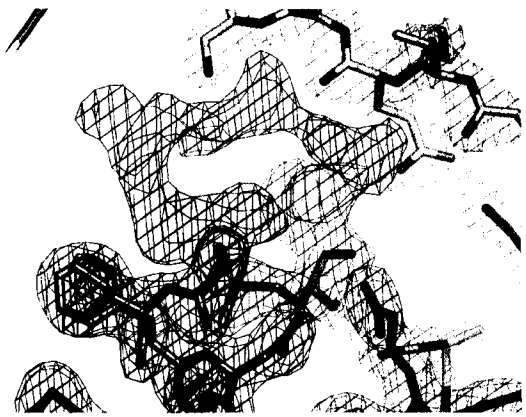


3DDD C290V Initial 2FoFc map.
2.2 Å map contoured at 1.6σ



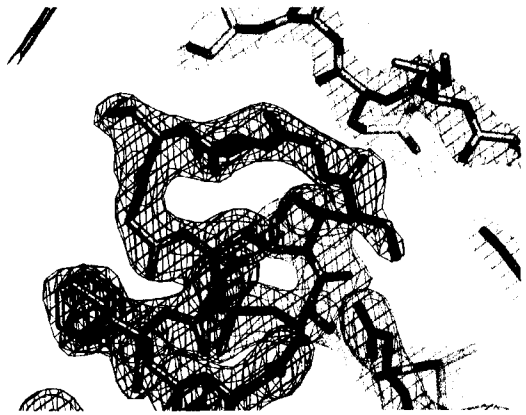
3DDD C290V Final 2FoFc map.
2.2 Å map contoured at 1.6σ

3DDD C290V composite simulated anneal omit map. 2.2 Å map Contoured at 1.6σ

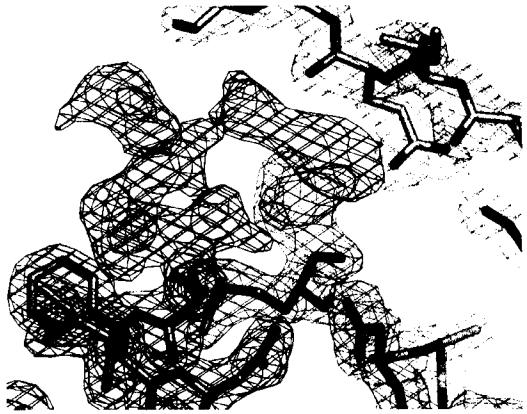


GTP bound 3DDD C290V Initial 2FoFc map.
2.4 Å map contoured at 1.6σ

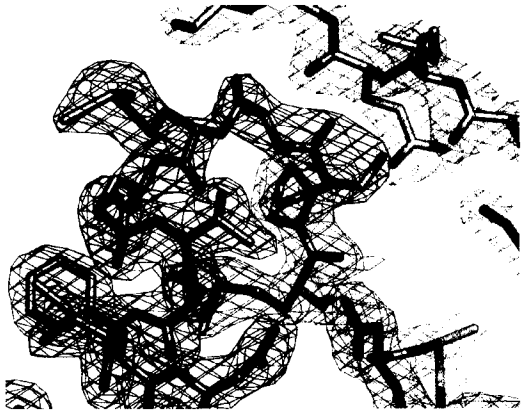
GTP bound 3DDD C290V Final 2FoFc map.
2.4 Å map contoured at 1.6σ



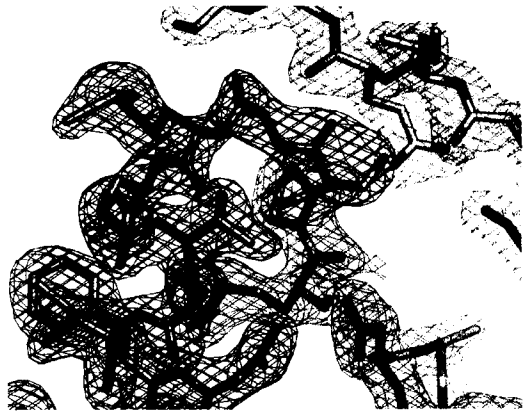
GTP bound 3DDD C290V composite simulated anneal omit map.
2.4 Å map contoured at 1.6σ



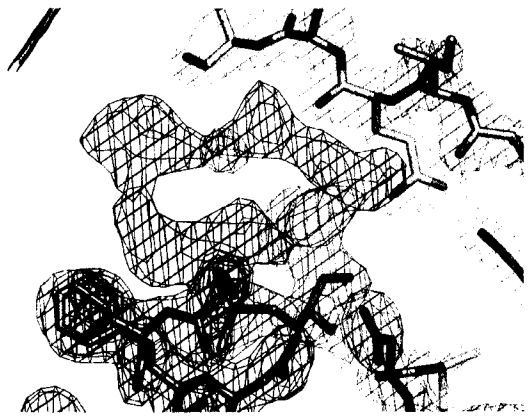
3DDD G289A Initial 2FoFc map.
2.1 Å map contoured at 1.6σ



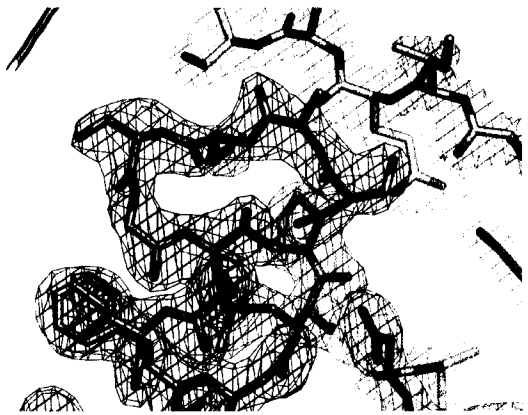
3DDD G289A Final 2FoFc map.
2.1 Å map contoured at 1.6σ



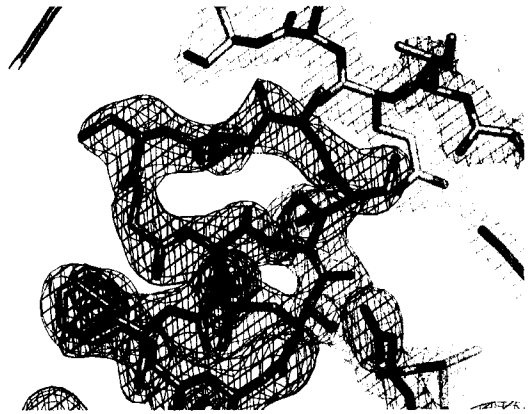
3DDD G289A composite simulated anneal omit map. 2.1 Å map contoured at 1.6σ



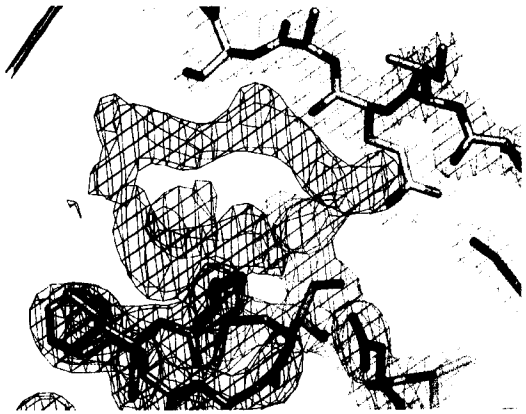
GTP bound 3DDD G289A Initial 2FoFc map.
2.3 Å map contoured at 1.6σ



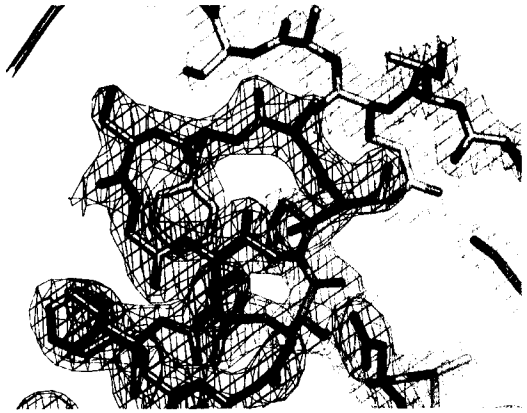
GTP bound 3DDD G289A Final 2FoFc map
2.3 Å map contoured at 1.6σ



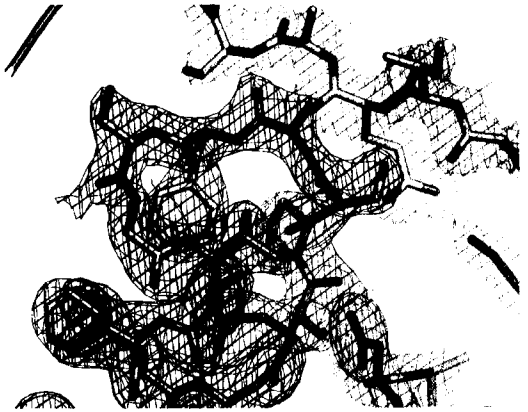
GTP bound 3DDD G289A composite simulated anneal omit map.
2.3 Å map contoured at 1.6σ



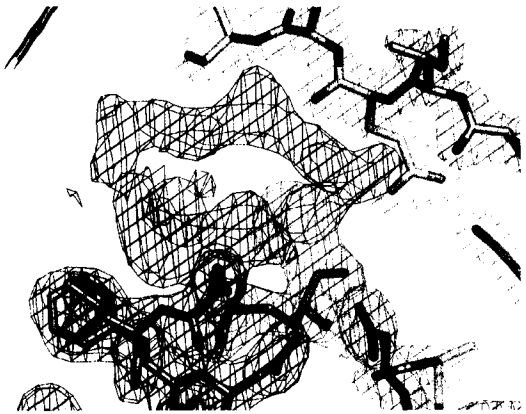
3DDD G289A/C290F Initial 2FoFc map.
2.3 Å map contoured at 1.6σ



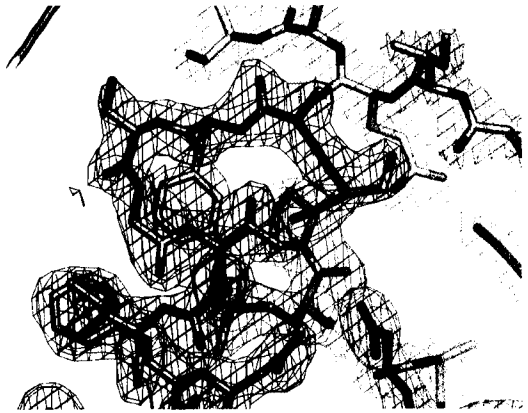
3DDD G289A/C290F Final 2FoFc map.
2.3 Å map contoured at 1.6σ



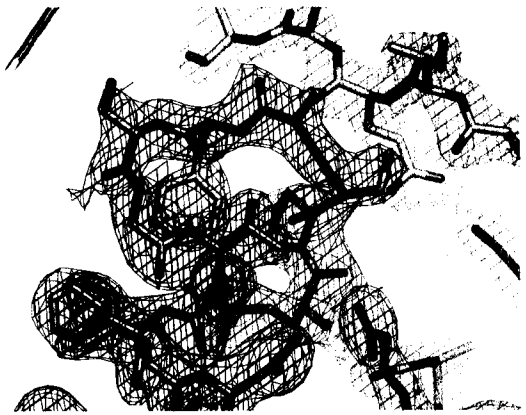
3DDD G289A/C290F composite simulated anneal omit map.
2.3 Å map contoured at 1.6σ



GTP bound 3DDD G289A/C290F Initial 2FoFc map.
2.3 Å map contoured at 1.6σ



GTP bound 3DDD G289A/C290F Final 2FoFc map.
2.3 Å map contoured at 1.6σ



GTP bound 3DDD G289A/C290F composite simulated anneal omit map.
2.3 Å map contoured at 1.6σ

3DDD G289A/C290I Initial 2FoFc map.
2.1 Å map contoured at 1.6σ

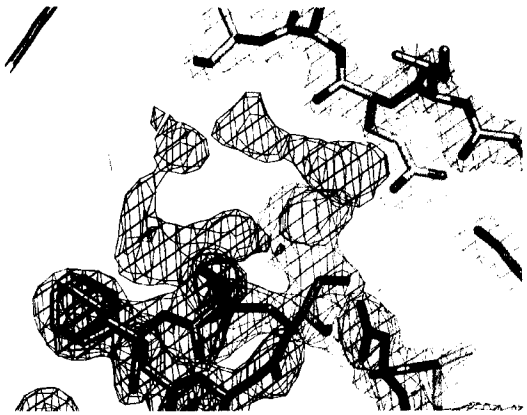
3DDD G289A/C290I Final 2FoFc map.
2.1 Å map contoured at 1.6σ

3DDD G289A/C290I composite simulated anneal omit map.
2.1 Å map contoured at 1.6σ

GTP bound 3DDD G289A/C290I Initial 2FoFc map.
2.2 Å map contoured at 1.6σ

GTP bound 3DDD G289A/C290I Final 2FoFc map.
2.2 Å map contoured at 1.6σ

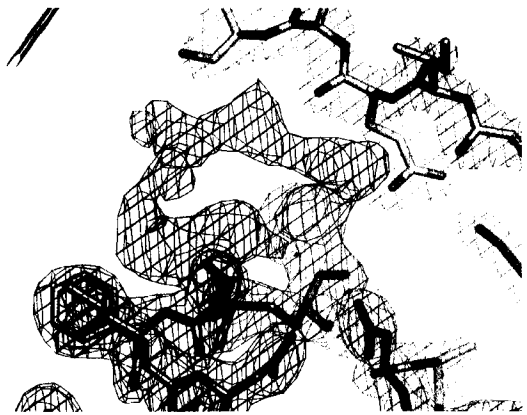
GTP bound 3DDD G289A/C290I composite simulated anneal omit map.
2.2 Å map contoured at 1.6σ



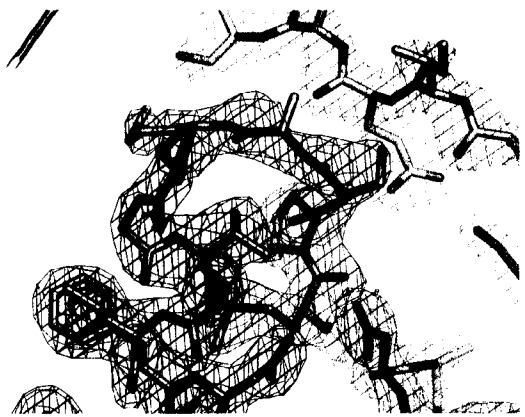
3DDD G289A/C290M Initial 2FoFc map.
2.2 Å map contoured at 1.6σ

3DDD G289A/C290M Final 2FoFc map.
2.2 Å map contoured at 1.6σ

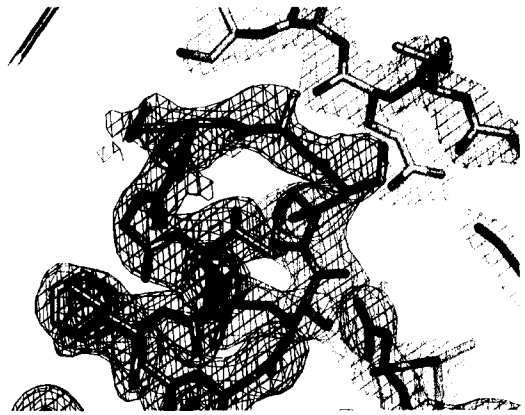
3DDD G289A/C290M composite simulated anneal omit map.
2.2 Å map contoured at 1.6σ



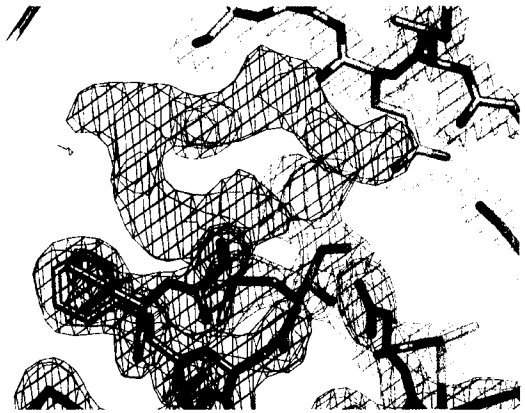
GTP bound 3DDD G289A/C290M Initial 2FoFc map.
2.3 Å map contoured at 1.6σ



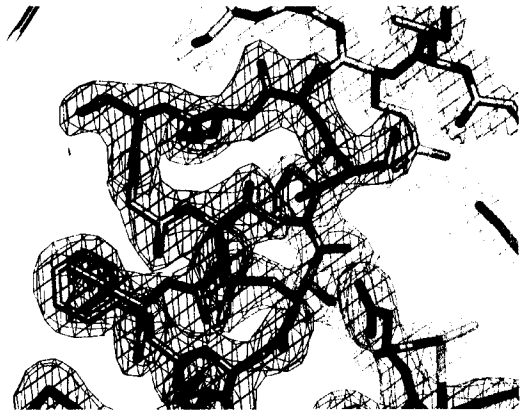
GTP bound 3DDD G289A/C290M Final 2FoFc map.
2.3 Å map contoured at 1.6σ



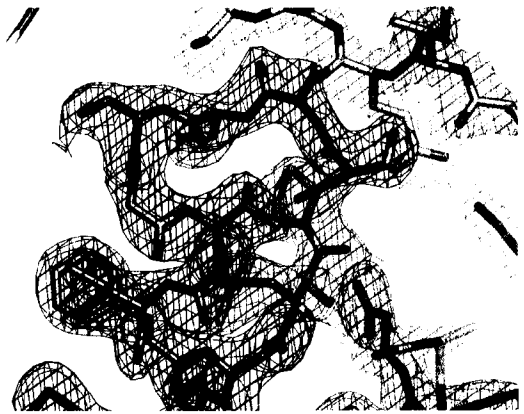
GTP bound 3DDD G289A/C290M composite simulated anneal omit map.
2.3 Å map contoured at 1.6σ



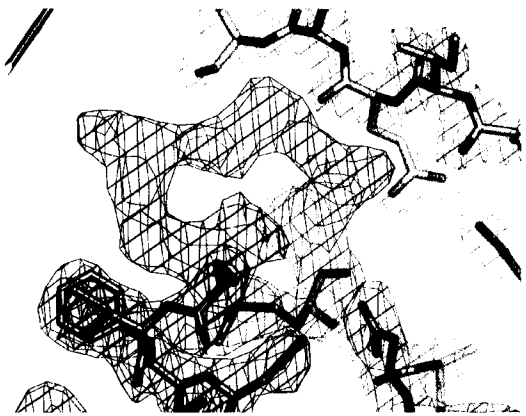
3DDD G289A/C290S Initial 2FoFc map.
2.4 Å map contoured at 1.6σ



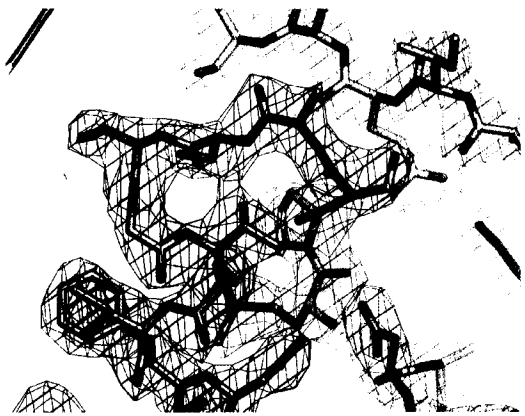
3DDD G289A/C290S Final 2FoFc map.
2.4 Å map contoured at 1.6σ



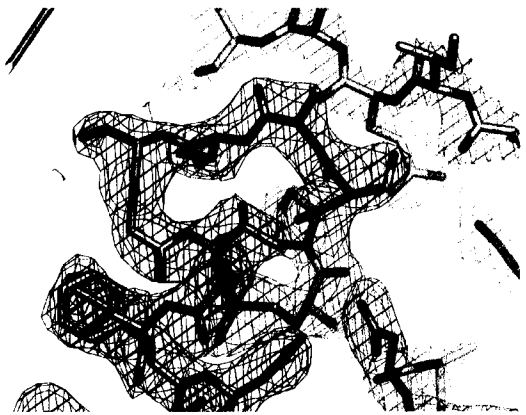
3DDD G289A/C290S composite
simulated anneal omit map.
2.4 Å map contoured at 1.6σ



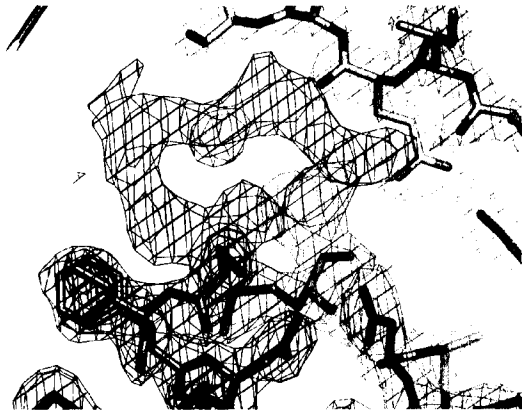
GTP bound 3DDD G289A/C290S Initial
2FoFc map.
2.8 Å map contoured at 1.6σ



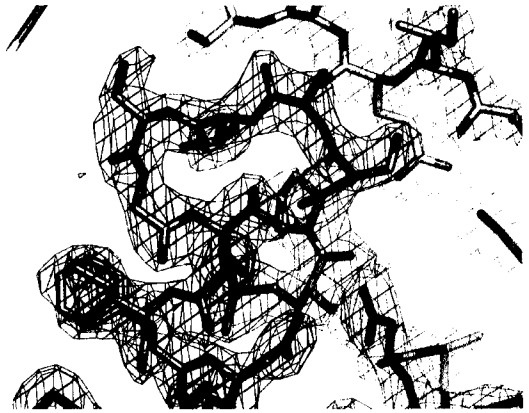
GTP bound 3DDD G289A/C290S Final
2FoFc map.
2.8 Å map contoured at 1.6σ



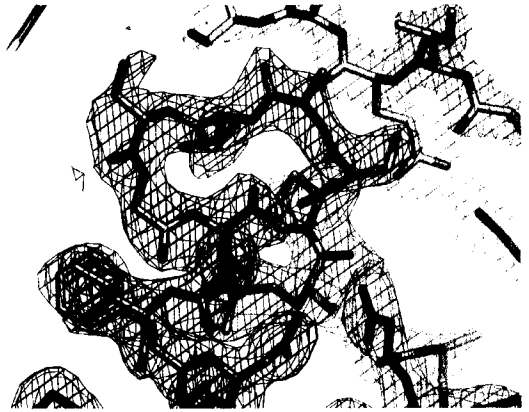
GTP bound 3DDD G289A/C290S composite
simulated anneal omit map.
2.8 Å map contoured 1.6σ



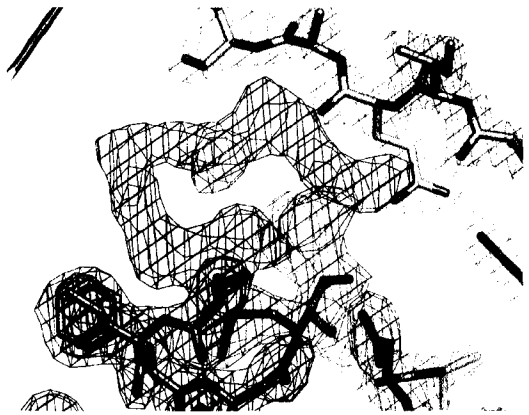
3DDD G289A/C290V Initial 2FoFc map.
2.6 Å map contoured at 1.6σ



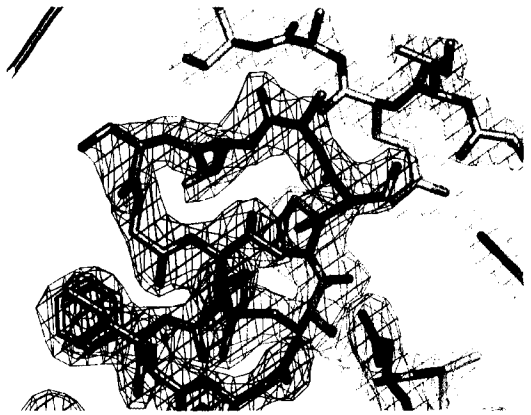
3DDD G289A/C290V Final 2FoFc map.
2.6 Å map contoured at 1.6σ



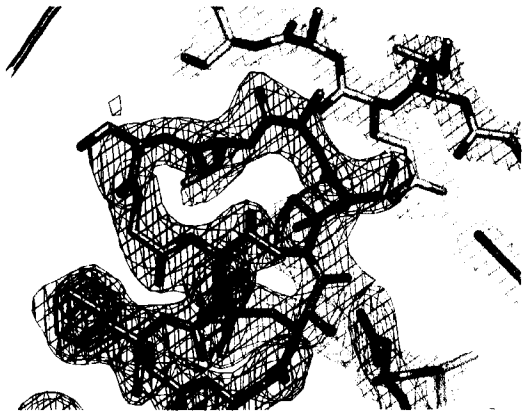
3DDD G289A/C290V composite simulated anneal omit map.
2.6 Å map contoured at 1.6σ



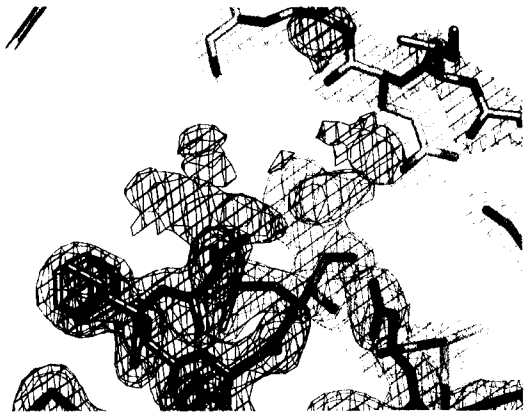
GTP bound 3DDD G289A/C290V Initial 2FoFc map.
2.5 Å map contoured at 1.6σ



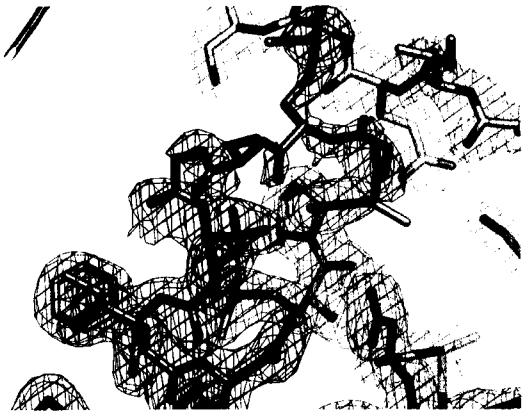
GTP bound 3DDD G289A/C290V Final 2FoFc map.
2.5 Å map contoured at 1.6σ



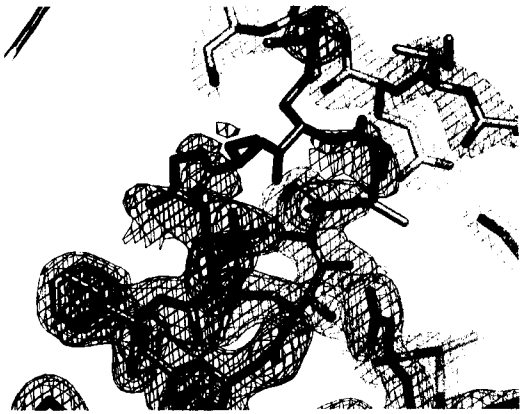
GTP bound 3DDD G289A/C290V composite simulated anneal omit map.
2.5 Å map contoured at 1.6σ



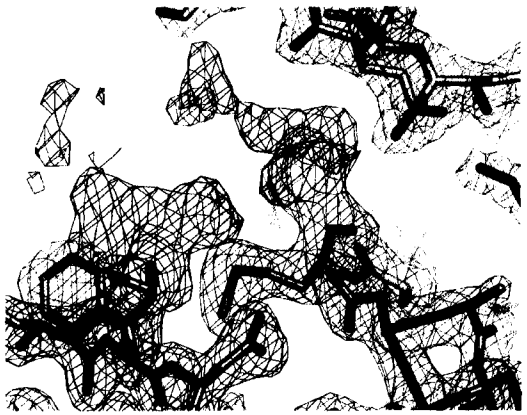
3DDD S288A Initial 2FoFc map.
2.0 Å map contoured at 1.6σ



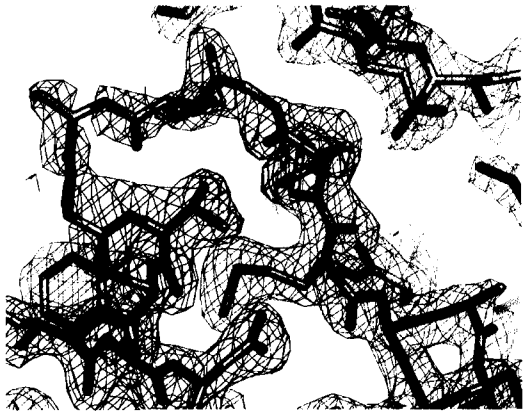
3DDD S288A Final 2FoFc map.
2.0 Å map contoured at 1.6σ



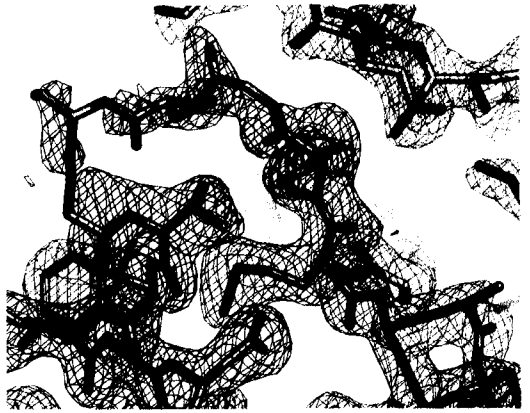
3DDD S288A composite simulated anneal omit map. 2.0 Å map contoured at 1.6σ



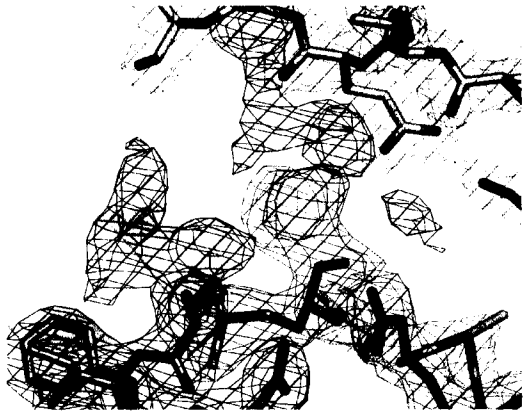
GTP bound 3DDD S288A Initial 2FoFc map.
2.2 Å map contoured at 1.6σ



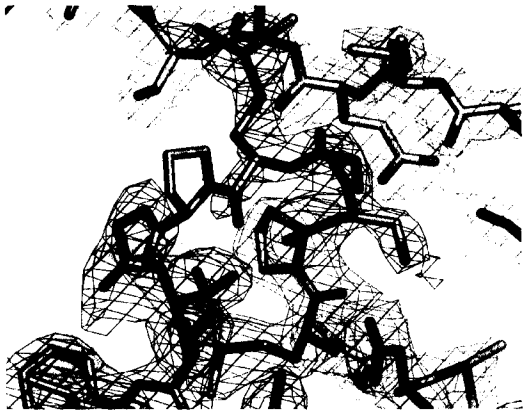
GTP bound 3DDD S288A Final 2FoFc map.
2.2 Å map contoured at 1.6σ



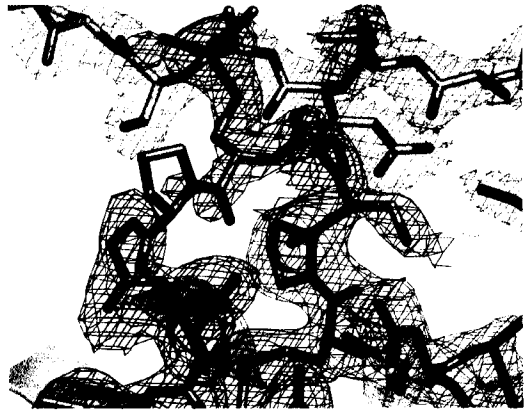
GTP bound 3DDD S288A composite simulated anneal omit map.
2.2 Å map contoured at 1.6σ



3DDD S291P Initial 2FoFc map.
2.5 Å map contoured at 1.6σ

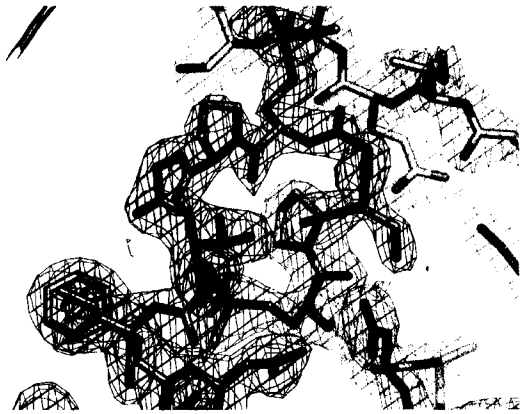


3DDD S291P Final 2FoFc map.
2.5 Å map contoured at 1.6σ



3DDD S291P composite simulated anneal omit map. 2.5 Å map contoured at 1.6σ

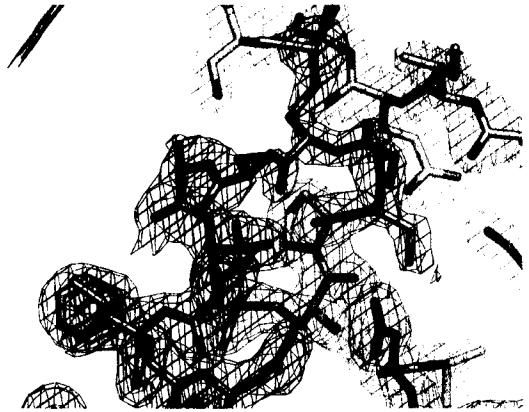
GTP bound 3DDD S291P Initial 2FoFc map.
2.2 Å map contoured at 1.6σ



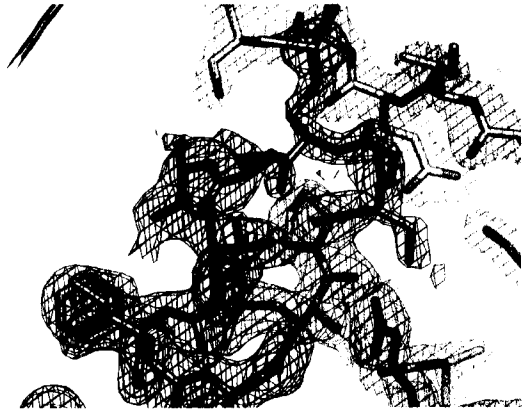
GTP bound 3DDD S291P Final 2FoFc map.
2.2 Å map contoured at 1.6σ

GTP bound 3DDD S291P simulated anneal composite omit map.
2.2 Å map contoured at 1.6σ

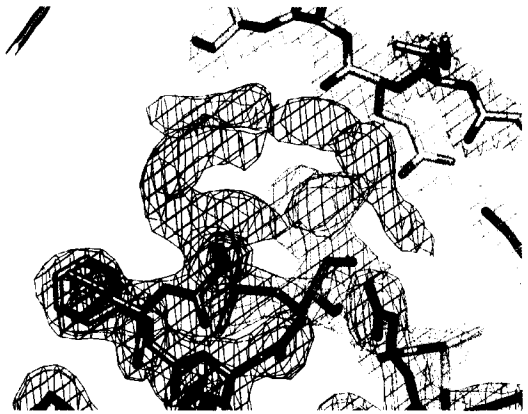
3DDD G292A Initial 2FoFc map.
2.1 Å map contoured at 1.6σ



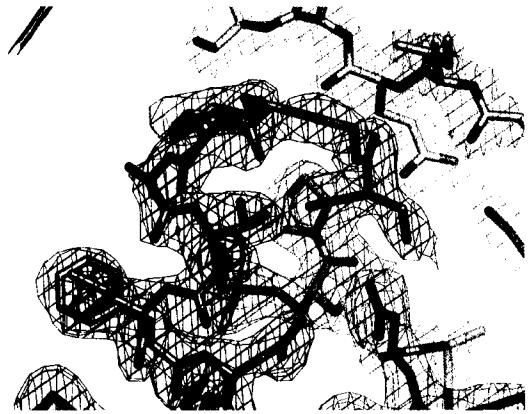
3DDD G292A Final 2FoFc map.
2.1 Å map contoured at 1.6σ



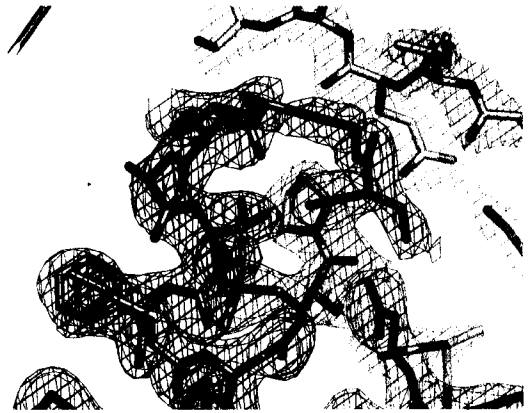
3DDD G292A composite simulated anneal omit map. 2.1 Å map contoured at 1.6σ



GTP bound 3DDD G292A Initial 2FoFc map.
2.4 Å map contoured at 1.6σ



GTP bound 3DDD G292A Final 2FoFc map.
2.4 Å map contoured at 1.6σ



GTP bound 3DDD G292A composite simulated anneal omit map.
2.4 Å map contoured at 1.6σ

References

Abraham, R., Minor, P., Dunn, G., Modlin, J.F. and Ogra, P.L. (1993) Shedding of virulent poliovirus revertants during immunization with oral poliovirus vaccine after prior immunization with inactivated polio vaccine. *J Infect Dis*, **168**, 1105-1109.

Ago, H., Adachi, T., Yoshida, A., Yamamoto, M., Habuka, N., Yatsunami, K. and Miyano, M. (1999) Crystal structure of the RNA-dependent RNA polymerase of hepatitis C virus. *Structure*, **7**, 1417-1426.

Agol, V. (2002) in *Molecular Biology of Picornaviruses*, eds. Semler,B. and Wimmer, E. (ASM Press), pp 127-148.

Ambros, V. and Baltimore, D. (1980) Purification and properties of a HeLa cell enzyme able to remove the 5'-terminal protein from poliovirus RNA. *J Biol Chem*, **255**, 6739-6744.

Ambros, V., Pettersson, R.F. and Baltimore, D. (1978) An enzymatic activity in uninfected cells that cleaves the linkage between poliovirion RNA and the 5' terminal protein. *Cell*, **15**, 1439-1446.

Andino, R., Rieckhof, G.E., Achacoso, P.L. and Baltimore, D. (1993) Poliovirus RNA synthesis utilizes an RNP complex formed around the 5'-end of viral RNA. *Embo J*, **12**, 3587-3598.

Andino, R., Rieckhof, G.E. and Baltimore, D. (1990) A functional ribonucleoprotein complex forms around the 5' end of poliovirus RNA. *Cell*, **63**, 369-380.

Andrei, G. and De Clercq, E. (1993) Molecular approaches for the treatment of hemorrhagic fever virus infections. *Antiviral Res*, **22**, 45-75.

Appleby, T.C., Luecke, H., Shim, J.H., Wu, J.Z., Cheney, I.W., Zhong, W., Vogeley, L., Hong, Z. and Yao, N. (2005) Crystal structure of complete rhinovirus RNA polymerase suggests front loading of protein primer. *J Virol*, **79**, 277-288.

Arnold, J.J. and Cameron, C.E. (2004) Poliovirus RNA-dependent RNA polymerase (3Dpol): pre-steady-state kinetic analysis of ribonucleotide incorporation in the presence of Mg2+. *Biochemistry*, **43**, 5126-5137.

Arnold, J.J., Gohara, D.W. and Cameron, C.E. (2004) Poliovirus RNA-dependent RNA polymerase (3Dpol): pre-steady-state kinetic analysis of ribonucleotide incorporation in the presence of Mn2+. *Biochemistry*, **43**, 5138-5148.

Arnold, J.J., Vignuzzi, M., Stone, J.K., Andino, R. and Cameron, C.E. (2005) Remote site control of an active site fidelity checkpoint in a viral RNA-dependent RNA polymerase. *J Biol Chem*, **280**, 25706-25716.

Barton, D.J., O'Donnell, B.J. and Flanegan, J.B. (2001) 5' cloverleaf in poliovirus RNA is a cis-acting replication element required for negative-strand synthesis. *Embo J*, **20**, 1439-1448.

Baumann, P. and Cech, T.R. (2001) Pot1, the putative telomere end-binding protein in fission yeast and humans. *Science*, **292**, 1171-1175.

Bougie, I. and Bisaillon, M. (2004) The broad spectrum antiviral nucleoside ribavirin as a substrate for a viral RNA capping enzyme. *J Biol Chem*, **279**, 22124-22130.

Bressanelli, S., Tomei, L., Roussel, A., Incitti, I., Vitale, R.L., Mathieu, M., De Francesco, R. and Rey, F.A. (1999) Crystal structure of the RNA-dependent RNA polymerase of hepatitis C virus. *Proc Natl Acad Sci U S A*, **96**, 13034-13039.

Butcher, S.J., Grimes, J.M., Makeyev, E.V., Bamford, D.H. and Stuart, D.I. (2001) A mechanism for initiating RNA-dependent RNA polymerization. *Nature*, **410**, 235-240.

Carpousis, A.J. and Gralla, J.D. (1980) Cycling of ribonucleic acid polymerase to produce oligonucleotides during initiation in vitro at the lac UV5 promoter. *Biochemistry*, **19**, 3245-3253.

Campagnola, G.C., Scoggin, K., Weigand, M., and Peersen, O. (2006). Structure of the Coxsackie virus RNA-Dependent RNA polymerase reveals common structural features among picornaviral polymerases. article in press.

Cheetham, G.M., Jeruzalmi, D. and Steitz, T.A. (1999) Structural basis for initiation of transcription from an RNA polymerase-promoter complex. *Nature*, **399**, 80-83.

Cheetham, G.M. and Steitz, T.A. (1999) Structure of a transcribing T7 RNA polymerase initiation complex. *Science*, **286**, 2305-2309.

Chen, C.Y. and Sarnow, P. (1995) Initiation of protein synthesis by the eukaryotic translational apparatus on circular RNAs. *Science*, **268**, 415-417.

Choi, K.H., Groarke, J.M., Young, D.C., Kuhn, R.J., Smith, J.L., Pevear, D.C. and Rossmann, M.G. (2004) The structure of the RNA-dependent RNA polymerase from bovine viral diarrhea virus establishes the role of GTP in de novo initiation. *Proc Natl Acad Sci U S A*, **101**, 4425-4430.

Counter, C.M., Avilion, A.A., LeFeuvre, C.E., Stewart, N.G., Greider, C.W., Harley, C.B. and Bacchetti, S. (1992) Telomere shortening associated with chromosome instability is arrested in immortal cells which express telomerase activity. *Embo J*, **11**, 1921-1929.

Crotty, S., Cameron, C.E. and Andino, R. (2001) RNA virus error catastrophe: direct molecular test by using ribavirin. *Proc Natl Acad Sci U S A*, **98**, 6895-6900.

Crotty, S., Maag, D., Arnold, J.J., Zhong, W., Lau, J.Y., Hong, Z., Andino, R. and Cameron, C.E. (2000) The broad-spectrum antiviral ribonucleoside ribavirin is an RNA virus mutagen. *Nat Med*, **6**, 1375-1379.

Crowder, S. and Kirkegaard, K. (2005) Trans-dominant inhibition of RNA viral replication can slow growth of drug-resistant viruses. *Nat Genet*, **37**, 701-709.

Cummings, K.J., Lee, S.M., West, E.S., Cid-Ruzafa, J., Fein, S.G., Aoki, Y., Sulkowski, M.S. and Goodman, S.N. (2001) Interferon and ribavirin vs interferon alone in the re-treatment of chronic hepatitis C previously nonresponsive to interferon: A meta-analysis of randomized trials. *Jama*, **285**, 193-199.

Dahlberg, M.E. and Benkovic, S.J. (1991) Kinetic mechanism of DNA polymerase I (Klenow fragment): identification of a second conformational change and evaluation of the internal equilibrium constant. *Biochemistry*, **30**, 4835-4843.

Domingo, E., Menendez-Arias, L., Quinones-Mateu, M.E., Holguin, A., Gutierrez-Rivas, M., Martinez, M.A., Quer, J., Novella, I.S. and Holland, J.J. (1997) Viral quasispecies and the problem of vaccine-escape and drug-resistant mutants. *Prog Drug Res*, **48**, 99-128.

Eigen, M., Winkler-Oswatitsch, R. and Dress, A. (1988) Statistical geometry in sequence space: a method of quantitative comparative sequence analysis. *Proc Natl Acad Sci U S A*, **85**, 5913-5917.

Fang, G. and Cech, T.R. (1993) Oxytricha telomere-binding protein: DNA-dependent dimerization of the alpha and beta subunits. *Proc Natl Acad Sci U S A*, **90**, 6056-6060.

Fernandez-Larsson, R. and Patterson, J.L. (1990) Ribavirin is an inhibitor of human immunodeficiency virus reverse transcriptase. *Mol Pharmacol*, **38**, 766-770.

Ferrer-Orta, C., Arias, A., Agudo, R., Perez-Luque, R., Escarmis, C., Domingo, E. and Verdaguer, N. (2006) The structure of a protein primer-polymerase complex in the initiation of genome replication. *Embo J*, **25**, 880-888.

Ferrer-Orta, C., Arias, A., Perez-Luque, R., Escarmis, C., Domingo, E. and Verdaguer, N. (2004) Structure of foot-and-mouth disease virus RNA-dependent RNA polymerase and its complex with a template-primer RNA. *J Biol Chem*, **279**, 47212-47221.

Flanegan, J.B. and Van Dyke, T.A. (1979) Isolation of a soluble and template-dependent poliovirus RNA polymerase that copies virion RNA in vitro. *J Virol*, **32**, 155-161.

Gamarnik, A.V. and Andino, R. (2000) Interactions of viral protein 3CD and poly(rC) binding protein with the 5' untranslated region of the poliovirus genome. *J Virol*, **74**, 2219-2226.

Gohara, D.W., Arnold, J.J. and Cameron, C.E. (2004) Poliovirus RNA-dependent RNA polymerase (3Dpol): kinetic, thermodynamic, and structural analysis of ribonucleotide selection. *Biochemistry*, **43**, 5149-5158.

Gohara, D.W., Crotty, S., Arnold, J.J., Yoder, J.D., Andino, R. and Cameron, C.E. (2000) Poliovirus RNA-dependent RNA polymerase (3Dpol): structural, biochemical, and biological analysis of conserved structural motifs A and B. *J Biol Chem*, **275**, 25523-25532.

Goodfellow, I., Chaudhry, Y., Richardson, A., Meredith, J., Almond, J.W., Barclay, W. and Evans, D.J. (2000) Identification of a cis-acting replication element within the poliovirus coding region. *J Virol*, **74**, 4590-4600.

Gottschling, D.E. and Zakian, V.A. (1986) Telomere proteins: specific recognition and protection of the natural termini of Oxytricha macronuclear DNA. *Cell*, **47**, 195-205.

Gray, J.T., Celander, D.W., Price, C.M. and Cech, T.R. (1991) Cloning and expression of genes for the Oxytricha telomere-binding protein: specific subunit interactions in the telomeric complex. *Cell*, **67**, 807-814.

Hansen, J.L., Long, A.M. and Schultz, S.C. (1997) Structure of the RNA-dependent RNA polymerase of poliovirus. *Structure*, **5**, 1109-1122.

Harris, K.S., Xiang, W., Alexander, L., Lane, W.S., Paul, A.V. and Wimmer, E. (1994) Interaction of poliovirus polypeptide 3CDpro with the 5' and 3' termini of the poliovirus genome. Identification of viral and cellular cofactors needed for efficient binding. *J Biol Chem*, **269**, 27004-27014.

Henderson, E.R., Moore, M. and Malcolm, B.A. (1990) Telomere G-strand structure and function analyzed by chemical protection, base analogue substitution, and utilization by telomerase in vitro. *Biochemistry*, **29**, 732-737.

Hogle, J.M. (2002) Poliovirus cell entry: common structural themes in viral cell entry pathways. *Annu Rev Microbiol*, **56**, 677-702.

Holland, J., Spindler, K., Horodyski, F., Grabau, E., Nichol, S. and VandePol, S. (1982) Rapid evolution of RNA genomes. *Science*, **215**, 1577-1585.

Hong, Z. and Cameron, C.E. (2002) Pleiotropic mechanisms of ribavirin antiviral activities. *Prog Drug Res*, **59**, 41-69.

Horvath, M.P. and Schultz, S.C. (2001) DNA G-quartets in a 1.86 A resolution structure of an Oxytricha nova telomeric protein-DNA complex. *J Mol Biol*, **310**, 367-377.

Horvath, M.P., Schweiker, V.L., Bevilacqua, J.M., Ruggles, J.A. and Schultz, S.C. (1998) Crystal structure of the Oxytricha nova telomere end binding protein complexed with single strand DNA. *Cell*, **95**, 963-974.

Hsieh, J.C., Zinnen, S. and Modrich, P. (1993) Kinetic mechanism of the DNA-dependent DNA polymerase activity of human immunodeficiency virus reverse transcriptase. *J Biol Chem*, **268**, 24607-24613.

Huang, Y., Eckstein, F., Padilla, R. and Sousa, R. (1997) Mechanism of ribose 2'-group discrimination by an RNA polymerase. *Biochemistry*, **36**, 8231-8242.

Jang, S.K., Krausslich, H.G., Nicklin, M.J., Duke, G.M., Palmenberg, A.C. and Wimmer, E. (1988) A segment of the 5' nontranslated region of encephalomyocarditis virus RNA directs internal entry of ribosomes during in vitro translation. *J Virol*, **62**, 2636-2643.

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

Klobutcher, L.A., Swanton, M.T., Donini, P. and Prescott, D.M. (1981) All gene-sized DNA molecules in four species of hypotrichs have the same terminal sequence and an unusual 3' terminus. *Proc Natl Acad Sci U S A*, **78**, 3015-3019.

Krausslich, H.G., Nicklin, M.J., Lee, C.K. and Wimmer, E. (1988) Polyprotein processing in picornavirus replication. *Biochimie*, **70**, 119-130.

Krilov, L.R. (2001) Respiratory Syncytial Virus: Update on Infection, Treatment, and Prevention. *Curr Infect Dis Rep*, **3**, 242-246.

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

Lawson, M.A. and Semler, B.L. (1990) Picornavirus protein processing--enzymes, substrates, and genetic regulation. *Curr Top Microbiol Immunol*, **161**, 49-87.

Lei, M., Podell, E.R., Baumann, P. and Cech, T.R. (2003) DNA self-recognition in the structure of Pot1 bound to telomeric single-stranded DNA. *Nature*, **426**, 198-203.

Lei, M., Podell, E.R. and Cech, T.R. (2004) Structure of human POT1 bound to telomeric single-stranded DNA provides a model for chromosome end-protection. *Nat Struct Mol Biol*, **11**, 1223-1229.

Leong, L., Cornell, C. and Semler, B., (2002) in *Molecular Biology of Picornaviruses*, eds. Semler,B. and Wimmer, E. (ASM Press), pp 187-198

Lesburg, C.A., Cable, M.B., Ferrari, E., Hong, Z., Mannarino, A.F. and Weber, P.C. (1999) Crystal structure of the RNA-dependent RNA polymerase from hepatitis C virus reveals a fully encircled active site. *Nat Struct Biol*, **6**, 937-943.

Lin, J.J. and Zakian, V.A. (1996) The Saccharomyces CDC13 protein is a single-strand TG1-3 telomeric DNA-binding protein in vitro that affects telomere behavior in vivo. *Proc Natl Acad Sci U S A*, **93**, 13760-13765.

Love, R.A., Maegley, K.A., Yu, X., Ferre, R.A., Lingardo, L.K., Diehl, W., Parge, H.E., Dragovich, P.S. and Fuhrman, S.A. (2004) The crystal structure of the RNA-dependent RNA polymerase from human rhinovirus: a dual function target for common cold antiviral therapy. *Structure*, **12**, 1533-1544.

Maag, D., Castro, C., Hong, Z. and Cameron, C.E. (2001) Hepatitis C virus RNA-dependent RNA polymerase (NS5B) as a mediator of the antiviral activity of ribavirin. *J Biol Chem*, **276**, 46094-46098.

Martin, C.T., Muller, D.K. and Coleman, J.E. (1988) Processivity in early stages of transcription by T7 RNA polymerase. *Biochemistry*, **27**, 3966-3974.

Maser, R.S. and DePinho, R.A. (2002) Connecting chromosomes, crisis, and cancer. *Science*, **297**, 565-569.

Mirmomeni, M.H., Hughes, P.J. and Stanway, G. (1997) An RNA tertiary structure in the 3' untranslated region of enteroviruses is necessary for efficient replication. *J Virol*, **71**, 2363-2370.

Mitton-Fry, R.M., Anderson, E.M., Hughes, T.R., Lundblad, V. and Wuttke, D.S. (2002) Conserved structure for single-stranded telomeric DNA recognition. *Science*, **296**, 145-147.

Muller, W.E., Maidhof, A., Taschner, H. and Zahn, R.K. (1977) Virazole (1-beta-D-ribofuranosyl-1,2,4-triazole-3-carboxamide; a cytostatic agent. *Biochem Pharmacol*, **26**, 1071-1075.

Murray, K.E. and Barton, D.J. (2003) Poliovirus CRE-dependent VPg uridylylation is required for positive-strand RNA synthesis but not for negative-strand RNA synthesis. *J Virol*, **77**, 4739-4750.

Murti, K.G. and Prescott, D.M. (2002) Topological organization of DNA molecules in the macronucleus of hypotrichous ciliated protozoa. *Chromosome Res*, **10**, 165-173.

Ng, K.K., Cherney, M.M., Vazquez, A.L., Machin, A., Alonso, J.M., Parra, F. and James, M.N. (2002) Crystal structures of active and inactive conformations of a caliciviral RNA-dependent RNA polymerase. *J Biol Chem*, **277**, 1381-1387.

Ng, K.K., Pendas-Franco, N., Rojo, J., Boga, J.A., Machin, A., Alonso, J.M. and Parra, F. (2004) Crystal structure of norwalk virus polymerase reveals the carboxyl terminus in the active site cleft. *J Biol Chem*, **279**, 16638-16645.

Novak, J.E. and Kirkegaard, K. (1991) Improved method for detecting poliovirus negative strands used to demonstrate specificity of positive-strand encapsidation and the ratio of positive to negative strands in infected cells. *J Virol*, **65**, 3384-3387.

Nugent, C.I., Hughes, T.R., Lue, N.F. and Lundblad, V. (1996) Cdc13p: a single-strand telomeric DNA-binding protein with a dual role in yeast telomere maintenance. *Science*, **274**, 249-252.

O'Farrell, D., Trowbridge, R., Rowlands, D. and Jager, J. (2003) Substrate complexes of hepatitis C virus RNA polymerase (HC-J4): structural evidence for nucleotide import and de-novo initiation. *J Mol Biol*, **326**, 1025-1035.

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

Paeschke, K., Simonsson, T., Postberg, J., Rhodes, D. and Lipps, H.J. (2005) Telomere end-binding proteins control the formation of G-quadruplex DNA structures in vivo. *Nat Struct Mol Biol*, **12**, 847-854.

Palmenberg, A. and Sgro, JV. (2002) in *Molecular Biology of Picornaviruses*, eds. Semler,B. and Wimmer, E. (ASM Press), pp 149-155

Parsley, T.B., Cornell, C.T. and Semler, B.L. (1999) Modulation of the RNA binding and protein processing activities of poliovirus polypeptide 3CD by the viral RNA polymerase domain. *J Biol Chem*, **274**, 12867-12876.

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.

Pathak, H.B., Ghosh, S.K., Roberts, A.W., Sharma, S.D., Yoder, J.D., Arnold, J.J., Gohara, D.W., Barton, D.J., Paul, A.V. and Cameron, C.E. (2002) Structure-function relationships of the RNA-dependent RNA polymerase from poliovirus (3Dpol). A surface of the primary oligomerization domain functions in capsid precursor processing and VPg uridylylation. *J Biol Chem*, **277**, 31551-31562.

Paul, A. (2002) in *Molecular Biology of Picornaviruses*, eds. Semler,B. and Wimmer, E. (ASM Press), pp 227-246.

Paul, A.V., Rieder, E., Kim, D.W., van Boom, J.H. and Wimmer, E. (2000) Identification of an RNA hairpin in poliovirus RNA that serves as the primary template in the in vitro uridylylation of VPg. *J Virol*, **74**, 10359-10370.

Peersen, O.B., Ruggles, J.A. and Schultz, S.C. (2002) Dimeric structure of the Oxytricha nova telomere end-binding protein alpha-subunit bound to ssDNA. *Nat Struct Biol*, **9**, 182-187.

Pelletier, J. and Sonenberg, N. (1988) Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA. *Nature*, **334**, 320-325.

Pfeiffer, J.K. and Kirkegaard, K. (2003) A single mutation in poliovirus RNA-dependent RNA polymerase confers resistance to mutagenic nucleotide analogs via increased fidelity. *Proc Natl Acad Sci U S A*, **100**, 7289-7294.

Pfeiffer, J.K. and Kirkegaard, K. (2005) Increased Fidelity Reduces Poliovirus Fitness and Virulence under Selective Pressure in Mice. *PLoS Pathog*, **1**, e11.

Poch, O., Sauvaget, I., Delarue, M. and Tordo, N. (1989) Identification of four conserved motifs among the RNA-dependent polymerase encoding elements. *Embo J*, **8**, 3867-3874.

Price, C.M. and Cech, T.R. (1987) Telomeric DNA-protein interactions of Oxytricha macronuclear DNA. *Genes Dev*, **1**, 783-793.

Price, C.M. and Cech, T.R. (1989) Properties of the telomeric DNA-binding protein from Oxytricha nova. *Biochemistry*, **28**, 769-774.

Rieder, E., Paul, A.V., Kim, D.W., van Boom, J.H. and Wimmer, E. (2000) Genetic and biochemical studies of poliovirus cis-acting replication element cre in relation to VPg uridylylation. *J Virol*, **74**, 10371-10380.

Rieder, E. and Wimmer E. (2002) in *Molecular Biology of Picornaviruses*, eds. Semler,B. and Wimmer, E. (ASM Press), pp 61-70.

Salgado, P.S., Makeyev, E.V., Butcher, S.J., Bamford, D.H., Stuart, D.I. and Grimes, J.M. (2004) The structural basis for RNA specificity and Ca2+ inhibition of an RNA-dependent RNA polymerase. *Structure*, **12**, 307-316.

Schaffitzel, C., Berger, I., Postberg, J., Hanes, J., Lipps, H.J. and Pluckthun, A. (2001) In vitro generated antibodies specific for telomeric guanine-quadruplex DNA react with Stylonychia lemnae macronuclei. *Proc Natl Acad Sci U S A*, **98**, 8572-8577.

Severson, W.E., Schmaljohn, C.S., Javadian, A. and Jonsson, C.B. (2003) Ribavirin causes error catastrophe during Hantaan virus replication. *J Virol*, **77**, 481-488.

Sousa, R., Chung, Y.J., Rose, J.P. and Wang, B.C. (1993) Crystal structure of bacteriophage T7 RNA polymerase at 3.3 A resolution. *Nature*, **364**, 593-599.

Steinhauer, D.A., de la Torre, J.C., Meier, E. and Holland, J.J. (1989) Extreme heterogeneity in populations of vesicular stomatitis virus. *J Virol*, **63**, 2072-2080.

Steinhauer, D.A., Domingo, E. and Holland, J.J. (1992) Lack of evidence for proofreading mechanisms associated with an RNA virus polymerase. *Gene*, **122**, 281-288.

Swanton, M.T., Heumann, J.M. and Prescott, D.M. (1980) Gene-sized DNA molecules of the macronuclei in three species of hypotrichs: size distributions and absence of nicks. DNA of ciliated protozoa. VIII. *Chromosoma*, **77**, 217-227.

Tahirov, T.H., Temiakov, D., Anikin, M., Patlan, V., McAllister, W.T., Vassylyev, D.G. and Yokoyama, S. (2002) Structure of a T7 RNA polymerase elongation complex at 2.9 A resolution. *Nature*, **420**, 43-50.

Tao, Y., Farsetta, D.L., Nibert, M.L. and Harrison, S.C. (2002) RNA synthesis in a cage--structural studies of reovirus polymerase lambda3. *Cell*, **111**, 733-745.

Theobald, D.L. and Schultz, S.C. (2003) Nucleotide shuffling and ssDNA recognition in Oxytricha nova telomere end-binding protein complexes. *Embo J*, **22**, 4314-4324.

Theobald, D. and Wuttke, D. (2006) Superpositioning by maximum likelihood analysis of the matrix Gaussian Procurstes problem. *PNAS*, article submitted.

Thompson, A.A. and Peersen, O.B. (2004) Structural basis for proteolysis-dependent activation of the poliovirus RNA-dependent RNA polymerase. *Embo J*, **23**, 3462-3471.

Thompson, A.A. (2006). X-ray crystal structure, proteolytic dependent activation, and functional analysis of the poliovirus RNA-Dependent RNA Polymerase (Doctoral dissertation, Colorado State University, 2006).

van Dijk, A.A., Makeyev, E.V. and Bamford, D.H. (2004) Initiation of viral RNA-dependent RNA polymerization. *J Gen Virol*, **85**, 1077-1093.

van Holde, KE. and Weishcet, WO. (1975). Boundary analysis of sedimentation velocity experiments with monodisperse and paucidisperse solutes. *Biopolymers* **17**, 1387-1403.

Vignuzzi, M., Stone, J.K., Arnold, J.J., Cameron, C.E. and Andino, R. (2006) Quasispecies diversity determines pathogenesis through cooperative interactions in a viral population. *Nature*, **439**, 344-348.

Vo, N.V., Young, K.C. and Lai, M.M. (2003) Mutagenic and inhibitory effects of ribavirin on hepatitis C virus RNA polymerase. *Biochemistry*, **42**, 10462-10471.

World Health Oraganization *Cessation of routine OPV use after global polio eradication.* Framework for National policy makers in OPV-using contries. 2005. http://www.polioeradication.org/content/publications/OPVCessationFrameworkEnglish.pdf

Wray, S.K., Gilbert, B.E. and Knight, V. (1985) Effect of ribavirin triphosphate on primer generation and elongation during influenza virus transcription in vitro. *Antiviral Res*, **5**, 39-48.

Wright, W.E. and Shay, J.W. (1992) The two-stage mechanism controlling cellular senescence and immortalization. *Exp Gerontol*, **27**, 383-389.

Yin, Y.W. and Steitz, T.A. (2002) Structural basis for the transition from initiation to elongation transcription in T7 RNA polymerase. *Science*, **298**, 1387-1395.

Yin, Y.W. and Steitz, T.A. (2004) The structural mechanism of translocation and helicase activity in T7 RNA polymerase. *Cell*, **116**, 393-404.

Ypma-Wong, M.F., Dewalt, P.G., Johnson, V.H., Lamb, J.G. and Semler, B.L. (1988) Protein 3CD is the major poliovirus proteinase responsible for cleavage of the P1 capsid precursor. *Virology*, **166**, 265-270.