DISSERTATION

X-RAY CRYSTAL STRUCTURE, PROTEOLYTIC DEPENDENT ACTIVATION, AND FUNCTIONAL ANALYSIS OF THE POLIOVIRUS RNA-DEPENDENT RNA POLYMERASE

Submitted by
Aaron A. Thompson
Graduate Program in Cell and Molecular Biology

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

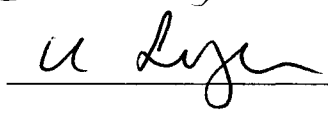
COLORADO STATE UNIVERSIY

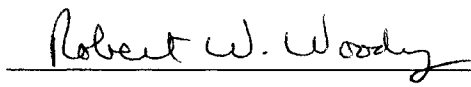
March 31, 2006

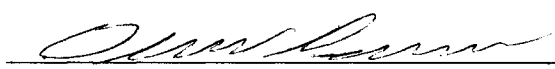
WE HEREBY RECOMMEND THAT THE DISSERTATION PREPARED UNDER OUR SUPERVISION BY AARON A. THOMPSON ENTITLED "X-RAY CRYSTAL STRUCTURE, PROTEOLYTIC DEPENDENT ACTIVATION, AND FUNCTIONAL ANALYSIS OF THE POLIOVIRUS RNA-DEPENDENT RNA POLYMERASE" BE ACCEPTED AS FULFILLING IN PART REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

Committee on Graduate Work


(Oren Anderson)


(Karolin Luger)


(Robert Woody)


(Olve Peersen)

Advisor


(Norman Curthoys)

Department Head/Director

ABSTRACT OF DISSERTATION

X-ray Crystal Structure, Proteolytic Dependent Activation and Functional Analysis of the Poliovirus RNA-dependent RNA Polymerase

Plus-strand RNA viruses are responsible for a large number of human and animal diseases including poliomyelitis, yellow fever, hepatitis C, foot-and-mouth disease, and dengue fever. Upon infection, the single plus-strand RNA molecule contained within the viral particle is translated by cellular ribosomes of the host into a single large polyprotein that is then cleaved by viral proteases into 11 distinct proteins and several functional precursors. The last enzyme in this polyprotein is a RNA-dependent RNA polymerase that is required for replication of the viral genome. Using a novel method of intentionally disrupting a known and persistent crystal packing interaction, we have solved the crystal structure of the complete poliovirus polymerase and complexes containing all four ribonucleotides and Mn^{2+}. The structures show the conformation of the previously unresolved "fingers" domain and reveal that the very N-terminus of the protein is buried in a pocket on the surface of the protein. This proteolytic dependent burial of the N-terminus forms important hydrogen bonds that are required for the proper positioning of a key residue in the active site, Asp238, which in turn forms a required hydrogen bond with the 2' OH group of the incoming NTP. Additionally, we provide data showing that interdomain

interactions responsible for the "enclosed" active site of this enzyme are important for thermal stability. These data suggest that poliovirus $3D^{pol}$ and other viral RNA-dependent RNA polymerases are unusually rigid, and we thus propose an exposed portion of the protein (pinky finger; residues 96-149) may play a significant role in the nucleotide polymerization activity of the poliovirus polymerase.

Aaron A. Thompson
Graduate Program in Cell
and Molecular Biology
Colorado State University
Fort Collins, CO 80523
Spring 2006

ACKNOWLEDGEMENTS

Foremost I thank Olve Peersen for his support and enthusiasm throughout my graduate career. As my mentor, he has shown me how to critically approach science while giving me the freedom to become an independent thinker. The knowledge I have gained while working in the Peersen laboratory will certainly be a positive influence on the rest of my scientific career. I truly hope that he remains a friend and becomes a future collaborator.

I would also like to thank all of the past and present members of the Peersen laboratory. Vandy Johnson introduced me into the Peersen lab by teaching me how to clone and purify proteins. Aaron Sholders has been a wonderful bench-mate and has been the source for countless stimulating discussions. Numerous other past and present members have augmented the scientific and social atmosphere of the lab including: Michelle Fodor, Kirsten Scoggin, Santano Mestes, Grace Campagnola, and Becky Albertini.

Most of all, I'd like to thank my entire family for their support and encouragement. My mother and father always stressed the importance of education and never let me loose sight of my goals.

DEDICATION

I dedicate this dissertation to my family, who have supported and encouraged me as long as I can remember. I will not forget drives through Over the Rhine.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER I

INTRODUCTION

POLIOVIRUS OVERVIEW

Poliovirus is perhaps the most recognized and studied member of the *Picornaviridae* family. Before the Salk and Sabin vaccines were developed in the 1950's, polio occupied a similar place in the national consciousness as that occupied by HIV today. This family of small RNA viruses includes many other notorious pathogens, such as the heart disease-causing coxsackie virus, hepatitis A virus, foot and mouth disease viruses, and the rhinoviruses that are the major cause of the common cold. The study of these picornaviruses has led to many significant advances in the field of virology. In 1898, foot and mouth disease virus was the first animal virus to be discovered (Loeffler and Frosch 1964), with poliovirus being isolated ten years later (Landsteiner and Popper, 1908). Development of the first infectious DNA clone of an animal RNA virus (Racaniello and Baltimore, 1981a) and the plaque assay (Dulbecco 1952), which is used for the quantification of viral infectivity, were pioneered with poliovirus. In the early '60s, the first RNA-dependent RNA polymerase (RdRp) was identified in mengovirus, a picornavirus that infects pancreatic beta cells of mice (Baltimore and Franklin, 1963). More recently, the first three-dimensional X-ray

crystallographic structures of viruses were determined from the picornaviruses poliovirus and rhinovirus (Hogle *et al.*, 1985; Rossmann *et al.*, 1985), followed by the first crystal structure of an RdRp, which was from poliovirus (Hansen *et al.*, 1997). Because the initial structure of the poliovirus polymerase was incomplete, my work has focused on obtaining the complete three-dimensional structure of this enzyme which in-turn has led to novel findings.

Poliovirus transmission is typically through the fecal-oral route, leading to viral multiplication in the upper gastrointestinal tract. Viral particles then enter the blood stream where they may seed infections at various sites within the host, including the central nervous system (CNS). Paralytic poliomyelitis, resulting from viral invasion of motor neurons, occurs rarely in 0.1 - 1 % of infections. Viral entry into the cell is achieved by binding the cellular receptor CD155, a member of the immunoglobulin (Ig) superfamily (Mendelsohn *et al.*, 1989; Koike *et al.*, 1990). The CD155 receptor has four splice variants, two of which (CD155β and CD155γ) are secreted and have unknown function (Koike *et al.*, 1990). The other two variants (CD155 α and CD155δ) are membrane-bound and have recently been found to function as cell-cell and cell-matrix adhesion molecules (reviewed in Mueller, 2005). By poorly understood mechanisms, poliovirus docking with its receptor causes significant structural changes of the viral capsid, resulting in the formation of a membrane pore through which the viral genome can enter the cytoplasm of its host (Hogle, 2002; Tuthill et al., 2006).

Following insertion of the 7.5-kilobase single-stranded positive-sense RNA molecule into the cell (Figure 1.1), ribosomes translate a single large open reading frame in a cap-independent manner (Jang *et al.*, 1988; Pelletier and Sonenberg, 1988). This produces a single large 247 KDa polyprotein that is then cleaved by cis-acting viral proteases into various structural and non-structural viral gene products, enabling propagation of the virus (Summers *et al.*, 1968; Wimmer et al., 1993). The last protein in the viral polyprotein is $3D^{pol}$, an RNA-dependent RNA polymerase (RdRP) responsible for both replicating the infecting positive-sense genome into minus-sense complements, and then using these as templates for the synthesis of positive-sense genomes that are packaged into new virions. RNA replication and viron assembly are coupled and occur at large viral replication center structures that are found on the surfaces of small vesicles. These vesicles are derived from disrupted golgi and endoplasmic reticulum by the action of viral proteins 2B/2BC, 2C, and 3A/3AB (Bienz *et al.*, 1983; Schlegel *et al.*, 1996; Teterina *et al.*, 1997; Jackson et al., 2005). All aspects of the viral life cycle, including viral replication, take place in the cytoplasm.

GENOME ORGANIZATION AND REPLICATION

Nucleotide sequence analysis reveals common genome organization among all picornaviruses. Flanking the open reading frame of these viruses are the 5' and 3' non-translated regions (NTR) (Kitamura et al., 1981; Racaniello and Baltimore, 1981b; Dorner et al., 1982), which are important for controlling replication and translation by mechanisms that are not fully understood. The

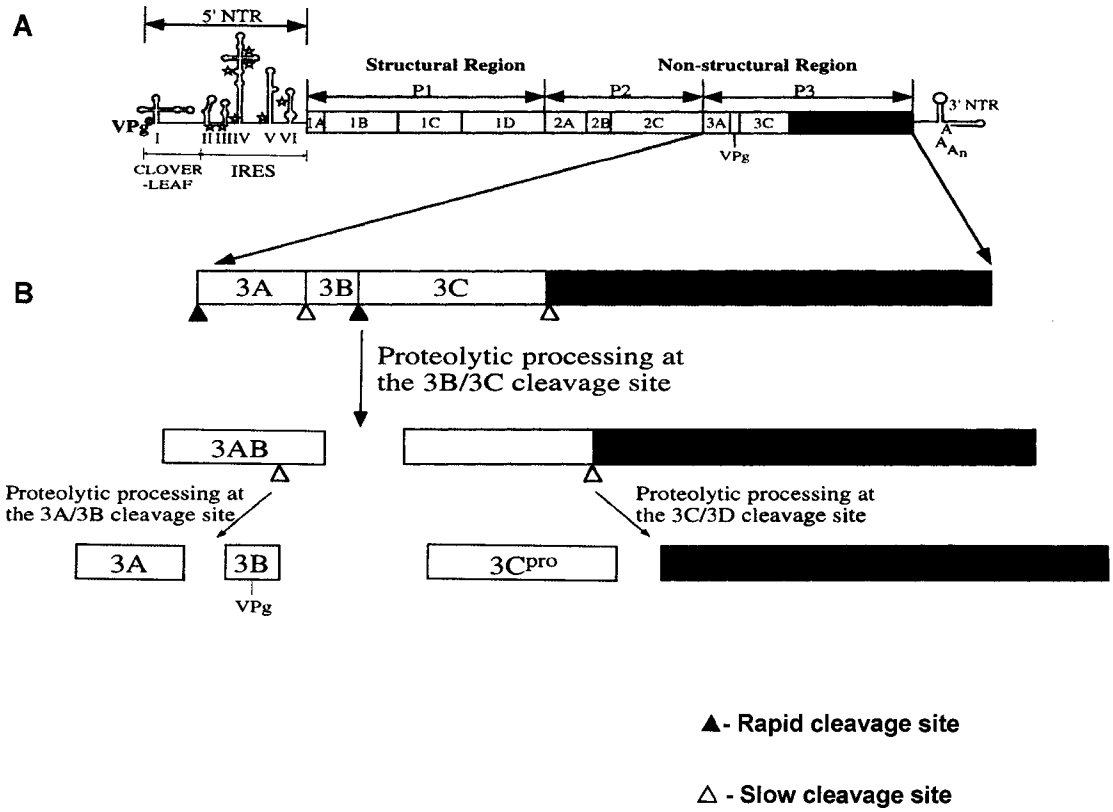


Figure 1.1 Genome Overview

(**A**) Structural organization of the poliovirus single-stranded positive-sense RNA genome. The single open reading frame is flanked by the 5' and 3' non-coding regions (NTR). The open reading frame encodes for both structural proteins (P1 region) which compose the viral capsid and non-structural proteins (P2 & P3 regions) that enable the virus to replicate. (**B**) Enlargement of the P3 nonstructural precursor and its processing products. Closed and open triangles indicate rapidly and slowly processed Q/G cleavage sites, respectively. Figure adapted from Xiang *et al.* J Virol. 1998 Aug;72(8):6732-41.

poliovirus 5' NCR is typically 742 nucleotides long and is predicted to have significant secondary structure consisting of six large stem-loop structures. The first of these stem loops forms a cloverleaf structure (Rivera *et al.*, 1988; Andino *et al.*, 1990), which interacts with both viral and host proteins to direct RNA synthesis, and the remaining five stem loops form the internal ribosome entry site (IRES), which enables cap-independent translation (Jang *et al.*, 1988, 1989; Molla *et al.*, 1992). The IRES has been shown to be responsible for initiation of translation by binding the 40S ribosome subunit. This mechanism of recruitment is necessary since the 5' end of the viral RNA lacks the traditional 7-methyl guanosine cap found in cellular mRNAs. This method is advantageous and necessary because the $2A^{pro}$ viral protease has been shown to cleave p220, a component of the translation initiation factors eIF-4F, thereby halting all translation from cellular capped mRNAs (Krausslich et al., 1987; Sonenberg, 1987; Bonneau and Sonenberg, 1987). Additionally, the $3C^{pro}$ protease assists in hijacking the cell by inactivating the transcription factor TFIIIC and TATA box binding protein, resulting in inhibition of host cell transcription (Clark *et al.*, 1991; Yalamanchili *et al.*, 1996).

The 5' cloverleaf structure is composed of three small predicted stem loops from the first ~90 nucleotides of the 5' NCR, and is known to bind the cellular protein poly(rC) binding protein (PCBP) and the viral protease and RNA-binding protein $3CD^{pro}$. This ribonucleoprotein (RNP) complex is thought to stabilize the positive-strand viral RNA against degradation and play a role in the initiation of negative-strand RNA synthesis following translation through interaction with the

3' NCR that leads to circularization of the genome. This is supported by data showing that an intact 5' cloverleaf structure is required for negative-strand synthesis, which is initiated at the 3' end of the genome (Lyons *et al.*, 2001).

The 3' NCR is small in comparison to the 5' NCR, and is composed of a 70-nt segment forming two predicted stem loops that may interact to form a tertiary pseudoknot structure (Pilipenko *et al.*, 1992; Jacobson *et al.*, 1993), and a poly(A) tail with an average length of 60 nucleotides (Yogo and Wimmer, 1972). It has been shown that reducing the length of the poly(A) tail results in deleterious consequences on viral infectivity (Spector and Baltimore, 1974; Sarnow, 1989). Comparison of poliovirus RNA transcripts with 12- and 80-nt poly(A) tails in translation-replication reactions revealed severely inhibited negative-strand RNA synthesis when the tail length is reduced, but no effect on translation (Barton *et al.*, 1996; Silvestri et al., 2006). Because the cellular poly(A)-binding protein (PABP) binds to both the poly(A) tail and $3CD^{pro}$ (bound at the cloverleaf structure), reducing the poly(A) tail length likely precludes potential interactions of the 3' NCR with the 5' NCR mediated by PABP and $3CD^{pro}$. It has been proposed that this circularization of the viral genome is responsible for the switch from translation to negative-strand RNA replication by preventing access of the host's ribosomes to the IRES (Herold and Andino, 2001; Barton *et al.*, 2001; Gamarnik and Andino, 1998).

The mechanisms of positive-strand synthesis are less well understood even though there are typically 20-50 (+) strands made in the cell for every (-) strand (Novak and Kirkegaard, 1991; Murray and Barton, 2003). The evolution of this

stoichiometric asymmetry is not surprising since the minus-sense RNAs are replication intermediates while larger amounts of plus-strand RNAs are required for packaging into the viral capsids. Work from the Barton laboratory suggests that one important difference between positive- and negative-strand RNA synthesis lies in the mechanism of initiation.

Both negative- and positive-strand synthesis is initiated and primed from Tyr3 of the viral peptide VPg (viral peptide genome-linked), a 22-amino acid peptide encoded by the 3B gene. However, in contrast with negative strand initiation, positive-strand initiation will not proceed until the VPg peptide has two uridines covalently added to Tyr3 to generate the molecule VPg-pU-pU$_{OH}$ (Murray and Barton, 2003). This reaction can be catalyzed *in vitro* by 3D^{pol} in the presence of a poly(A) template (Paul *et al.*, 1998). Because of this observation, it was believed that the poly(A) tail of the positive-sense RNA molecule served as the template for the uridylylation reaction (Paul *et al.*, 1998). However, investigators working with human rhinovirus 14 found that a predicted internal RNA hairpin structure termed cis-replicating element (*cre*) instead served as the template (McKnight and Lemon, 1996, 1998). A *cre* was then discovered in the 2C coding region of the poliovirus genome with a conserved 5'-AAAC-3' sequence templating the uridylylation reaction more efficiently than a poly(A) oligomer (Paul *et al.*, 2000a,b; Rieder *et al.*, 2000). Uridylylation studies using preinitiation RNA-replication complexes have shown that mutations of the *cre* abolish the synthesis of VPg-pU-pU$_{OH}$ despite the presence of a poly(A) tail on the constructs tested (Murray and Barton, 2003). It has also been shown that the

5' (A) of the conserved 5'-AAAC-3' sequence serves as the template for the addition of both uridines by a proposed slide-back mechanism, and that this reaction is dramatically stimulated by $3CD^{pro}$, thus providing a potential mechanism for specificity *in vivo* (Rieder *et al.*, 2000; Paul *et al.*, 2000b)

PROTEOLYTIC PROTEIN PROCESSING

Because the poliovirus virion contains no proteins required for viral replication, its single open reading frame must first be translated for successful infection. The product of translation is a 2209-amino acid polyprotein that is organized into three regions designated P1, P2, and P3 (see figure 1.1). The P1 region contains structural proteins that compose the viral capsid, and the P2 and P3 region contain non-structural proteins required for propagation of the virus. Cleavage of the various proteins occurs post- and co-translationally by the cis-acting viral proteases $2A^{pro}$, $3C^{pro}$ and $3CD^{pro}$. Because of differences in the catalytic efficiency of these proteases at the various cleavage sites, there are disproportionate amounts of the mature viral proteins and their precursors. This temporal expression, which yields approximately 27 intermediate and mature cleavage products (Harris *et al.*, 1990; Wimmer *et al.*, 1993), is important because many of the precursors have functions which differ from that of the fully cleaved protein. This post-translational control facilitates progression of the viral life-cycle of these viruses with genomes of limited coding capacity.

The best example of this may be found with proteins derived from the non-structural P3 region of the poliovirus genome. The P3 precursor is first

generated by fast cleavage at the amino terminus of the 3A protein and is rapidly followed by another cleavage at the 3AB - $3CD^{pro}$ junction. This yields two stable and multifunctional precursor proteins (3AB and $3CD^{pro}$) that are important for the viral life cycle. Proteolytic processing of these precursors is less efficient and therefore occurs at a much slower rate, yielding less abundance of the cleavage end products, 3A, 3B, $3C^{pro}$ and $3D^{pol}$. As mentioned above, the 3B (VPg) protein serves as the primer for initiation, but its fusion with the hydrophobic 3A is likely the functionally relevant form of this peptide *in vivo*. The 3A protein is predicted to contain a peripheral-membrane binding domain that tethers it and it's precursor (3AB) to intra-cellular membranes (Semler *et al.*, 1982; Datta and Dasgupta, 1994; Towner *et al.*, 1996). 3AB has been shown to associate with $3D^{pol}$ in yeast two-hybrid assays (Xiang *et al.*, 1998), and it is thought this association may localize $3D^{pol}$ and/or its precursors to the membrane-bound replication centers (Hope *et al.*, 1997). Detergent-solubilized 3AB has also been shown to stimulate $3D^{pol}$ polymerase activity and $3CD^{pro}$ protease activity (Lama *et al.*, 1994; Paul *et al.*, 1994; Plotch *et al.*, 1995).

The viral protease $3C^{pro}$ is homologous to the serine proteinases such as chymotrypsin both in terms of structure and amino acid sequence. One important difference is found in the catalytic triad where the nucleophilic Ser195 of chymotrypsin is replaced with Cys147 in $3C^{pro}$. Because the 3BC junction site likely lies in the substrate binding site of $3C^{pro}$, it has been proposed that cleavage of this junction results in refolding of the N-terminus, creating an open substrate-accessible active site (Semler and Wimmer, 2002). This is reminiscent of the

classic chymotrypsin mode of activation where the inactive proenzyme chymotrypsinogen undergoes cleavage at Arg15-Ile16, thereby creating a new N-terminus that restructures the enzyme forming an active catalytic site. While $3C^{pro}$ is competent at cleaving Gln-Gly bonds, its precursor $3CD^{pro}$ is able to accomplish this on capsid proteins up to 1000-fold more efficiently when the $3D^{pol}$ domain is present (Ypma-Wong *et al.*, 1988; Parsley *et al.*, 1999). However, proper proteolytic processing of the $3D^{pol}$ N-terminus is required to activate the polymerase as $3CD^{pro}$ has absolutely no RNA polymerase activity despite the fact that it contains the entire polymerase domain (Flanegan and Van Dyke, 1979; Harris *et al.*, 1992). Interestingly, the addition of 11 residues from the C-terminal end of $3C^{pro}$ to the N-terminus of $3D^{pol}$ results in a total loss of polymerase activity (Rothstein *et al.*, 1988), as does the deletion of the first 6 residues (Hobson *et al.*, 2001) or of only Trp5 (Plotch *et al.*, 1989). $3CD^{pro}$ also functions as a key RNA-binding protein that orchestrates translation and replication of the viral genome by binding the 5' and 3' NTPs (Parsley *et al.*, 1999).

The RNA dependent RNA polymerase ($3D^{pol}$)

The C-terminal end of the viral polyprotein encodes a 53 KDa RNA-dependent RNA polymerase ($3D^{pol}$), which is known to catalyze two types of reactions, uridylylation of VPg, and template-primer dependent RNA polymerization. As discussed above, uridylylation of the viral 3B protein (VPg) ultimately results in the covalent attachment of positive-sense viral RNA to this

primer. $3D^{pol}$ is also the principal enzyme responsible for both replicating the infecting positive-sense genomes into minus-sense complements and then using these as templates for the synthesis of positive-sense genomes that are packaged into new virions.

Like other RNA viruses, the fidelity of replication is low in an effort to increase the evolutionary fitness. The error rate for $3D^{pol}$ is approximately 10^{-3} to 10^{-5} substitutions per nucleotide, a value approximately a millionfold higher that the rate estimated for cellular DNA replication (Drake, 1969; Drake, 1991). These error rates suggest that every viral RNA genome contains at least one mutated nucleotide (Drake and Holland, 1999), leading to the idea that an infected cell contains an assortment of related but slightly different viruses in what is called a quasispecies population. While it is possible that the architecture of RNA polymerases is more permissive of nucleotide misincorporation, the high error rate is undoubtedly influenced by the absence of proofreading-repair activities that are found with many other types of polymerases. With no known viral or cellular processivity factors, the promiscuity of $3D^{pol}$ is further demonstrated by its ability to switch templates (Arnold and Cameron, 1999).

A partial crystal structure of poliovirus $3D^{pol}$ (Hansen *et al.*, 1997) showed that the RNA-dependent RNA polymerase shared the "right hand" analogy used to describe the structure of the DNA polymerase I Klenow fragment (Ollis et al., 1985) with fingers, palm, and thumb domains (Figure 1.2). This original $3D^{pol}$ structure and three other crystal forms (Hobson, 2000) crystallized in similar space groups whose lattices were dominated by a persistent crystal contact called

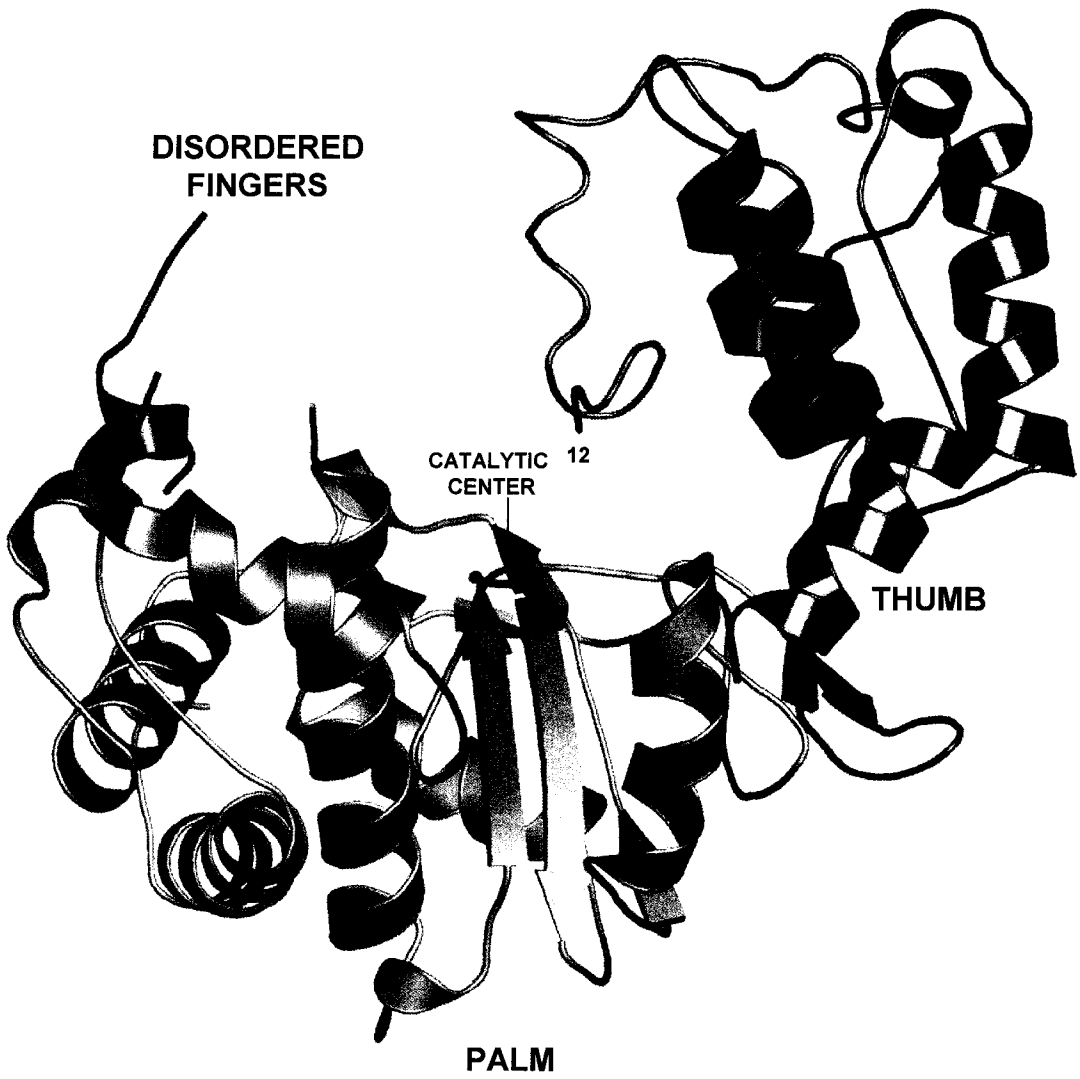


Figure 1.2 Partial wild-type Structure
Partial $3D^{pol}$ structure (PDB code 1RDR) showing the palm in gray, thumb in blue, catalytic center in magenta, base of the fingers domain in red, and an N-terminal portion from the fingers domain in green. Most of the fingers domain (145 residues) is disordered in this structure.

Interface I (see figure 1.3). Interface I results in a head-to-tail oligomerization of the polymerase that potentially reflects biologically relevant interactions in the membrane-bound replication complexes (Hobson *et al.*, 2001; Lyle *et al.*, 2002a). Unfortunately, the electron density for the fingers domain was missing from all these structures, precluding a structural explanation of many aspects of polymerase function such as the critical importance of having a proper N-terminus in order for the polymerase to function. The original structure does show a small segment of the fingers domain where residues 25-37 interact with the top of the thumb and residues 12-24 descend toward the active site in the palm domain. It was originally proposed that this stretch of residues (12-37) was contributed in *trans* from another molecule through an interaction called Interface II (Hansen *et al.*, 1997). Although the first eleven residues were not observed, the structure suggested that the N-terminus could be an integral component of the catalytic site. Activation of the polymerase would therefore require the intermolecular N-terminal contribution of one polymerase to the complete the active site of another, an attractive hypothesis for supporting replication in densely packed replication centers.

The palm domain contains the catalytic center and is highly conserved across all categories of polymerases. This domain consists of a four-stranded antiparallel β-sheet packed against two α-helices (Figure 1.4), a topology very similar to RNA recognition motifs (RRM) found in various types of RNA-binding proteins (Hansen *et al.*, 1997). The active site is centered around a type II' β-turn from the β-sheet consisting of the residues Tyr-Gly-Asp-Asp which is very

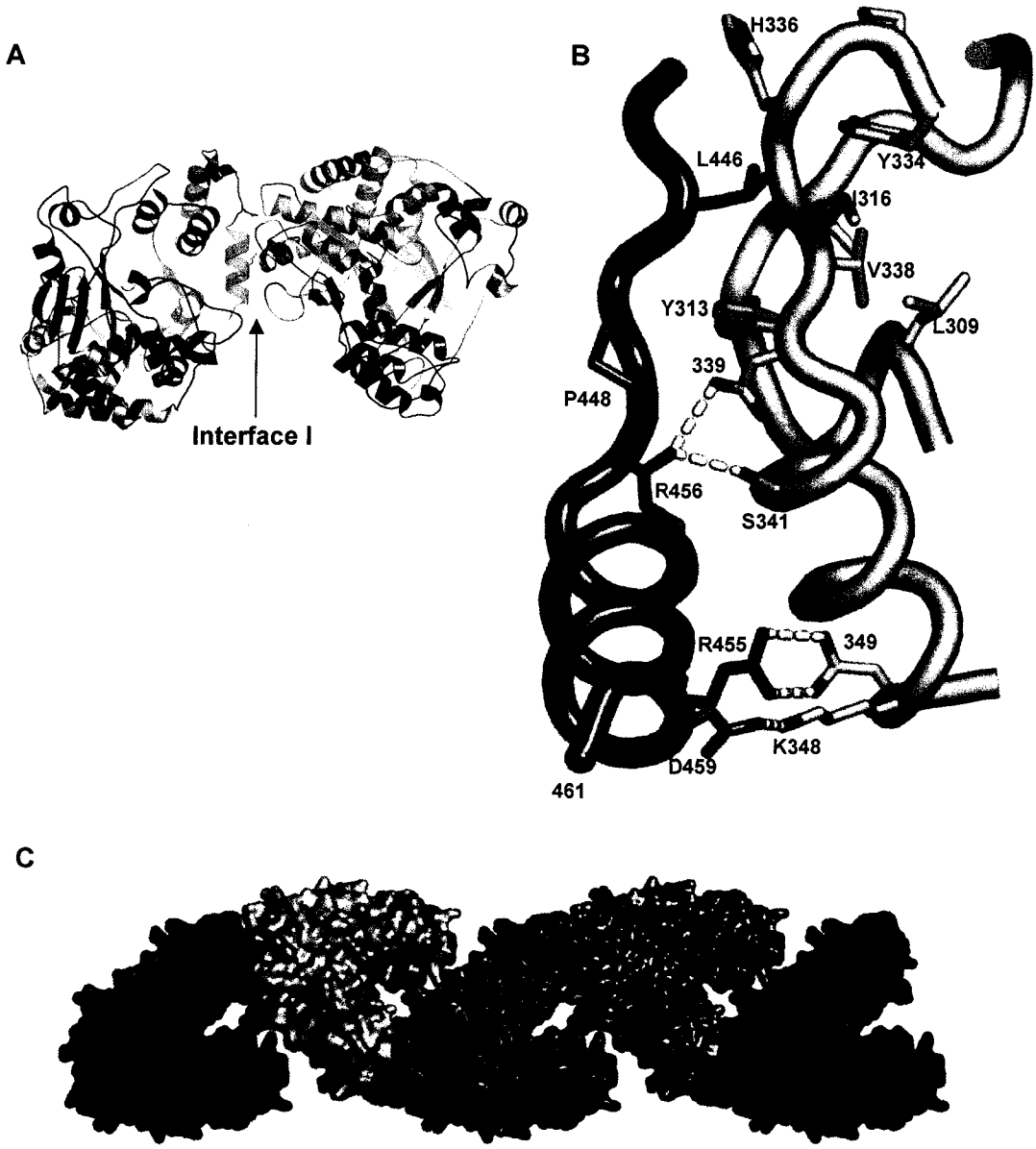


Figure 1.3 interface I
(**A**) Two complete 3D^{pol} molecules interacting at the putative Interface I which involves the thumb (blue) of one polymerase and the palm (grey) of another. (**B**) Detailed view of the Interface I interaction shown in panel (A). Residues colored in magenta (L446, Arg455) were mutated to aspartates to disrupt this interaction. (**C**) Wild-type 3D^{pol} polymerase fiber consisting of five molecules interacting via Interface I. Each polymerase-polymerase interaction buries ~2000 Å^2 of solvent exposed surface area. The complete 3D^{pol} structure discussed in Ch. II is used to represent the Interface I fiber (C) and dimer (A).

conserved among RdRps (colored magenta in all figures – Figure 1.2). Tyrosine 326 can be mutated to a phenylalanine or methionine without loss of *in vitro* activity, although only the Y326F mutation yields viable virus in poliovirus cDNA transfection trials (Jablonski and Morrow, 1993). While the flexibility of glycine 327 at the tight β-turn likely accounts for the conservation of this amino acid, residual activity has been observed when this residue is mutated to alanine or serine (Jablonski *et al.*, 1991). Aspartates 328 and 329 occupy positions 3 and 4 of the β-turn and are critical for chelating two divalent metals required for catalysis. Aspartate 328 is structurally conserved in all polymerases studied (Gohara *et al.*, 2000) and it is not surprising that mutation of this residue completely inactivates the enzyme (Jablonski *et al.*, 1995). However, mutation of aspartic acid 329 to an asparagine displayed *in vitro* activity when Mn^{2+} or Fe^{2+} were supplied while other divalent cations, including Mg^{2+}, did not support activity (Jablonski *et al.*, 1995). Located on an adjacent β-strand on the "back" side of the palm is Asp233, another highly conserved constituent of the metal binding site. In the partial polio polymerase structure, a strong electron density peak was found cradled between these aspartates, which was subsequently modeled as a Ca^{2+} ion since it was the only divalent metal cation present in the crystallization conditions (Figure 1.4; Hansen *et al.*, 1997).

Only recently has the two-metal mechanism of catalysis been firmly established (Brautigam and Steitz, 1998; Steitz, 1998; Sosunov *et al.*, 2003, 2005). This model involves the binding of two divalent cations termed metals A and B, which have distinct roles. Metal A is proximal to the 3' end of the primer

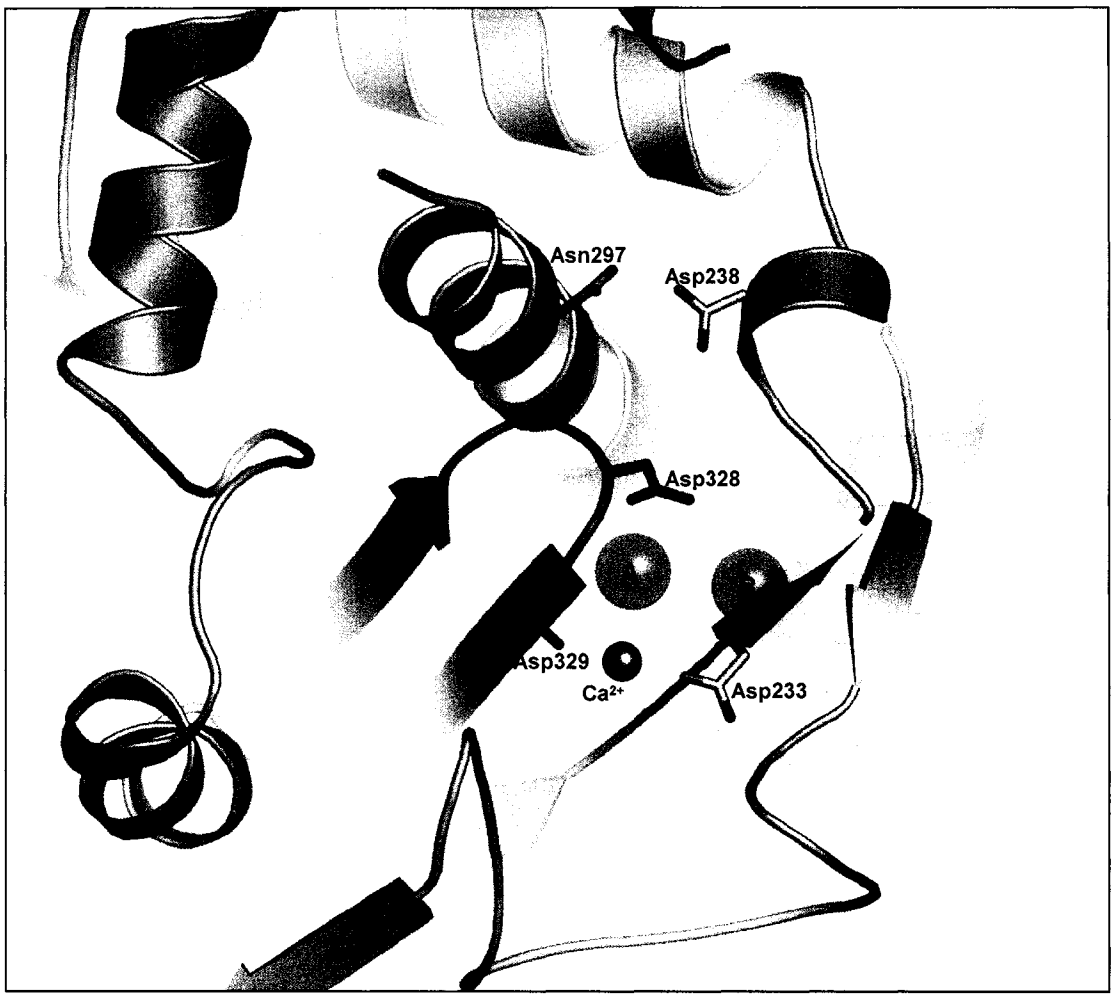


Figure 1.4 Catalytic Center of $3D^{pol}$
Top view of the palm from the partial $3D^{pol}$ structure (Hansen *et al.*, 1997). The active site center is colored magenta and a single bound Ca^{2+} ion is colored orange. The two transparent cyan spheres represent the canonical divalent cation binding configuration (Metals A & B). Side chain residues involved with metal binding (Asp328, Asp329, Asp233) and rNTP sugar selection (Asn297, Asp238) are depicted with sticks.

and its thought to function as a Lewis acid that activates the 3' OH group to an O^- for nucleophilic attack on the α-phosphate of the incoming rNTP. Metal B is distal to the 3' end of the primer and is thought to interact with and stabilize the negatively charged triphosphate moiety in a catalysis competent conformation. This metal may be utilized *ad hoc* and supplied to the active site in a complex with the triphosphate moiety of incoming rNTPs. While Mg^{2+} is generally believed to the cation used *in vivo* (Kornberg and Baker, 1991), several other metals are capable of supporting catalysis. Craig Cameron and colleagues tested several metals and found the following preference with respect to $3D^{pol}$ catalytic efficiency: $Mn^{2+} > Co^{2+} > Ni^{2+} > Fe^{2+} > Mg^{2+} > Ca^{2+} > Cu^{2+}$ (Arnold *et al.*, 1999). However, catalytic efficiency is not intended to imply fidelity as Mn^{2+} is known to permit the increase use of DNA templates and incorporation of incorrect nucleotides (both noncognate and dNTPs) as discussed further in Chapter V. Interestingly, providing only Zn^{2+} to the reaction did not support activity, although in the presence of Mg^{2+} it has a stimulatory effect at low concentrations (60 μM used in activity assays, see appendix 1). It is possible that Zn^{2+} plays a structural role and this is discussed in Chapter II.

There is no question that correct positioning of the incoming nucleotide is essential for nucleophilic attack on the α-phosphate resulting in phosphodiester bond formation. As discussed above, the divalent metals play this role in coordination of the triphosphate moiety. While cations are necessary for catalysis, they are not sufficient for coordination of rNTPs. Two other regions of the polymerase are required for correct positioning of the ribonucleotides. By

analogy to other polymerase structures, basic residues from the fingers domain are believed to be involved and this is discussed in chapter V. Coordination and selection of the correct 2' OH ribose moiety involves two residues situated on the palm domain, Asp238 and Asn297 (Figure 1.4). In the partial polymerase structure, these residues are hydrogen bonded to each other, and mutation of either residue to an alanine reduces $3D^{pol}$'s ability to distinguish between ATP and 2' dATP (Gohara *et al.*, 2000). While both of these mutants have negative effects on polymerase activity, the impact of each mutant is quite different. In the presence of Mg^{2+}, mutating Asn297 to an alanine results in a 20-fold reduction in nucleotide incorporation rate constant and a moderate 20% reduction in overall elongation activity, whereas the Asp238 to alanine mutant reduces the rate constant 2000-fold and activity by over 90%. Although the differences in polymerase activity potentially reflect the importance of each residue in ribose discrimination and orientation, both Asp238 and Asn297 have been proposed to interact with the 2' OH on rNTPs (Gohara et al., 2000, 2004; Huang et al., 1997; Castro et al., 2005).

Aspartate 238 is completely conserved both in terms of sequence and structure in all viral RdRPs for which information is available (Gohara et al., 2000, 2004; Hansen et al., 1997; Koonin, 1991). In DNA polymerases such as the reverse transcriptases from HIV and Moloney murine leukemia virus (MMLV), this position is occupied by hydrophobic phenylalanine or tyrosine amino acids which are thought to sterically interfere with the 2' hydroxyl of rNTPs (Astatke et al., 1998a, 1998b; Bonnin et al., 1999). Interestingly, mutation of Phe155 of the

MMLV reverse transcriptase to a valine results in an enzyme capable of acting as an RNA polymerase (Gao et al., 1997). The converse experiment of mutating the poliovirus Asp238 to an aromatic residue to generate a DNA polymerase was unsuccessful (Gohara et al., 2000; Peersen Lab, *unpublished*). Nonetheless, an aspartate at position 238 could favor rNTPs over dNTPs by directly interacting with the 2' OH group of incoming ribonucleotides.

DISSERTATION CONTENT

My work has been focused on determining the complete structure of the poliovirus polymerase and ascertaining the structural and functional role of the N-terminus in the proteolytic activation of $3D^{pol}$. To achieve this, we intentionally disrupted the persistent Interface I observed in the previous structures and found new conditions for crystallizing the protein in a new packing arrangement where the entire structure could be solved at 2.0 Å resolution. The complete $3D^{pol}$ structure reveals a novel polymerase activation motif where the very N-terminus of the protein is buried in a pocket on the back of the fingers domain (discussed in Chapters II & III). The buried terminus stabilizes a structure that directly positions Asp238 for binding to the 2' hydroxyl group of the incoming nucleoside in the active site. This interaction was confirmed by solving the co-crystal structures of $3D^{pol}$ complexed with all four rNTPs, which are bound with their 2' OH groups making the anticipated hydogen bond with Asp238 (discussed in Chapter V). The final focus of this dissertation examines the role of the fully enclosed polymerase structure on the thermal stability of the

protein. In contrast with DNA polymerases and DNA-templated RNA polymerases, the overall fold of $3D^{pol}$ is similar to other RdRps in adopting a conformation where the active site is fully enclosed as a result of extensive interactions between the fingers and thumb domains. Using circular dichroism, I found that mutations and truncations that disrupt the interactions between the thumb and fingers domains have severe consequences on the thermal stability and activity of this polymerase by precluding formation of a stable enclosed structure (discussed in Chapter V).

CHAPTER II

CRYSTALLIZATION AND STRUCTURE DETERMINATION OF THE POLIOVIRUS POLYMERASE

Determination of the wild-type poliovirus polymerase structure was an important milestone because it was the first structure of an RNA-dependent RNA polymerase (Hansen et al., 1997). It was also very significant in the world of picornaviruses since poliovirus and its replication is the paradigm for this entire family of viruses. Although one-third of the protein's structure was disordered, the enzyme appeared to have the same overall shape as other polymerases that have been described by analogy with a right hand (Ollis et al., 1985). However, as is often the case in science, this structure led to many more questions. The most obvious question concerned the architecture of the disordered fingers domain that comprises almost a third of the molecule. The original structure does show a small segment of the fingers domain where residues 25-37 interact with the top of the thumb and residues 12-24 descend towards the active site. Although the first eleven residues were not observed, the structure suggested that the N-terminus could be an integral component of the catalytic site thus providing a possible explanation for the critical importance of native N-terminus for proper function of the polymerase (discussed in Chapters I

& III). It was also proposed that this stretch of residues (12-37) was contributed in *trans* from another polymerase molecule through an interaction called Interface II (Hansen *et al.*, 1997; Hobson *et al.*, 2001). To identify the important structural role of the N-terminus and its molecular source, my initial goal in the lab was to solve the complete structure of $3D^{pol}$.

Results

Crystallization

Interface I is formed by an interaction between the thumb domain of one polymerase and the back of the palm domain of a second that buries ~2000 $Å^2$ of solvent accessible surface area (Hansen et al, 1997; Hobson et al, 2001). A major molecular interaction along this interface is the insertion of Leu446 from the thumb into a hydrophobic pocket at the bottom of the palm of another polymerase molecule (Figure 1.3). The interface is further stabilized by intermolecular salt bridges linking Arg455–Asp349, Arg456–Asp339, and Asp459–Lys384. Mutations at these residues have been shown to disrupt the formation of $3D^{pol}$ sheet structures observed in electron microscopy studies (Lyle *et al.*, 2002a) and affect viral viability (Hobson *et al.*, 2001). In addition, it is also possible that residues 12–36, observed to bind the top of the thumb domain in the original structure, were contributed in *trans* from another molecule through an interaction called Interface II (Hansen *et al.*, 1997). Based on these observations, four different $3D^{pol}$ mutants were made in an effort to eliminate the native crystal contacts and force the formation of a new crystal lattice.

First, Interface I was disrupted by mutating Leu446 to either alanine or aspartic acid and Arg455 to aspartic acid, i.e. L446A-R455D and L446D-R455D. Second, these mutations were combined with deletion of the first 68 residues of $3D^{pol}$, i.e. Δ68-L446A-R455D and Δ68-L446D-R455D, to eliminate the proposed Interface II by initiating protein synthesis at the first ordered residue of the original structure. All four proteins expressed well and were purified at

significantly higher yield than the wild type protein. This is probably due to the higher solubility of the mutants, which increases from ~3 mg/ml for the wild type protein to ~20 mg/ml for $3D^{pol}$-Δ68-L446A-R455D and greater than 40 mg/ml for $3D^{pol}$-L446D-R455D in 200 mM NaCl. New crystallization conditions were found for two of the mutants by hanging drop vapor diffusion (Figure 2.1).

The $3D^{pol}$-Δ68-L446A-R455D crystals grew in 8 days with a precipitant/well solution containing 1.5 M ammonium formate, 0.1 M sodium chloride, 50 mM HEPES (pH 7.0), 2 mM DTT, and 0.02 % (w/v) sodium azide. $3D^{pol}$-L446D-R455D crystals grew in 4 days with a precipitant/well solution containing 2 M sodium acetate, 0.1 M cacodylic acid (pH7.1), 2 mM DTT, and 0.02 % sodium azide. All crystals were transferred into corresponding precipitant solutions containing 30 % (v/v) glycerol over the course of 30 minutes prior to flash freezing with liquid nitrogen.

Structure Determination

Data were initially collected at Colorado State University using a Rigaku Automated X-ray Imaging System (R-AXIS) IV with Cu-Kα radiation. Crystals were maintained at 100°K using a nitrogen cryostream during data collection. The oscillation range varied between 0.5 and 1 degrees and the reflections were integrated, merged, and scaled using Denzo/Scalepack (Otwinowski and Minor, 1997) and d*TREK (Pflugrath, 1999). The search model for molecular replacement was the 2.6 Å-resolution partial structure of $3D^{pol}$ (PDB code 1RDR) without water molecules and ions (Hansen *et al.*, 1997). Cross rotation and

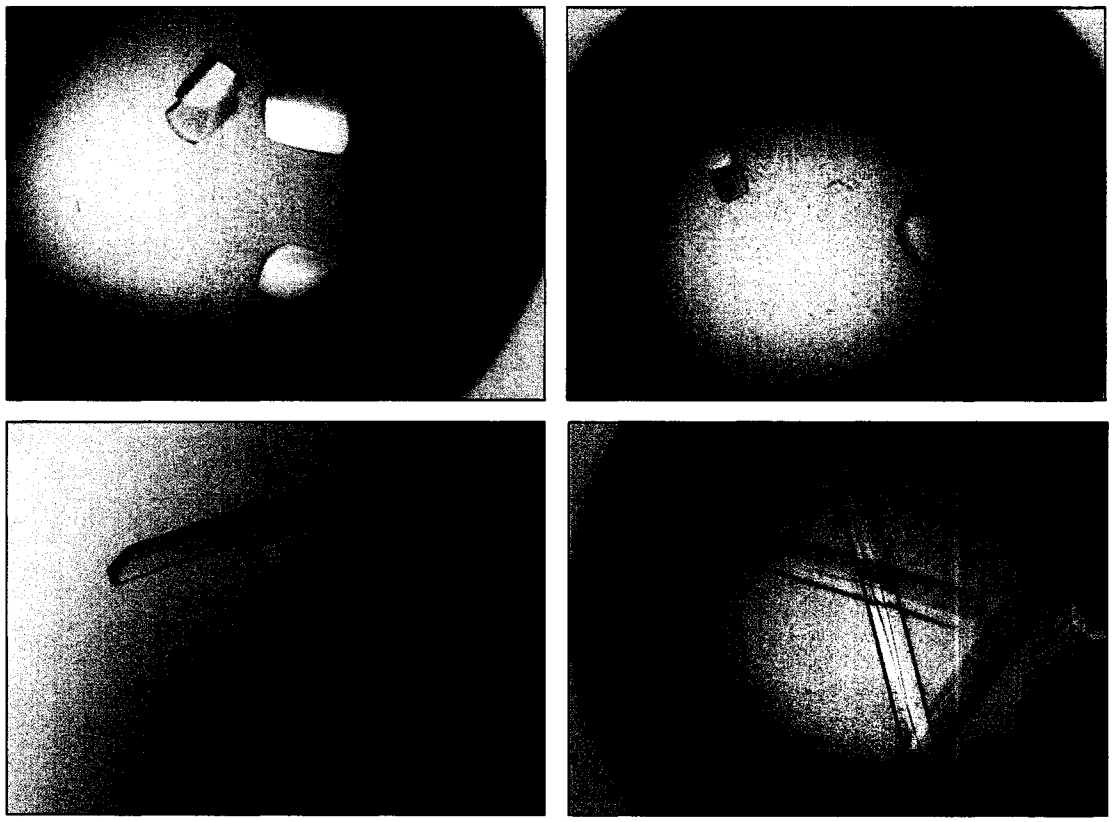


Figure 2.1 Crystals of 3D^{pol}-Δ68-L446A-R455D and 3D^{pol}-L446A-R455D
(**A**) (**B**) Crystals of the 3D^{pol}-Δ68-L446A-R455D mutant in a $P3_121$ lattice. Typical crystals are 0.4 X 0.3 mm and diffract to ~2.6 Å resolution
(**C**) (**D**) Crystals of the 3D^{pol}-L446D-R455D mutant in a $P6_5$ lattice. Typical crystals are 1.5 X 0.2 mm and diffract to ~2.0 Å resolution.

translation searches were performed with the CNS software package (Brunger *et al.*, 1998) using 15 - 4 Å resolution data. 2Fo-Fc and Fo-Fc electron density maps were calculated after 30 rounds of rigid-body refinement, and inspection of the $3D^{pol}$- L446D-R455D maps using the program O (Jones *et al.*, 1991) revealed density for the entire fingers domain.

$3D^{pol}$-L446D-R455D Refinement

The first round of refinement involved deletion of amino acids from the initial structure solution that did not fit well into electron density, resulting in a slight increase in the R-free (Figure 2.2). Amino acids were then manually fit into electron density over the next 12 rounds of refinement using the full resolution range. After resetting all the atomic B-factors to 40 $Å^2$, refinement of the atomic positions and B-factors generally consisted of a B-group, conjugate gradient energy minimization, and B-individual using a maximum likelihood intensity refinement target. Anisotropic B-factor corrections were applied after the R-factor and R-free dropped to approximately 30 percent. During the course of refinement, three different crystal data-sets of increasingly superior quality were utilized (Figure 2.2). To eliminate any systematic model bias in the phases, simulated anneal omit maps were calculated throughout the model building process.

Water molecules were added to the structure during the 14th round of refinement to positive Fo-Fc peaks greater than or equal to 3σ and when this position was within hydrogen bonding distance (2.4 - 3.1 Å) of appropriate

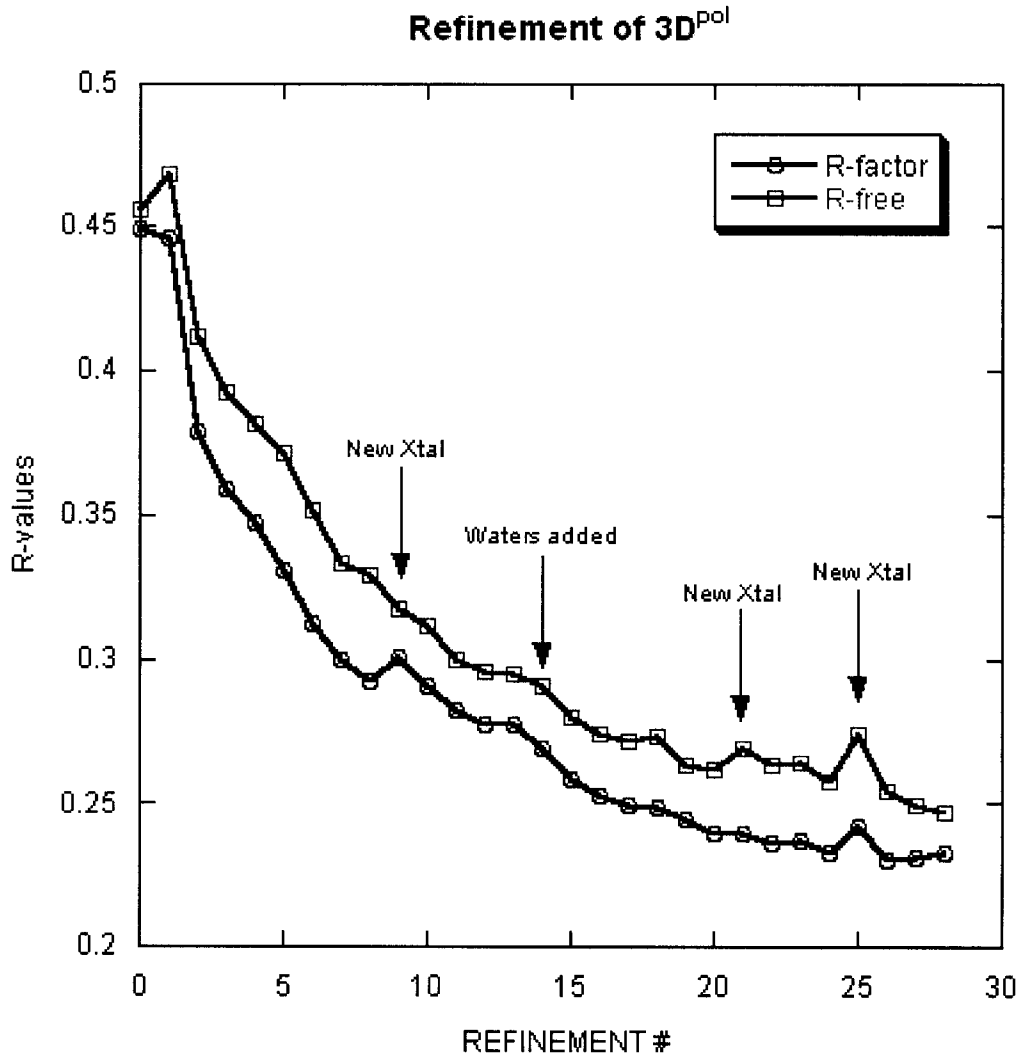


Figure 2.2 R-values during refinement of 3D^{pol}.
Graph plotting the decrease in R-factor and R-free as a function of refinement round.

hydrogen bond donors or acceptor groups. Following refinement, waters were removed if the B-factors exceeded ~60 $Å^2$ or if the 2Fo-Fc peak could not be observed when contoured to ~1.2 - 1.4σ. While several large electron density peaks were successfully modeled in as acetate ions, the density around several solvent-exposed cysteine residues suggested the presence of an electron-rich ligand. Anomalous difference maps further indicated this ligand was an anomalous scatterer. Crystallization of the $3D^{pol}$-L446D-R455D mutant required cacodylic acid and DTT, consistent with a proposed mechanism (Tsao and Maki, 1991) for the formation of the dimethyl arsenic adducts covalently bound to cysteine residues that had previously been observed in several structures (Maignan *et al.*, 1998; Tete-Favier *et al.*, 2000; Raman *et al.*, 2001). Cysteine residues modified with the dimethyl arsenic adducts were refined in CNS as CAS residues using a parameter set from the HIC-up database (Kleywegt and Jones, 1998). This iterative process resulted in convergence of the R-values around 25 % (Figure 2.1 and Table 2-1). All data reductions and refinements included anomalous scattering components.

	$3D^{pol}$-Δ68-L446A-R455D	$3D^{pol}$-L446D-R455D
PDB Code	1RAJ	1RA6
Space Group	$P3_121$	$P6_5$
Unit Cell (Å)	a=b=87.7 c=107.3	a=b=127.6 c= 113.0
X-Ray Source	Cu-Kα	NSLS X25 – 1.1Å
Resolution Limits	30 - 2.5 (2.59-2.50)	30 - 2.0 (2.07-2.00)
Reflections		
Total Collected	169949	236338
Unique	16973	67129
Redundancy	10 (8)	3.5 (3.1)
I/σ [1]	35 (3.3)	16.5 (2.5)
Completeness (%)	99.9 (99.6)	96.3 (93.7)
R_{merge} (%)	6.2 (56.9)	5.4 (45.6)
Refinement		
Resolution Range	30 - 2.5	30 - 2.0
R	24.2	24.3
R-free (10 % of data)	26.2	25.6
Model Statistics		
Number of atoms (waters)	2138 (67)	4091 (298)
Average B factor	56.9	51.4
rmsd bond lengths	0.008 Å	0.007 Å
rmsd bond angles	1.4°	1.4°
Ramachandran statistics [2]		
Favored	209	380
Allowed	16	28
Generous	0	2
Disfavored	0	1

Table 2-1 Data Collection and Refinement Statistics
Data collection and refinement statistics for the $3D^{pol}$-Δ68-L446A-R455D and full-length $3D^{pol}$-L446D-R455D mutants. Data in parentheses are for the highest resolution shell, and Ramachandran statistics do not include glycine and proline residues

Discussion

3D^{pol}-Δ68-L446A-R455D Structure

The 3D^{pol}-Δ68-L446A-R455D mutant crystallized in the space group $P3_121$ and provided a 2.5 Å-resolution structure (Table 2-1). Despite the increased solubility of the mutant due to the destabilization of Interface I, this protein crystallized with Interface I intact. The fingers domain continued to be disordered and the structure is nearly identical to that of the wild type 3D^{pol} with an r.m.s. difference in C_α positions of 0.6 Å (Figure 2.3).

There is one area of significant structural deviation between the 3D^{pol}-Δ68-L446A-R455D and wild type 3D^{pol} structures located at the tip of the thumb. The original wild-type structure showed that a small segment of the fingers domain (residues 25-37) interacts with this region of the thumb, and as discussed below, this interaction is also observed in the full-length complete structure. However, the polymerase structure containing the 68-residue N-terminal truncation shows that Trp403 and the flanking residues (amino acids 401-406) are flipped out in a more exposed conformation (Figure 2.4). It should be noted that the flipped-out Trp403 observed in the 3D^{pol}-Δ68-L446A-R455D structure is an integral component of a crystal lattice contact. The insertion of Trp403 into a hydrophobic pocket located at the bottom of the palm of a symmetry related molecule constitutes a major component of this lattice contact. Nonetheless, the ability of Trp403 to adopt an alternate conformation demonstrates flexibility at the top of the thumb domain. As discussed in

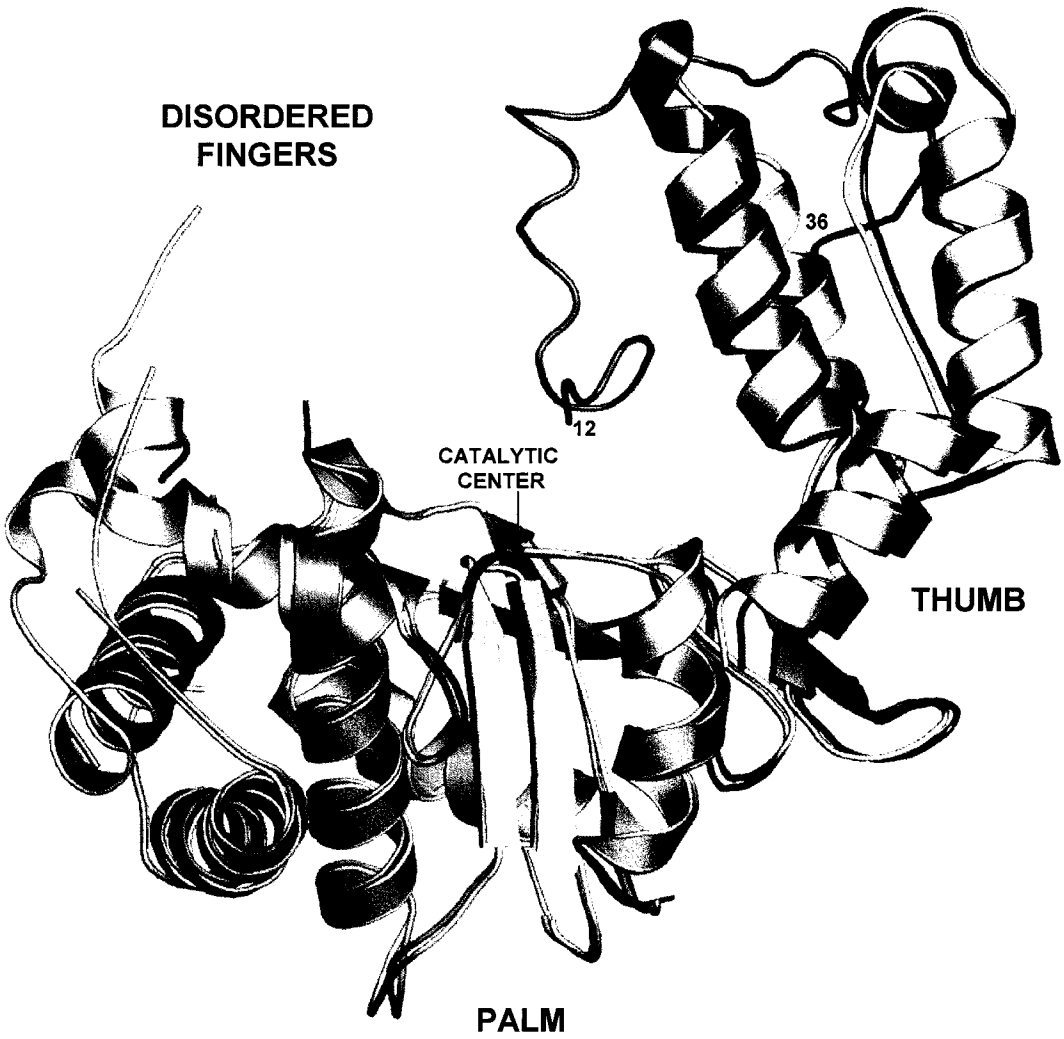


Figure 2.3 3D^{pol}-Δ68-L446A-R455D structure superimposed on wild-type 3D^{pol}
Structure of the3D^{pol}-Δ68-L446A-R455D mutant with the palm colored grey, thumb in blue, and the catalytic center in magenta superimposed onto the partial wild-type structure colored yellow. The N-terminal strand (residues 12-36) of the original wild-type structure that descended toward the active site is shown in green. The fingers domain is disordered in both of these structures.

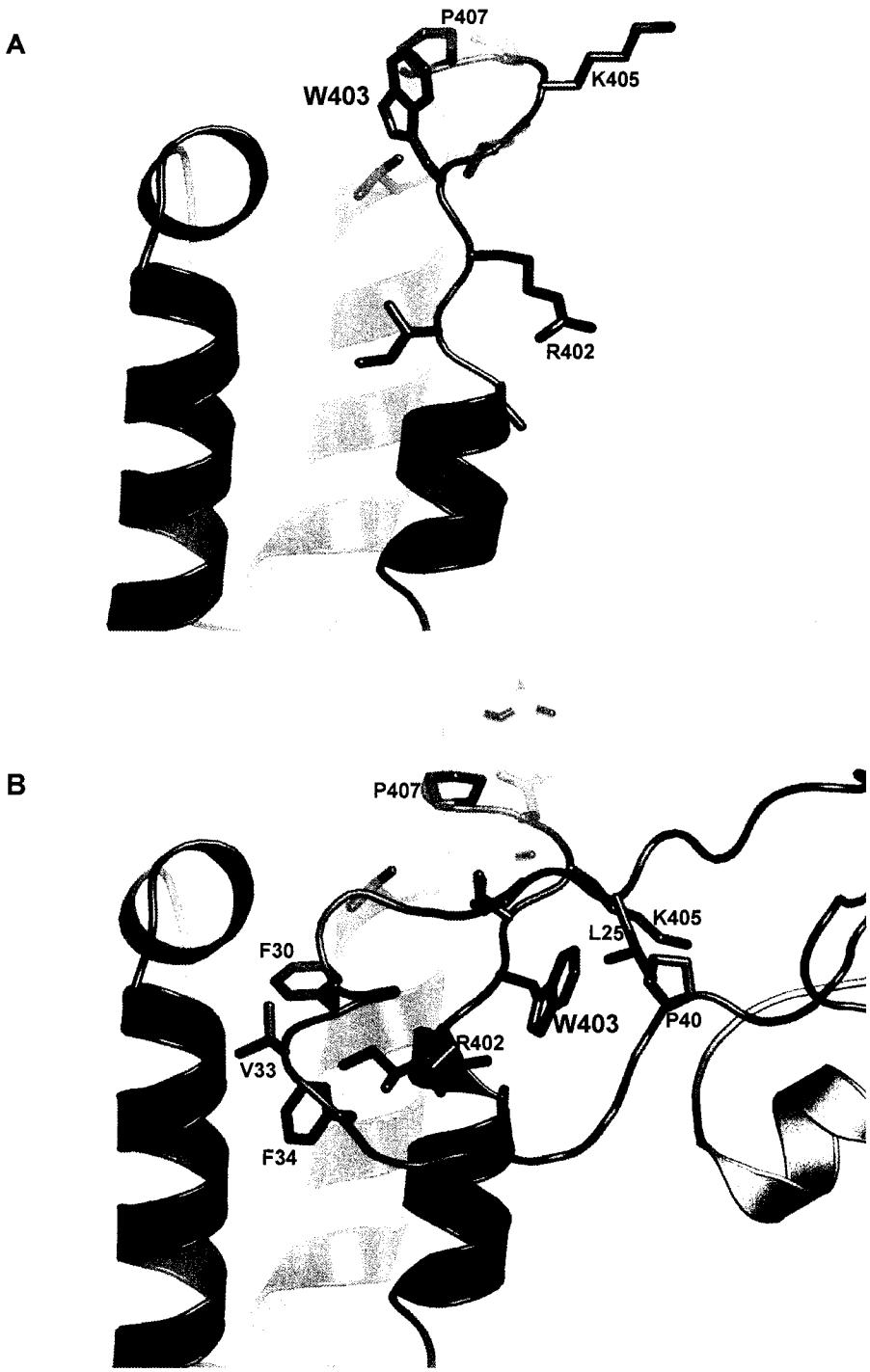


Figure 2.4 Comparison of the thumbtip from the 3D^{pol}-Δ68-L446A-R455D and complete structures.
(**A**) Back view of the top of the thumb from the 3D^{pol}-Δ68-L446A-R455D structure. (**B**) Same view of the complete structure shown with the index finger in green and ring finger in yellow. In the presence of the index finger, the conformation of the thumb-tip (centered around Trp403) undergoes an induced fit conformational change.

Chapter VI, this region plays a major role in stabilizing the enclosed $3D^{pol}$ structure.

$3D^{pol}$-L446D-R455D Complete Structure

The full-length $3D^{pol}$-L446D-R455D mutant crystallized in a new $P6_5$ lattice (Table 2-1) where polymerase molecules did not interact across Interface I. Electron density for the missing fingers domain was clearly visible in the maps after molecular replacement and the entire structure was solved at 2.0 Å resolution. The domain structure of this enzyme was consistent with the usual right-hand analogy of a thumb, palm, and fingers domain that was first used to describe the structure of the DNA polymerase I Klenow fragment (Ollis *et al*, 1985). Like other RdRP structures, $3D^{pol}$ adopts a "closed" conformation where extensive interactions between the thumb and fingers domains completely encircle the active site and create a NTP entry tunnel at the back of the polymerase (Figure 2.5). The structures of the palm and thumb domains are essentially identical to those seen in the initial partial structure (Hansen *et al*, 1997). There is a minor ~2.5° change in the relative orientations of these domains that may be due to the removal of crystal packing constraints along Interface I (Figure 2.6). The new aspect of our complete $3D^{pol}$ structure is the fingers domain, which is composed of four separate stretches of amino acid sequence. These form four tightly intertwined finger structures that we describe by analogy to primate anatomy where there are index, middle, ring, and pinky fingers in addition to a thumb (Figure 2.7).

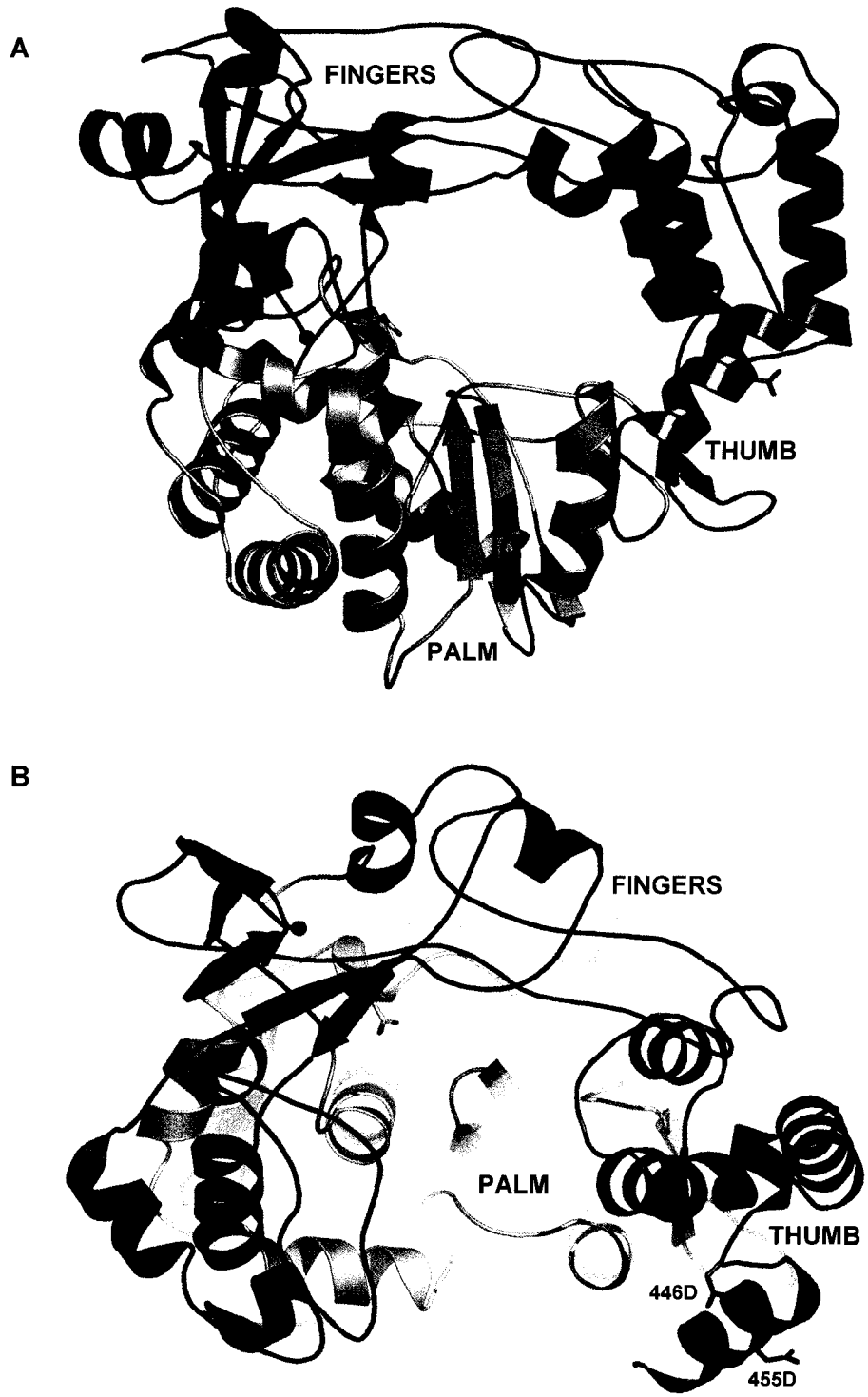


Figure 2.5 Complete 3D^{pol}-L446D-R455D structure.
(**A**) Front view of the complete 3D^{pol} structure highlighting domains of the right hand analogy. The fingers, palm, thumb domains are colored red, grey and blue, respectively and the catalytic centers are colored magenta. The side chain of Asp238, Asp446, and Asp455 are depicted as sticks and the N-terminus is shown with a blue sphere. (**B**) Top view of panel (A).

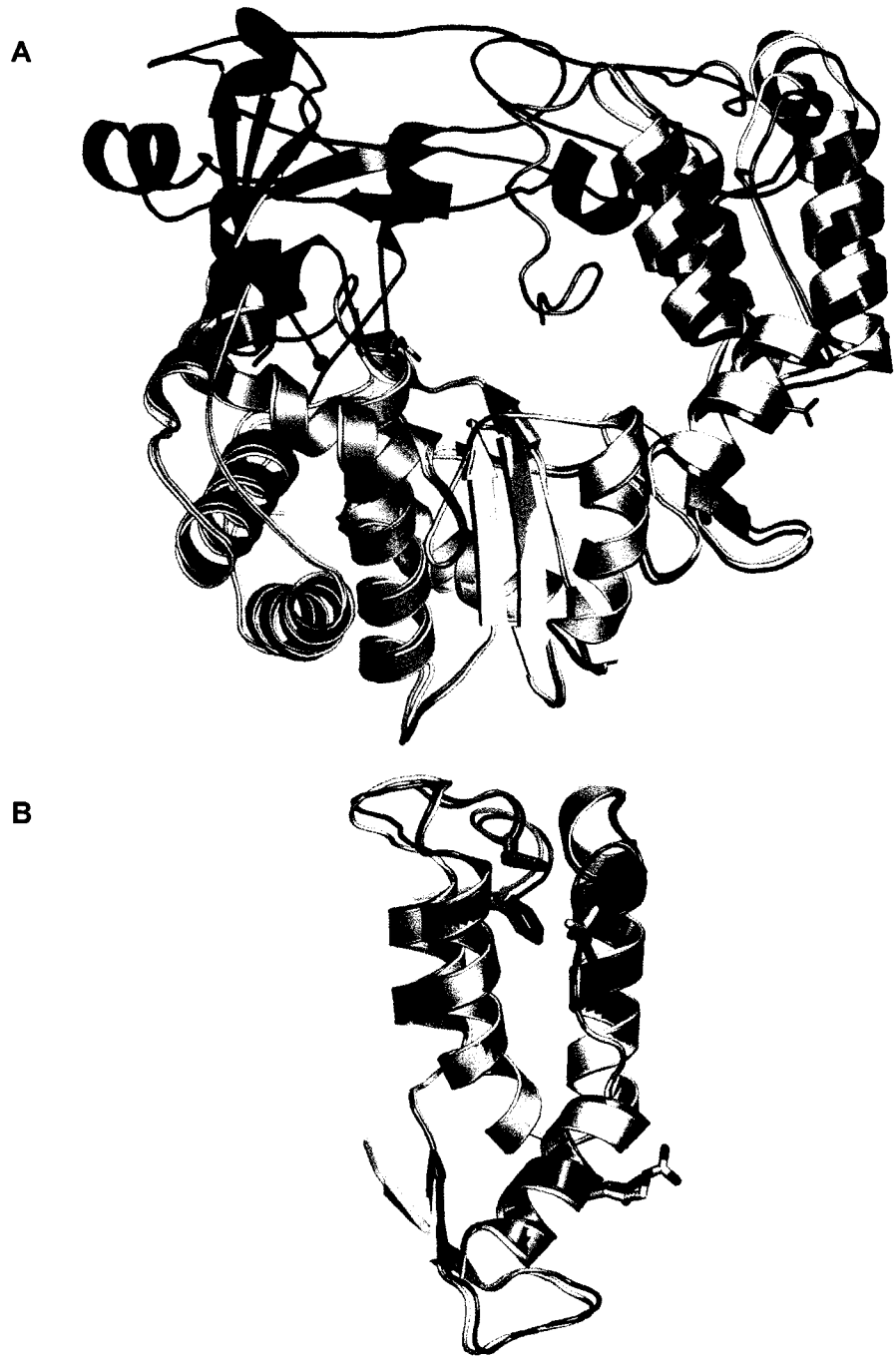


Figure 2.6 Comparison of the Complete 3D^{pol}-L446D-R455D structure with the Partial Wild-type Polymerase Structure

(**A**) Comparison of the original partial structure (yellow) with the complete 3D^{pol} structure shown with the fingers domain in red, the palm in grey, the thumb in blue, and the active site colored magenta. The N-terminal strand (residues 12-36) of the original structure that descended toward the active site is shown in green. These structures are superimposed using the backbone of the active site (residues 324-332) showing a 2.5 degree difference in the orientation of the thumb domains. .(**B**) Superposition of the thumb domains from from the original structure (yellow) and complete structure (blue) showing that the thumb is not perturbed by the two mutations L446D and R455D.

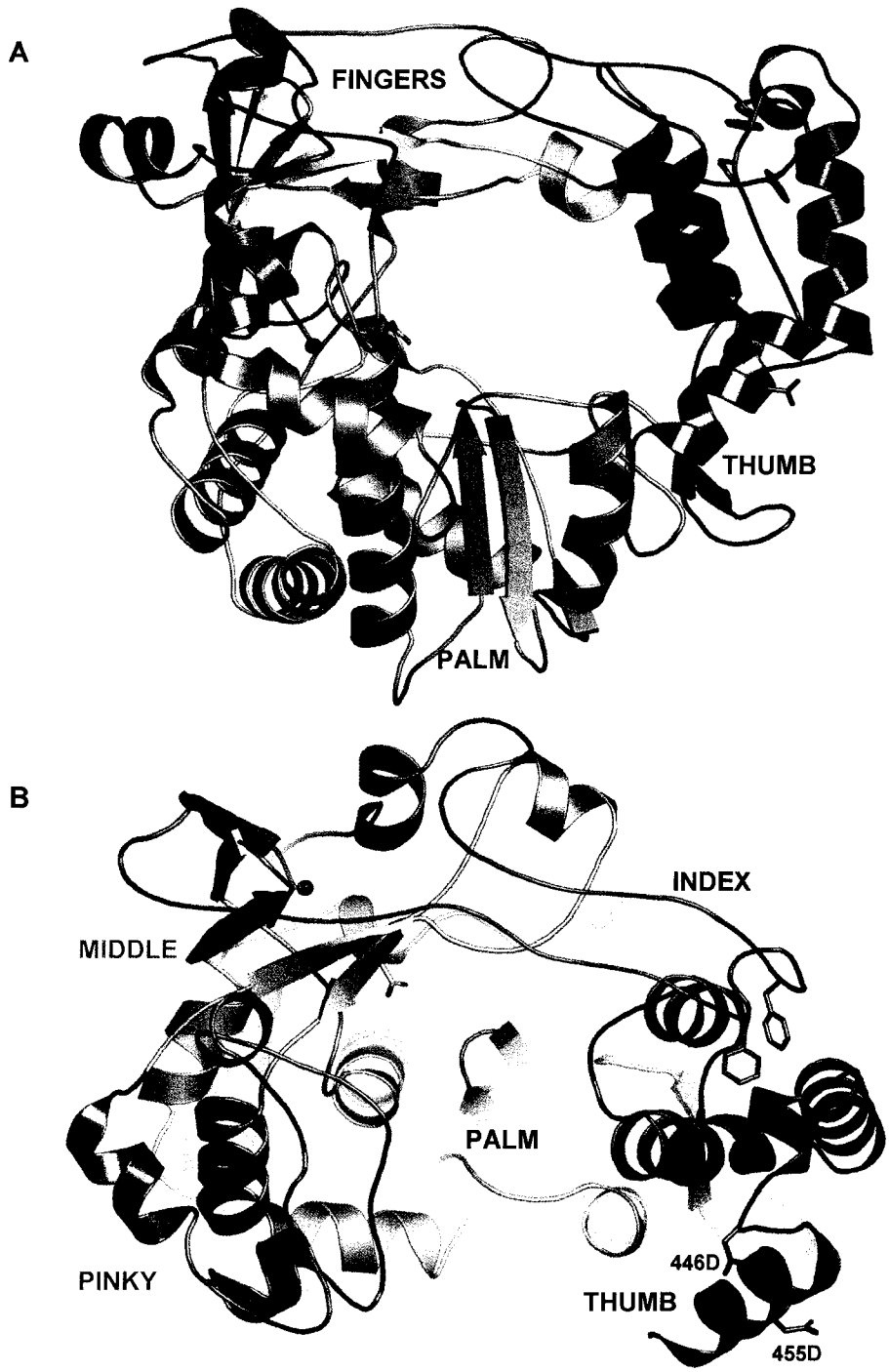


Figure 2.7 3D^{pol} finger structures
(**A**) Front view of the complete 3D^{pol} structure highlighting the individual fingers of the fingers domain. The index finger is shown in green, the middle finger in orange, the ring finger in yellow, and the pinky finger in pink. As in fig. 2.5, the palm is shown in grey, the thumb is in blue, and the active site is colored magenta. Asp238, Asp446, Asp455, Phe30 and Phe34 are shown in sticks and the N-terminus is shown with a blue sphere. (**B**) Top view of panel (A).

The fingers begin with a buried N-terminus at the back of the palm that plays a critical role in structuring the active site (see below). The index finger (residues 1-68, green in Figure 2.7) then rises from the palm domain for eight residues before folding into a loop and reaching across the palm to interact with the thumb. This conformation is anchored by the insertion of Phe30 and Phe34 into the hydrophobic core at the top of the thumb. Following this, residues 35–68 fold back toward the palm domain in a generally extended structure that completes the index finger. The middle finger (residues 269-285, orange) consists of an anti-parallel β-sheet with a β-turn at its fingertip. This sheet also includes the first eight residues of the index finger, and the tip of the middle finger is inserted into the loop formed by residues 9-17 of the index finger. Lys276, the only Ramachandran plot outlier in the structure, is located at the tip of the middle finger. There is not an obvious structural reason for the distorted geometry, but there may be a functional requirement for a basic residue at this position because a K276L mutation reverts *in vivo* to an arginine residue (Richards and Ehrenfeld, 1997). At the base of the middle finger we observe a large positive electron density peak that appears to be from a cation, based on its coordination to four main-chain carbonyl oxygens and a serine, cation-oxygen distances, and B-factors. We have modeled this as a Na^+ ion, but perhaps this site is normally occupied by Zn^{2+}, providing an explanation for the stimulation of activity in the presence of this cation (data not shown). Next, the ring finger (residues 150-179, yellow) forms an elongated β-sheet structure that reaches across the active site and protrudes from the polymerase by traversing under the

index finger to present a short α-helix on the surface of the protein. The ring finger forms the roof of the NTP entry tunnel and contains conserved basic residues (Arg163, Lys167, Arg174) that are poised for interactions with the incoming NTP. The pinky finger (residues 96-149 and 180-190, pink) is a fairly large structure that is separated from the other fingers by a cleft at the top of the fingers domain (Figures 2.7 & 2.8). Lining the cleft on the pinky is a proline (Pro119) that is 100 % conserved and is found in the *cis* conformation. The tip of the pinky fingers appears to be quite flexible by virtue of weak electron density. The folding of the pinky is likely dependent on the ring finger adopting its proper conformation (and vice versa) as the roof of the NTP entry tunnel because the ring finger is effectively an insertion in the middle of the pinky sequence (Figure 2.7).

The intertwined structures of the four fingers are stabilized by a multitude of interactions where one set of interactions will form the platform upon which another set of interactions is built. For example, there is an extensive hydrogen bonding network in which the first eight residues of the index finger and the entire β-turn middle finger together form a three-stranded anti-parallel β-sheet (Figures 2.7 & 2.8). The edge of this sheet is then hydrogen bonded to the backbone of the ring finger, stabilizing the ring finger as it crosses the top of the active site to form the roof of the NTP entry tunnel. The finger structure is further anchored by a well-ordered salt bridge between Lys61 on the index finger and Asp177 on the ring finger that lines the side of the NTP entry tunnel. A mutation of Lys61 to leucine abolishes polymerase activity (Richards *et al*, 1996). At the top of the NTP entry tunnel there are a number of hydrogen bonding and van der Waals

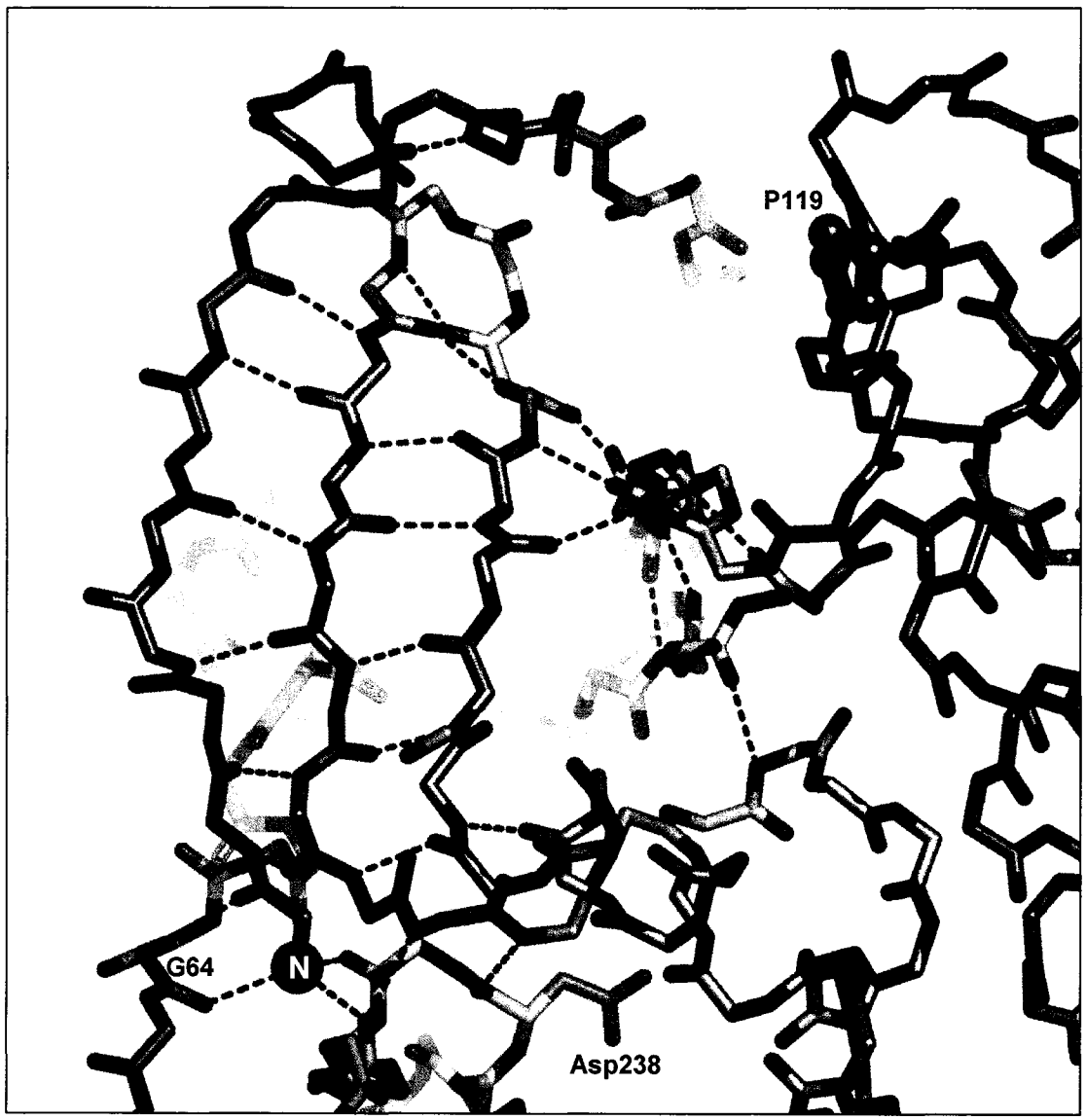


Figure 2.8 Hydrogen bond network around the N-terminus
Structure of the fingers domain highlighting the extensive network of hydrogen bonds linking the N-terminus (blue sphere lower left) and the N-terminal strand of the index finger to the middle and ring fingers. Note the putative template entry channel separating the pinky finger (pink) from the rest of the fingers domain and how the ring finger (yellow) is an insertion in the pinky finger structure. Proline 119, which may play a role in template binding is highlighted, and Asp238 is colored cyan.

packing interactions between the index finger and the ring finger that are dependent on a kink in the index finger structure. This kink is stabilized by hydrogen bonding interactions involving Pro40, Glu47, and Arg49 and these residues are 100% conserved across picornaviruses, suggesting that the kink is common to all these viral polymerases. Correct positioning of the ring finger may be dependent on the conserved hydrophobic amino acids (Phe30, Val 33, Phe34) of the index finger responsible for anchoring the fingers to the top of the thumb, thereby providing a stable structural support platform for the fingers domain.

Buried N-terminus

The most remarkable and unique feature of the poliovirus $3D^{pol}$ structure is an elegant proteolytic processing-dependent allosteric switch for polymerase activation that involves burying the N-terminal glycine residue in a pocket at the base of the fingers domain (Figure 2.9). A comparison of the complete structure with the original wild type and the $3D^{pol}$-Δ68-L446A-R445D structures shows that the buried N-terminus is involved in positioning Asp238 in the active site. When the three $3D^{pol}$ structures are superimposed using the protein backbone of the active site Gly-Asp-Asp motif and three residues on either side of it (residues 324-332), there is a clear 1.4 Å movement of Asp238 toward the active site that is apparent only in the complete structure (Figure 2.10). Importantly, this movement is relative to the other active-site residues, which are highly superimposable among the three structures. In RNA polymerases this asparate hydrogen bonds to the 2' OH of the incoming NTP in an interaction that is important for the selection of rNTPs over dNTPs (Gohara et al, 2000; Huang et al, 1997). Asp238 is

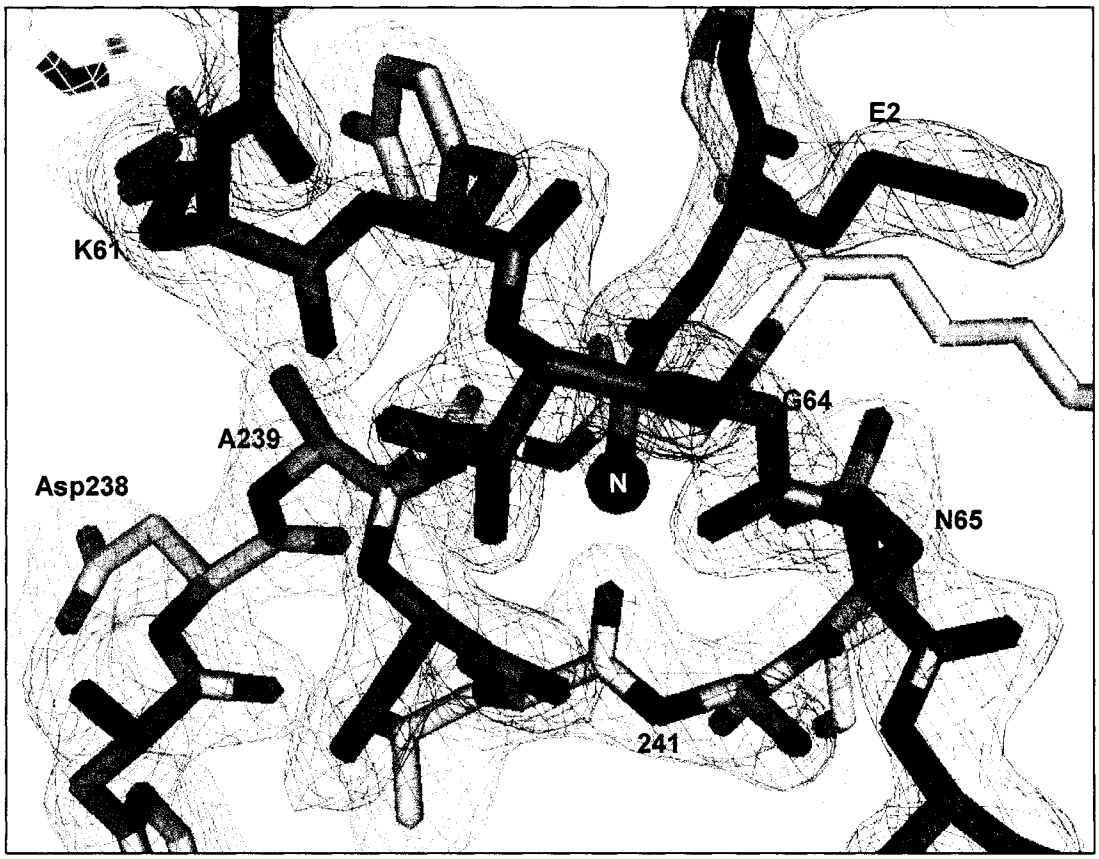


Figure 2.9 Composite Omit Electron Density Map Contoured Around the N-terminus

Electron density map of the region surrounding the buried N-terminus. The map is a 2.0 Å resolution simulated annealing (1500 K) composite omit $2F_o$-F_c map contoured at 1.6σ. The carbon atoms are colored as in Figure 2.7 and the corresponding sections of the density map are colored differently for clarity.

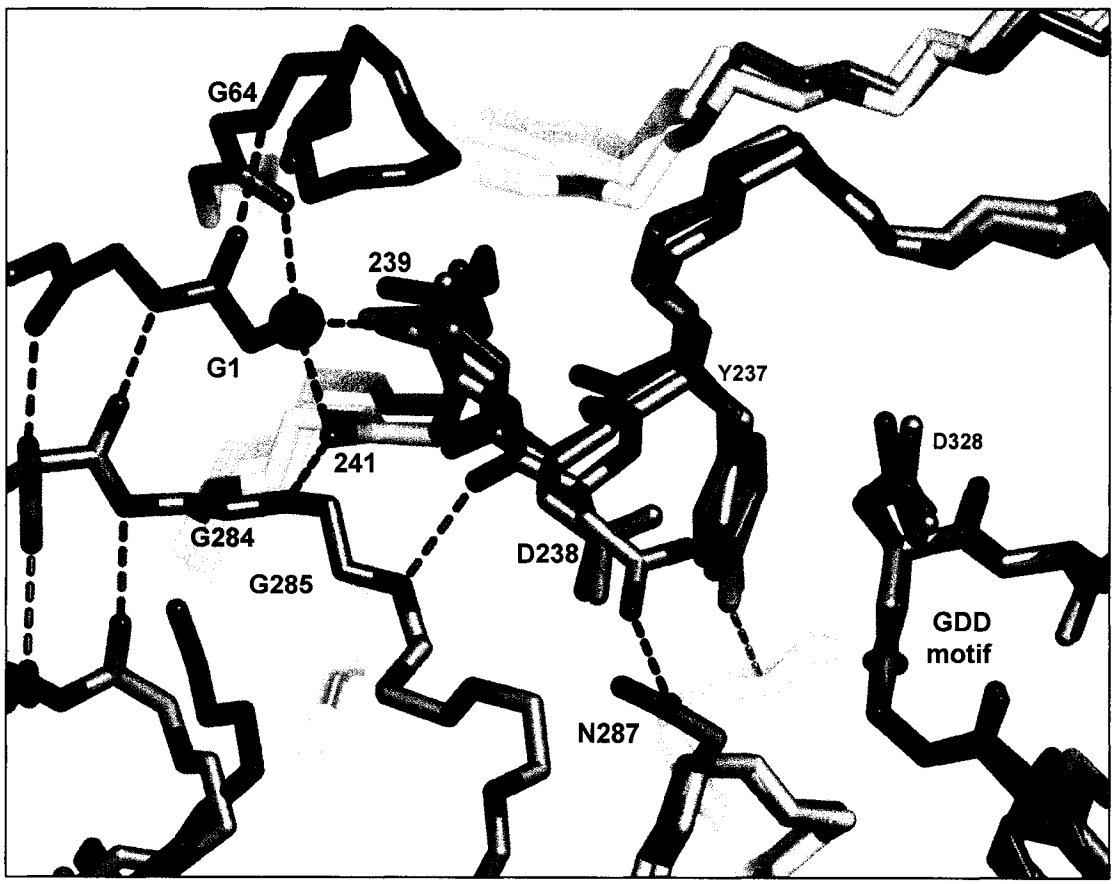


Figure 2.10 Selective Movement of Aspartate 238
Superposition of three $3D^{pol}$ structures showing the selective ~1.4 Å movement of Asp238 toward the active site when the N-terminus is properly positioned. The original partial wild-type structure is in pink, the $3D^{pol}$-Δ68-L446D-R455D is in salmon, and the complete structure is colored by atom type with carbons colored according to structural motifs as in Figure 2.7. Most side chains have been omitted for clarity and residues 324-332 of the active site (magenta) were used for the superpositions.

essential in poliovirus $3D^{pol}$ as a mutation to alanine abolishes poliovirus polymerase activity and viral viability (Gohara *et al*, 2000). The movement of Asp238 is primarily hinged on a change in the backbone between Asp238 and the adjacent Tyr237, which does not move significantly. The source of the movement can be traced to a pair of hydrogen bonds between the buried N-terminus and the backbone carbonyls of residues 239 and 241. These act to pull the polypeptide backbone of residues 239-242 up toward the buried N-terminus, effectively pushing Asp238 into the active site. This conformation is further stabilized by pairs of backbone hydrogen bonds linking residues 285–241 and 286–238 and a side-chain interaction between Asp238 and Asn297.

The binding pocket for the N-terminal glycine residue is itself almost entirely composed of glycines and the structure makes use of both the small size and backbone torsional flexibility of this amino acid. There is no formal charge counter-ion for the buried amino terminus. Instead, the terminal amino group makes hydrogen bonds with the backbone carbonyls of residues 64, 239, and 241 in almost perfect tetrahedral geometry. Additionally, the Gly1 amide is also hydrogen bonded to the Gly64 carbonyl, and the Asn65 side chain is within hydrogen-bonding distance of the very N-terminus. One side of the binding pocket for the N-terminus is made up of glycines 284 and 285 from the base of the middle finger. These glycines are 100% conserved among picornaviral polymerase sequences and their backbone conformations are in the disallowed region of the Ramachandran plot. The outside of the pocket consists of Gly64 from the end of the index finger sequence that is also in the disallowed region of

the Ramachandran plot. Gly64 makes two reciprocal amide-carbonyl backbone hydrogen bonds with Gly1 that link together the two ends of the index finger sequence. Overall, the structure around the N-terminus relies on backbone contacts and glycine flexibility to make interactions that leave little room to accommodate mutations in the protein or improper proteolytic processing of the polymerase.

Conclusions

The complete structure of $3D^{pol}$ provides an explanation for why the fingers domain was disordered in the original partial wild-type structure solved by Hansen *et al.* (1997). When our complete structure is superimposed into the $P3_221$ crystal lattice of the original structure there are two significant steric clashes involving the fingers domain; one is at the loop of the index finger as it wraps around the middle finger and the other involves the pinky finger (Figure 2.11). Neither of these crystal contacts is a direct result of interactions across Interface I, the interface that was intentionally disrupted to obtain the complete structure. Rather, the clashes arise from secondary crystal packing interactions between parallel Interface I fibers. Based on this, we find it unlikely that the proposed Interface II of the original structure exists in solution. This interface, where residues 12-35 interacting with the top of the thumb were proposed to come in *trans* from another polymerase molecule, involves the burial of two large hydrophobic residues into the top of the thumb (Phe30 and Phe34, see Figure 2.7). We feel that this interaction is likely too strong to have been disrupted by the unfolding of the fingers domain and therefore remained intact in the crystal lattice of the original structure.

The overall fold of poliovirus $3D^{pol}$ is similar to that of other viral RNA-dependent RNA polymerases in adopting a conformation where the active site is fully enclosed as a result of extensive interactions between the fingers and thumb domains. This conformation has now been observed in all other RdRP structures and is discussed further in Chapter VI. A number of important interactions in the

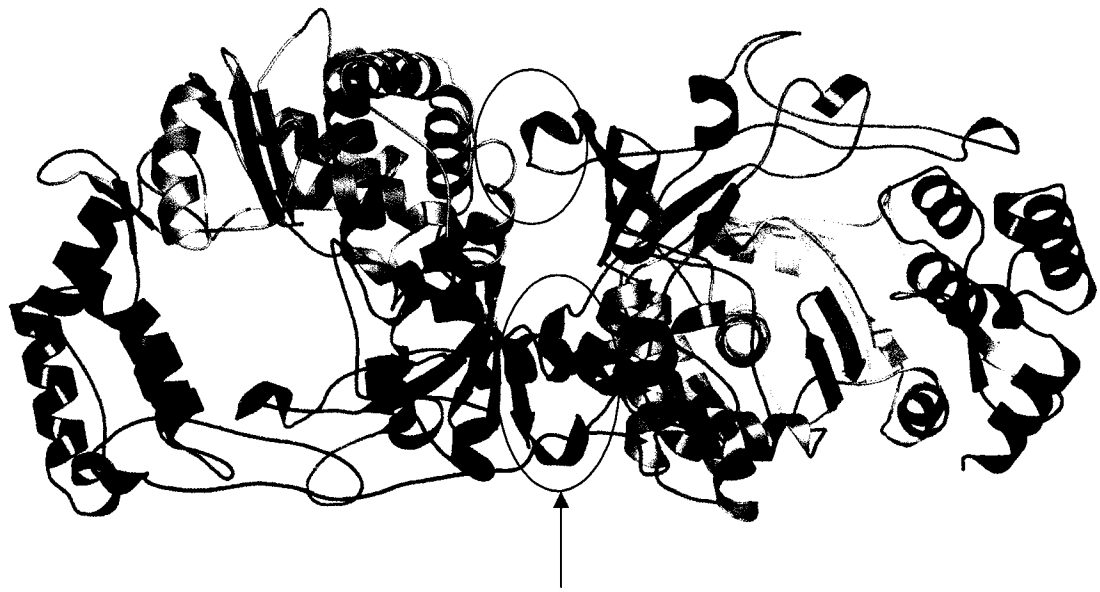


Figure 2.11 Steric Clash in $P3_221$ Crystal Lattice
Steric clashes observed when the complete structure is superimposed into the P3221 crystal lattice of the original partial structure. The clashes of the fingers domain provides an explanation for the disorder of this domain in this crystal lattice. The fingers domains are colored red and pink on the different molecules to contrast these regions

poliovirus $3D^{pol}$ structure are dependent on residues that are highly conserved among the picornaviruses, suggesting that many aspects of the poliovirus $3D^{pol}$ structure are also retained among these viral polymerases. These interactions include the Lys61–Glu177 salt bridge between index and ring fingers that lines the NTP entry tunnel, the hydrophobic residues on the index fingertip that are inserted into the top of the thumb, and the kink in the index finger structure where it interacts with the ring finger to form the top of the NTP entry tunnel.

The $3D^{pol}$ fingers domain has its own hydrophobic core that includes residues from the index, middle, and ring fingers and extends from the base of the fingers, across the top of the NTP entry tunnel, and over to the thumb. The pinky finger also has a small self-contained hydrophobic core that is separate from that of the other fingers. As noted previously, the ring finger is effectively an insertion in the pinky finger sequence (Figure 2.7) and as a result the folding of the two hydrophobic cores within the fingers domain is probably linked. Importantly, there are relatively few interactions between the palm and fingers domains. A number of observations suggest that the poliovirus polymerase palm and thumb domains may fold independently of its fingers domain. The lattice of the original partial structure showed that the fingers can be selectively disordered while leaving the palm and thumb intact. Our structure of the 68-residue N-terminal deletion mutant further showed that the entire index finger can be deleted without affecting the folding of the palm and thumb. The buried N-terminus sits at the heart of an extensive array of hydrogen bonds and may play a significant role in nucleating the structure of the entire fingers domain (Figure 2.8). If this is

the case, then the structure of the fingers domain may be quite different in the context of $3CD^{pro}$ where the proper $3D^{pol}$ N-terminus does not exist. In other words, the structure of $3CD^{pro}$ may not simply be a fusion of the $3C^{pro}$ structure (Mosimann et al., 1997) and the complete $3D^{pol}$ structure.

The poliovirus $3D^{pol}$ structure provides the first structural insight into the molecular mechanism responsible for the proteolysis dependent activation of a polymerase. It shows that the newly created $3D^{pol}$ N-terminus is buried in a pocket at the junction of the fingers and palm domains. This in turn appears to position the essential Asp238 residue in the active site for interactions with the 2' hydroxyl group of the incoming rNTP. The original proposal suggesting the N-terminus structures the active site turns out to be correct, albeit in a more indirect manner.

Because the $3D^{pol}$-L445D-R455D construct significantly increased the solubility of this protein, this double point mutation is found on all subsequent mutants generated for this dissertation (unless otherwise noted). Additionally, the complete structure of this polymerase construct will hereby be referred to as the "native" structure throughout the rest of the dissertation.

CHAPTER III

MUTATIONAL ANALYSIS OF THE N-TERMINUS AND ITS BINDING POCKET

It is known that proper proteolytic processing of $3D^{pol}$ is required to activate the polymerase as the $3CD^{pro}$ precursor has absolutely no RNA polymerase activity despite the fact that it contains the entire polymerase domain (Flanegan and Van Dyke, 1979; Harris *et al*, 1992). As noted in chapter I, modifications of the N-terminus were found to be deleterious for the polymerase activity of this enzyme. Examination of the buried N-terminus of the complete structure shows that the conserved glycine 1 of $3D^{pol}$ fits tightly into a pocket, which itself is lined with conserved glycines (Figure 3.1). A comparison of the complete native structure with the $3D^{pol}$-L446A-R455D-Δ68 and partial wild-type structures show that the buried N-terminus is involved in positioning Asp238 in the active site (Figure 2.9). The focus of this study was to probe the sequence requirements of this allosteric activation using various amino acid modifications of both the N-terminus and its binding pocket.

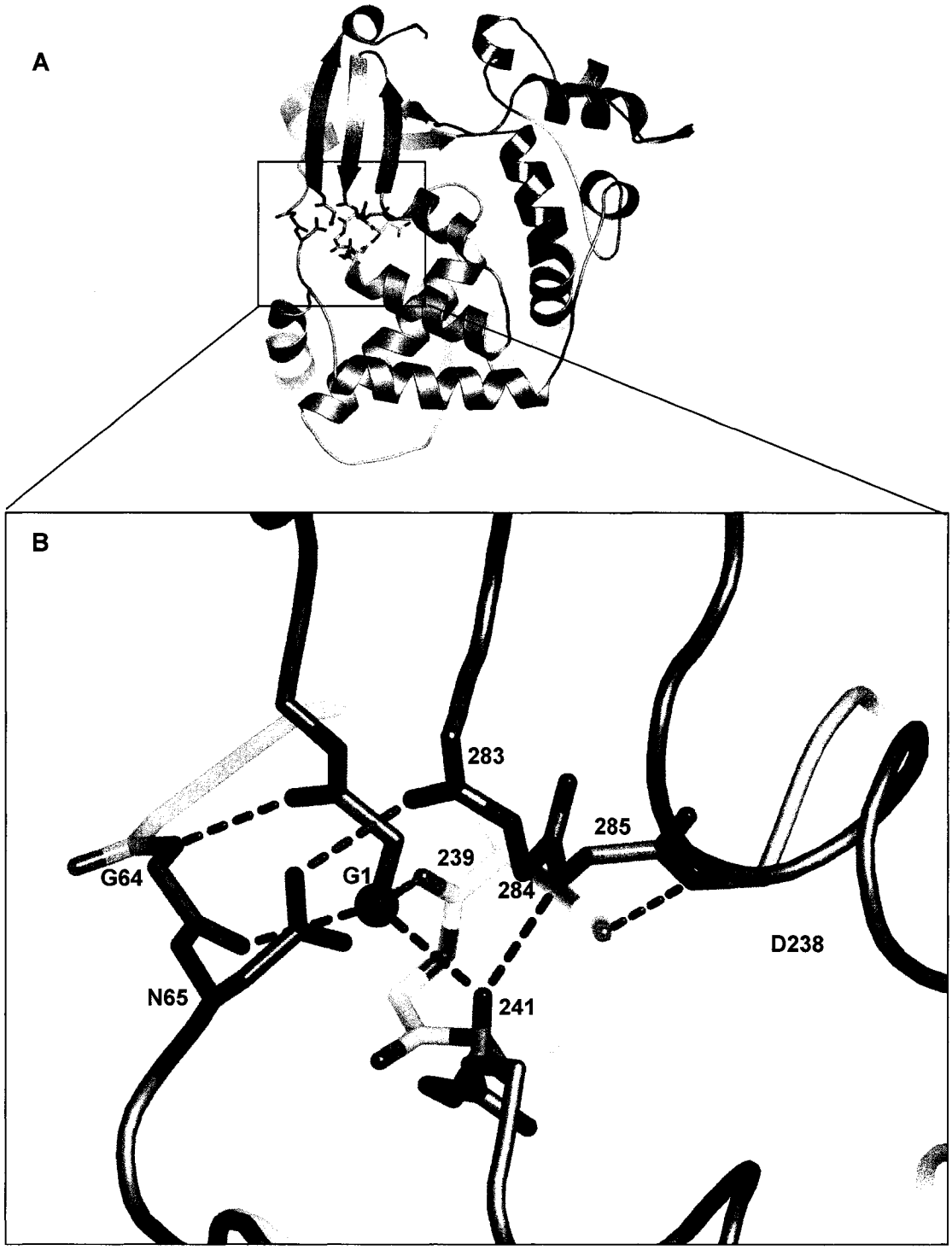


Figure 3.1 The N-terminus and its Binding Pocket
(**A**) Overview of the N-terminus and it's binding pocket colored according to structural motifs as in Fig. 2.7. (**B**) Close up view of the N-terminus and its binding pocket. Hydrogen bonds (Magenta) formed between the N-terminus and its binding pocket correctly position Asp238 into the active sie.

.

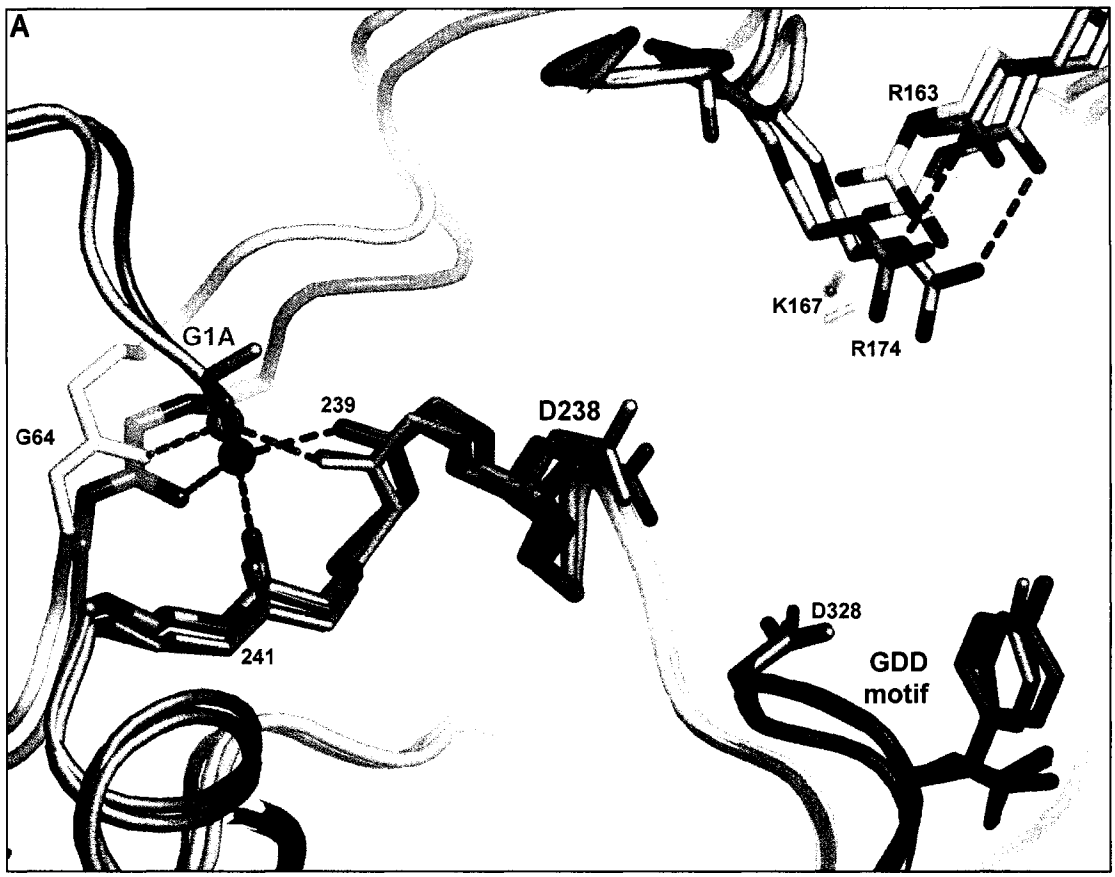
Results

Mutations of the N-terminus

To investigate the importance of the N-terminus for the structure and function of $3D^{pol}$, we made a series of mutations that added or removed residues from the protein or altered Gly1. The elongation activities of these mutants were tested using a poly-A/oligo-dT extension assay to measure ^{32}P-UMP incorporation into a product RNA strand (see Appendix 1). We generated two mutants that added a single glutamine residue or a serine-glutamine dipeptide to the N-terminus. These correspond to the last two residues of $3C^{pro}$ and thus mimic the natural precursor junction sequence of $3CD^{pro}$. Both mutants had less than 0.2 % the activity of the full-length protein (Table 3-1), demonstrating that the addition of even a single residue completely abolishes enzymatic activity. Likewise, deletion of the first residue, Gly1, also inactivates the polymerase. We then made two more subtle mutations of Gly1 to alanine and serine. These mutations retained partial polymerase activity, with $3D^{pol}$-G1A displaying ~54 % activity and $3D^{pol}$-G1S having only 1.6 % activity. With the exception of $3D^{pol}$-G1A, none of the N-terminus mutants could be crystallized under the conditions used for the full-length protein (Figure 3.2A & Table 3-2). However, circular dichroism wavelength spectra of these mutants show that the secondary structure is not significantly perturbed (Figure 3.2B)

Protein	% Activity
$3D^{pol}$ WT	215 ± 7
$3D^{pol}$ L446D/R455D	100 ± 4
$3CD^{pro}$ WT	0.1 ± 0.2
Mutants below are in addition to L446D/R455D	
$3D^{pol}$ Δ68	0.0 ± 0.2
$3D^{pol}$ ΔG1	-0.1 ± 0.2
$3D^{pol}$ +Q	0.1 ± 0.2
$3D^{pol}$ +SQ	-0.4 ± 0.2
$3D^{pol}$ G1A	54 ± 2
$3D^{pol}$ G1S	1.6 ± 0.3
$3D^{pol}$ G64A	~25
$3D^{pol}$ G64S	~25
$3D^{pol}$ N65A	9.3 ± 1.2
$3D^{pol}$ A239G	59 ± 0.4
$3D^{pol}$ G284A	5.0 ± 0.2
$3D^{pol}$ G285A	131 ± 3
$3D^{pol}$ G284A-G285A	18 ± 0.2

Table 3-1 Polymerase Extension Activities of N-terminus and N-terminus Binding Pocket Mutants



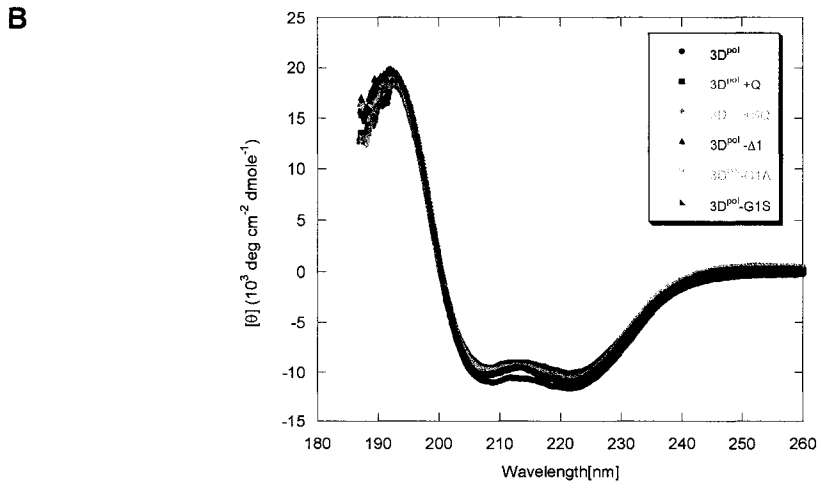


Figure 3.2 3D^{pol}G1A Structure Comparison
(**A**) View showing how the buried N-terminus of 3D^{pol} positions Asp238 for rNTP interactions. The N-terminus forms three hydrogen bonds with the carbonyl oxygens of residues 64, 239, 241 (magenta bonds) that act to position Asp238 for interaction with the 2' OH of rNTPs. The structures of the G1A mutant (orange), D238A mutant (teal, only residues 238-241 are shown), and the original partial structure without a buried N-terminus (red) are superimposed using the active site. (**B**) Wavelength CD spectra 3D^{pol} (all these constructs contain the L446D-R455D mutations) and the various N-terminal mutants showing that the secondary structure is not perturbed.

	$3D^{pol}$-G1A	$3D^{pol}$-G284A
PDB Code	1TQL	N/A
Space Group	$P6_5$	$P6_5$
Unit Cell (Å)	a=b=127.8 c=113.3	a=b=127.6 c=113.3
X-Ray Source	ALS 4.2.2 – 1.0 Å	Cu-Kα – 1.54 Å
Resolution Limits	30 – 2.30 (2.38 – 2.30)	30 – 2.40 (2.49 – 2.40)
Reflections		
Total Collected	279462	261890
Unique	46290	40581
Redundancy	6.0 (4.4)	6.45 (6.37)
I/σ [1]	13.8 (2.3)	11.8 (3.5)
Completeness (%)	99.2 (96.1)	100 (99.9)
R_{merge} (%)	7.4 (50.4)	7.7 (46.6)
Refinement		
Resolution Range	30 – 2.30	30 – 2.40
R	23.7	24.6
R-free (10 % of data)	26.2	27.7
Ramachandran statistics [2]		
Favored	373	369
Allowed	36	40
Generous	1	2
Disfavored	1	0

Table 3-2 Data collection and refinement statistics.
Data collection and refinement statistics for the $3D^{pol}$-G1A and $3D^{pol}$-G284A mutants. Data in parentheses are for the highest resolution shell, and Ramachandran statistics do not include glycine and proline residues

Mutations of N-terminus binding pocket

Several mutations of the N-terminal binding pocket were made to dissect its role in modulating the structure and activity of $3D^{pol}$. Lining the outside of the binding pocket are residues Gly64 and Asn65, which both interact with Gly1. Aaron J. Sholders has done extensive mutational analysis of Gly64 as a primary component of his dissertation, and therefore I will briefly report his findings of only two constructs, $3D^{pol}$-G64S and $3D^{pol}$-G64A. Both of these mutants displayed polymerase activity approximately ~25% that of the wild-type protein, and he has been able to obtain crystals of both although only the $3D^{pol}$-G64A is shown (Figure 3.3A) (Sholders, 2006). The side chain of asparagine 65 from the outside of the pocket appears to be hydrogen bonded to both the middle finger and to the amide nitrogen of the N-terminus. To determine the importance of these interactions, Asn65 was mutated to an alanine by Grace Campagnola. The activity of this mutant was only ~10 % of the wild-type protein, and she was able to obtain crystals and a structure (see figure 3.3B) (Sholders *et al.*, 2006).

Alanine 239 is an important residue in the binding pocket whose carbonyl oxygen is directly hydrogen bonded to the N-terminus (Figures 3.1 & 2.9). Because we propose that the burial of the N-terminus and its interaction with the binding pocket is partially responsible for "pushing" Asp238 into proper position in the active site, we decided to mutate this residue to a flexible glycine residue to examine the effect. We predict that a flexible residue at this position will not fully drive Asp238 into the active site compared to an amino acid with a more rigid polypeptide backbone. The activity of this mutant was ~60% of the wild-

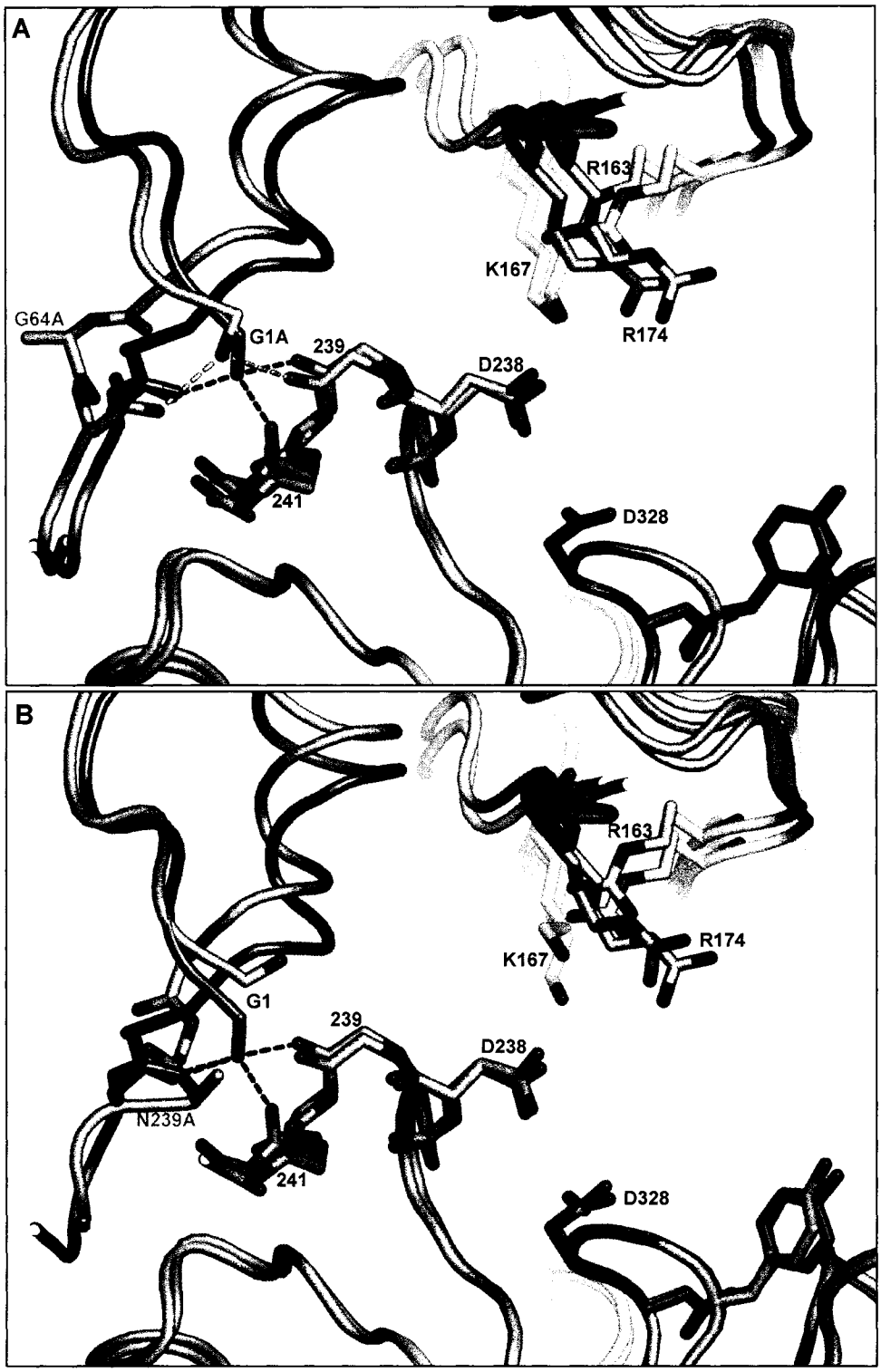


Figure 3.3 Mutant structures of the N-terminal Binding Pocket (outside) Compared with the Native Structure

(**A**) Structure of the $3D^{pol}$-G64A mutant colored orange and superimposed onto the active site (magenta) of native $3D^{pol}$ with its fingers and palm colored according to structural motifs described in Fig. 2.7. (**B**) Structure of $3D^{pol}$-N65A colored orange superimposed onto the active site (magenta) of native $3D^{pol}$ colored according to structural motifs as in Fig. 2.7.

type protein and Grace Campagnola was able to obtain crystals of this construct (Figure 3.4A).

Lining the inside of the pocket at the base of the middle finger are two glycines, 284 and 285, that are 100 % conserved in picornaviruses (Ann Palmenberg, personal communication; Hobson, 2000). These residues not only accommodate binding of the N-terminus, they also form important hydrogen bonds with residues from the bottom of the pocket including Asp238 (Figure 3.1 & 2.9). The single mutation of both Gly284 and Gly285 to alanines results in polymerases that are 5% and 131% active respectively. The hyperactivity of the $3D^{pol}$-G285A mutant is mitigated with the double mutant $3D^{pol}$-G284A-G285A resulting in 18% activity. Of these three constructs, only the $3D^{pol}$-G285A mutant crystallized (Figure 3.4B & Table 3-2).

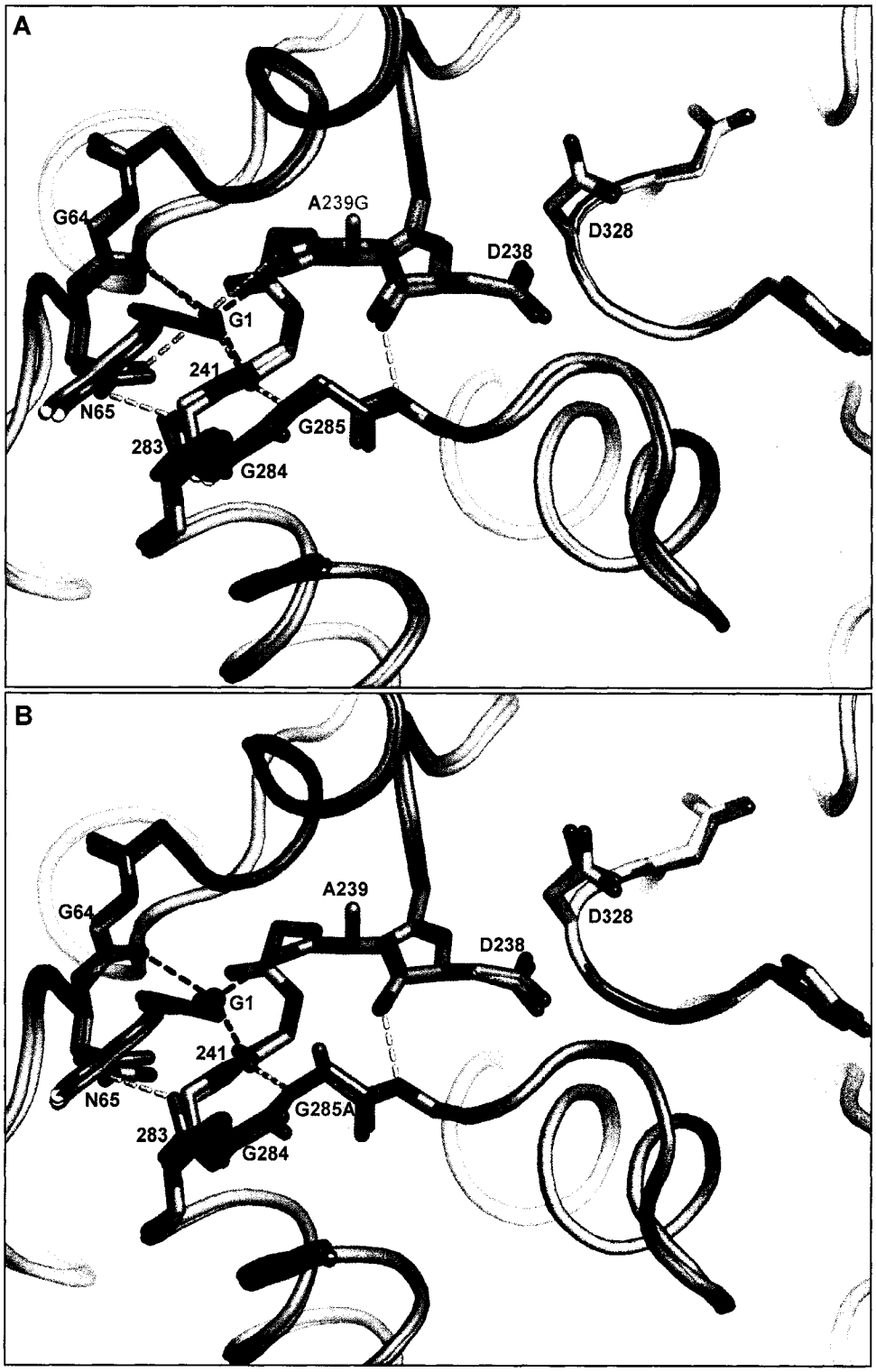


Figure 3.4 Mutant structures of the N-terminal Binding Pocket (inside) Compared with the Native Structure

(**A**) Structures of the $3D^{pol}$-A239G mutant colored orange superimposed onto the active site (magenta) of native $3D^{pol}$ with the fingers and plam colored according to structural motifs as in Fig. 2.7. (**B**) Structure of $3D^{pol}$-G285A colored orange superimposed onto the active site (magenta) of native $3D^{pol}$ colored according to structural motifs as in Fig. 2.7, and Ala285 is colored cyan.

Discussion

Mutations of the N-terminus

It is apparent that mutation of the very N-terminus has deleterious consequences on the proper functioning of $3D^{pol}$. In almost all mutants, the addition, deletion or modification of the N-terminus abolished the ability of this enzyme to catalyze the polymerization of nucleotides. One exception to this was the $3D^{pol}$-G1A mutant which retains ~54 % activity. This is not entirely surprising since an alanine, which contains only a methyl group as its side chain, is a relatively small amino acid making it possible for this residue to fit into the N-terminal binding pocket. Crystals of the $3D^{pol}$-G1A mutant grew much slower than those with native N-termini, resulting in a "sick"-looking morphology. The failure of the catalytically dead polymerases to crystallize and the "sick"-looking $3D^{pol}$-G1A crystal morphology is likely due to the presence of a major crystal lattice contact centered around this region.

$3D^{pol}$-G1A Structure

In the $3D^{pol}$-G1A structure, the side chain methyl group of the alanine mutation is packed into the area normally occupied by the Gly1 alpha carbon by expanding the outside of the N-terminus binding pocket. As a result, the amino terminus is pushed out of its binding pocket by ~0.9 Å and the hydrogen bond to the backbone carbonyl of residue 241 is lost. This has a direct effect on the positioning of Asp238, which is now only pushed about halfway into the active site as compared to the native Gly1 residue (see figure 3.2). Thus, there could be

a direct correlation between enzymatic activity (~54 %) and the positioning of Asp238 by the N-terminus.

Recovery Mutants

An unsuccessful attempt was made at restoring activity of the N-terminal insertion and deletion constructs by mutating aspartate 238 to a glutamate. Due to the presence of an extra methylene group, the carboxylate group of a glutamate side chain is extended approximately ~1.5 Å further from the backbone when compared to an aspartate side chain. Although we do not have the structures of the $3D^{pol}$-ΔG1, $3D^{pol}$ +Q, and $3D^{pol}$ +SQ mutants, the positioning of Asp238 should resemble the $3D^{pol}$-Δ68 and partial wild-type structures (1.4 Å movement away from catalytic center) (see figure 2.9). We therefore hypothesized that mutating Asp238 to a glutamate with a longer side may circumvent the requirement of the N-terminal dependent positioning of Asp238. Unfortunately, the $3D^{pol}$-ΔG1-D238E, $3D^{pol}$+Q-D238E, $3D^{pol}$+SQ-D238E and full-length $3D^{pol}$ D238E mutants displayed no polymerase activity (data not shown). Modeling in a glutamic acid with a library of its most commonly observed rotamers (using the program O; Jones *et al.*, 1991), suggests that the carboxylate group will not be correctly positioned even if it is extended an extra 1.5 Å from the polypeptide backbone. The absence of activity for the full-length $3D^{pol}$ D238E construct underscores the importance of a correctly positioned aspartate at residue 238. Likewise, an Asp238Ala mutation completely abolished activity and electron

density maps from crystals of this mutant soaked with rNTPs suggested a lack of nucleotide binding (nucleotide binding is discussed in Chapter V).

Mutations of the N-terminus Binding Pocket

The Outside of the Pocket

The outside of the N-terminus binding pocket consists of residues Gly64 and Asn65 from the index finger which are both poised to interact with the Gly1 of the N-terminus, thereby linking both ends of this finger to the middle finger (Figure 3.1). Gly64 makes a pair of reciprocal amide-carbonyl hydrogen bonds with Gly1, and is interesting because mutation to a serine leads to a polymerase with increased fidelity, allowing it to discriminate against the nucleoside analog ribavirin (Pfeiffer and Kirkegaard, 2003). Gly64 is highly conserved and lies in the disallowed region of the Ramachandran plot, thus requiring it to be a flexible residue. When this residue is mutated to an alanine or serine, the outside conformation of the binding pocket is extended by approximately 2 Å (only $3D^{pol}$-G64A is shown in Figure 3.3A; comparing the C_α atom of residue 64) (Sholders et al., 2006). Surprisingly, the conformation of the N-terminus and Asp238 does not look very different from the native structure, but since these structures are static snapshots, it is plausible that this region is dynamic during nucleotide polymerization. In other words, the altered structure of the binding pocket could affect potential movements of the N-terminus and Asp238, which may be required during nucleotide polymerization or translocation.

The activity and fidelity of most nucleic acid polymerases can be altered by the use of different divalent cations during catalysis (discussed in chapters I & V). Because an optimal geometry for catalysis likely exists, the most common explanation typically lies with the ability of the metal ions, with various ionic radii, to properly coordinate substrates for formation of a phosphodiester bond. Interestingly, the $3D^{pol}$-G1A and both of the glycine 64 mutants show a consistent shift of the index finger backbone from the outside of the N-terminus binding pocket and across the back of the polymerase where it interacts with and stabilizes the ring finger (Figures 3.2 & 3.3). This movement in turn affects the positioning of the ring finger over the active site, which contains several basic residues that position the nucleotide for catalysis. This small conformational change within the fingers domain could therefore present the incoming nucleotide in a less favorable orientation, thus accounting for the 50 - 75 % decrease in activity. Correct positioning of the triphosphate moiety of (ribo)nucleotides is absolutely essential for proper nucleic acid polymerization (Beese *et al.*, 1993; Cheetham and Steitz, 1999; Doublie *et al.*, 1998; Johnson *et al.*, 2003; Li *et al.*, 1998a, 1998b, 1999; Yin and Steitz, 2002). It is also very possible that mutating the outside of the N-terminus binding pocket affect potential conformational changes of the ring finger for repositioning of NTPs during catalysis.

The complete structure with the native N-terminal region shows that the Asn65 side chain is hydrogen bonded to the carbonyl of residue 283 at the base of the middle finger and is in hydrogen bonding distance with the very N-terminus (~3.0 Å). This conformation appears to anchor and stabilize the outside of the

binding pocket to the fingers domain. The $3D^{pol}$-N65A structure shows significant conformational differences of the binding pocket exterior (see figure 3.3B). In this structure, mutating residue 65 to an alanine causes this amino acid to protrude into the binding pocket by approximately 2.7 Å, resulting in displacement of the entire N-terminus. The methyl group of this alanine now occupies the space where the native N-terminal nitrogen resided. Although the density for the N-terminal nitrogen is poor, it is sufficient enough to model this residue with its alpha-carbon nearly 1.5 Å away from its location in the native structure. Additionally, the terminal ammonium group of Gly1 now points towards the back of the polymerase at an acetate ion that now sits in this small cavity. As in the $3D^{pol}$-G1A and G64 mutant structures, the conformation of the index finger is altered enough to slightly change the position of the ring finger while having little effect on Asp238. Although the fidelity of this mutant is not known, its polymerase extension activity is reduced by ~90%.

The Bottom of the Pocket

The reduced activity of the $3D^{pol}$-A239G mutant confirms the importance of a residue with a rigid backbone at this location (see figure 3.4A). As noted above, the backbone carbonyls of residues 239 and 241 form a pair of hydrogen bonds with Gly1 that act to pull on the polypeptide backbone of residues 239-242, thereby pushing Asp238 into the active site. The $3D^{pol}$-A239G structure does not show any significant conformational differences of this region when compared to the static picture of native structure. The isomorphous nature of these structures suggests that altered dynamics of this region may best explain the ~40% decrease

in polymerase activity. While Asp238 appears to be correctly pre-positioned when compared to the native structure, the potential to "push and pull" on Asp238 during nucleotide polymerization is likely affected when a glycine residue resides adjacent to this aspartate.

The Inside of the Pocket

Glycines 284 and 285, which line the inside of the binding pocket, are 100 % conserved among picornaviruses and are in the disallowed region of the Ramachandran plot, requiring them to be glycines. The small size of this amino acid may prevent steric clashes with both the N-terminus and the surrounding region, which is important since the amide nitrogens of residues 284 and 286 hydrogen bond with the backbone of residues 241 and Asp238. Assuming minimal changes to the polypeptide backbone, mutating Gly284 to an alanine would put the side chain methyl group in very close proximity with the Asn65 side-chain providing a possible explanation of why this mutant is almost completely dead. Unfortunately, $3D^{pol}$-G284A would not crystallize, precluding a conclusive structural explanation for the enzymatic inactivation of this mutant.

The Gly285Ala mutation interestingly resulted in hyperactivity at 131 % that of the native polymerase. This construct crystallized and the structure shows that the alanine side chain fits into the binding pocket without any steric clashes. The conformation of the binding pocket, including the N-terminus and aspartate 238, has not changed and so the cause of hyperactivity is unclear. Once again, altered dynamics may provide the best explanation. The hyperactivity of the

$3D^{pol}$-G285A is diminished when glycine 284 is also mutated to an alanine. Although the $3D^{pol}$-G284A-G285A mutant did not crystallize, the 82 % reduction in activity compared to the native polymerase is likely due to compensatory effects of an alanine at position 284.

Conclusions

The polymerase extension activity data and structures presented above show that the native N-terminus and its binding pocket are essential for optimal polymerase activity. Unfortunately, the N-terminal region forms a major crystal lattice contact, thereby precluding crystallization of mutants that show significantly reduced polymerase activity due to perturbations of this region. The mutants that successfully crystallized yielded crystals that grew much more slowly and with "sick"-looking morphologies when compared to the native crystals. Of all the N-terminal mutants, only the G1A construct with partial polymerase activity crystallized. This structure highlights the correlation between the burial of the native N-terminus, the positioning of Asp238 in the active site, and the enzymatic activity of $3D^{pol}$. We hypothesize that the absence of crystallization and polymerase activity of the other N-terminal mutants is due to significantly perturbed N-terminal regions, suggesting $3D^{pol}$ is very sensitive to modifications of Gly1. While the mechanism of allosteric activation provides ample selective pressure for conservation at this amino acid position, additional pressure is exerted by virtue of this site comprising a conserved Glu-Gly junction on the viral polyprotein that $3CD^{pro}$ recognizes for proteolytic processing.

The N-terminus binding pocket is also highly conserved, although not to the extent of the N-terminus. In picornaviruses, the glycine at residue 64 is over 75 % conserved, with a serine and even threonine occasionally found at this position. A flexible glycine appears to be required for the native conformation, activity, and fidelity of poliovirus $3D^{pol}$. An asparagine is ~90 % conserved at

residue 65 with an aspartate found in several viruses. Both structural and enzymatic data show that this residue is important for the conformation of the binding pocket exterior, a function that both aspartates and asparagines can apparently perform. Alanine 239 is ~70 % conserved among picornaviruses with small serine and even glycine residues found in the remaining viruses. While we found $3D^{pol}$-A239G moderately inhibited activity, it seems that small residues are conserved at this location, thereby preventing steric clashes with the N-terminus. Finally, the glycines at position 284 and 285 are probably 100 % conserved due to their flexibility and small size. Although its function is unknown, it should also be noted that the carbonyls of glycines 284 and 285 are involved in chelating a metal ion at the base of the thumb (in close proximity with the N-terminus). All mutations of the N-terminal binding pocket caused significant aberations in the activity and conformation of $3D^{pol}$. It is likely that similar interactions between the N-terminus and its binding pocket are found in other picornaviruses resulting in the allosteric activation of the polymerase.

There appears to be an emerging inverse correlation between processivity and fidelity. The Cameron lab has proposed that metals (Mn^{2+}) which increase the rate of phosphoryl transfer confer a decrease in fidelity due to the reduction in the life-time of the interrogation step (Arnold *et al.*, 2004; Castro *et al.*, 2005). While the generation of quasispecies from the high mutation frequency of animal viruses is essential for evolutionary fitness, an emerging model suggests that these viruses may be susceptible to "error catastrophe" where the increase in nucleotide misincorporation exceeds the ability to produce viable viral progeny (Crotty et al.,

2001; Crotty and Andino, 2002). This model likely explains why the identification of hyperactive polymerase mutants by our laboratory and others are not integrated into evolving poliovirus quasispecies. The $3D^{pol}$-G64S mutation, may cause the polymerase to replicate at a slower rate, but the higher fidelity enables escape from the selective pressure of the ribavirin nucleoside analogue.

Specific proteolytic processing is a common strategy employed by many classes of proteins to regulate proper activation of proteins at the desired time and location. The classic examples are digestive enzymes such as trypsin and chymotrypsin that are initially released as inactive proenzymes whose cleavage results in a slight refolding of the active site to yield the active enzyme. Such allosteric effects are also common among RNA viruses that have evolved mechanisms to take advantage of variable proteolytic processing to regulate many aspects of the viral life cycle. In poliovirus, the $3CD^{pro}$ protein plays an integral role in orchestrating viral RNA replication by binding to both ends of the viral RNA genome in the context of large membrane-associated replication centers (Murray and Barton, 2003). This has the effect of localizing the inactive polymerase precursor to the sites of negative- and positive-strand RNA synthesis, where it is poised to initiate RNA replication after cleavage to yield the mature $3D^{pol}$ protein. These structures and biochemical data of poliovirus $3D^{pol}$ show the molecular details of the switch responsible for the complete allosteric activation of a RNA polymerase, and provide insight into the structural conservation of picornaviral polymerases.

CHAPTER IV

REVIEW OF RNA-DEPENDENT RNA POLYMERASE STRUCTURES

Since publication of the first partial RNA-dependent RNA polymerase (RdRP) crystal structure from poliovirus, the coordinates of at least eight other viral polymerases from this class of enzyme have been released. These include three serotypes from human rhinovirus (HRV) (Love *et al.*, 2004) and foot-and-mouth disease virus (FMDV) (Ferrer-Orta *et al.*, 2004), which are all in the *Picornaviridae* family with poliovirus (PV); two members of the *Caliciviridae* family including rabbit hemorrhagic disease virus (RHDV) (Ng *et al.*, 2002) and Norwalk virus (NV) (Ng *et al.*, 2004); two strains of the hepatitis C virus (HCV) NS5B polymerase (Ago *et al.*, 1999; Bressanelli *et al.*, 1999; Lesburg *et al.*, 1999; O'Farrell *et al.*, 2003) and the RdRP from bovine viral diarrhea virus (BVDV) (Choi *et al.*, 2004) representing members of the *Flaviviridae* family; and the double-stranded RNA bacteriophage ϕ6 (Butcher and Grimes *et al.*, 2001) and reovirus λ3 (Tao *et al.*, 2002). In addition, the picornaviral coxsackie virus polymerase structure is currently being solved by the Peersen laboratory. This chapter consists of a review that compares and contrasts the structures of these polymerases with the complete poliovirus $3D^{pol}$ structure.

OVERALL RDRP ARCHITECTURE

The domain structure of these viral RdRPs follows the usual right hand analogy consisting of a thumb, palm, and fingers domain that was first used to describe the structure of the DNA polymerase I Klenow fragment (Ollis *et al*, 1985). An emerging structural feature of these polymerases is the presence of an "enclosed" active site characterized by extensive interactions between the fingers and thumb domains, resulting in a large tunnel through the molecule (see figure 4.1). This is in contrast with DNA polymerases and DNA-templated RNA polymerase structures (see figure 4.2), which do not show interdomain interactions between the fingers and thumb domains, thereby permitting large ratchet-like conformational changes during nucleotide polymerization. In chapter VII, we show that the enclosed RdRP structure is important for the thermal stability of PV polymerase. The tunnel that is created forms a positively charged funnel on the "back" of the enzyme for the entry of ribonucleotides and exit of pyrophosphate.

There are several structures available of the DNA-dependent RNA polymerase from bacteriophage T7 in complex with a number of nucleic acid templates and template-product pairs. A superposition of the $3D^{pol}$ active site with the T7 polymerase active site in that enzyme's elongation complex structure (Yin and Steitz, 2002) shows that the template strand of the T7 complex interacts with a groove formed by the fingers domain after roughly threading through a cleft at the "top" of this domain (see figure 4.3A). Nucleic acid binding with the fingers domain is also observed with many of the other well-studied polymerases

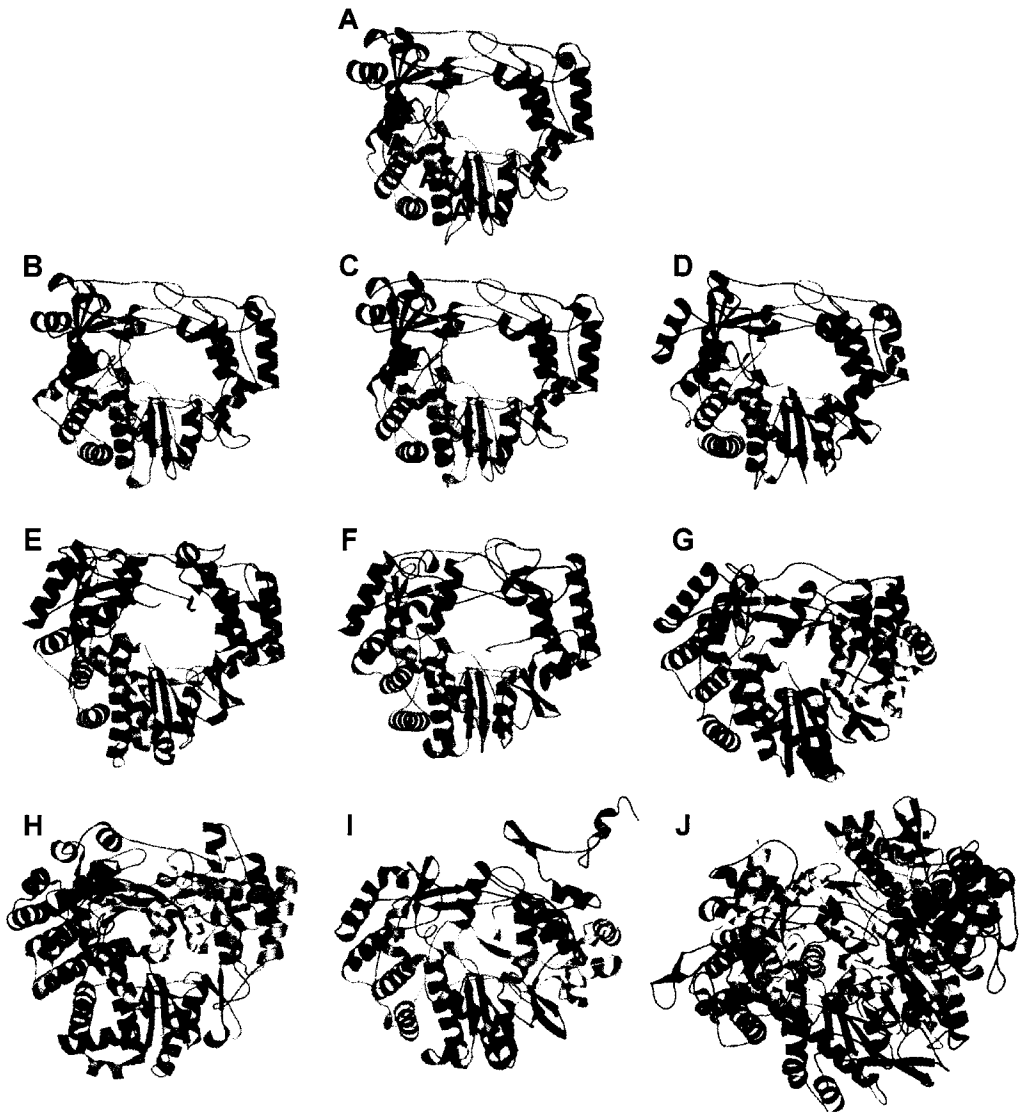


Figure 4.1 Gallery of viral RNA-Dependent RNA Polymerases
Front view of polymerase structures from (**A**) poliovirus (Thompson and Peersen, 2004), (**B**) coxsackie virus (O. Peersen, *personal communication*), (**C**) rhinovirus (Love *et al.*, 2004), (**D**) foot-and-mouth disease virus (Ferrer-Orta *et al.*, 2004), (**E**) rabbit hemorrhagic disease virus (Ng *et al.*, 2002), (**F**) Norwalk virus (Ng *et al.*, 2004), (**G**) hepatitis C virus (Ago *et al.*, 1999; Bressanelli *et al.*, 1999; Lesburg *et al.*, 1999; O'Farrell *et al.*, 2003), (**H**) bacteriophage ϕ6 (Butcher and Grimes *et al.*, 2001), (**I**) BVDV (Choi *et al.*, 2004), and (**J**) reovirus (Tao *et al.*, 2002). The fingers, palm, thumb domains are colored red, grey and blue, respectively and the catalytic centers are colored magenta. C-terminal priming platforms are colored yellow and extra N-terminal regions are colored green.

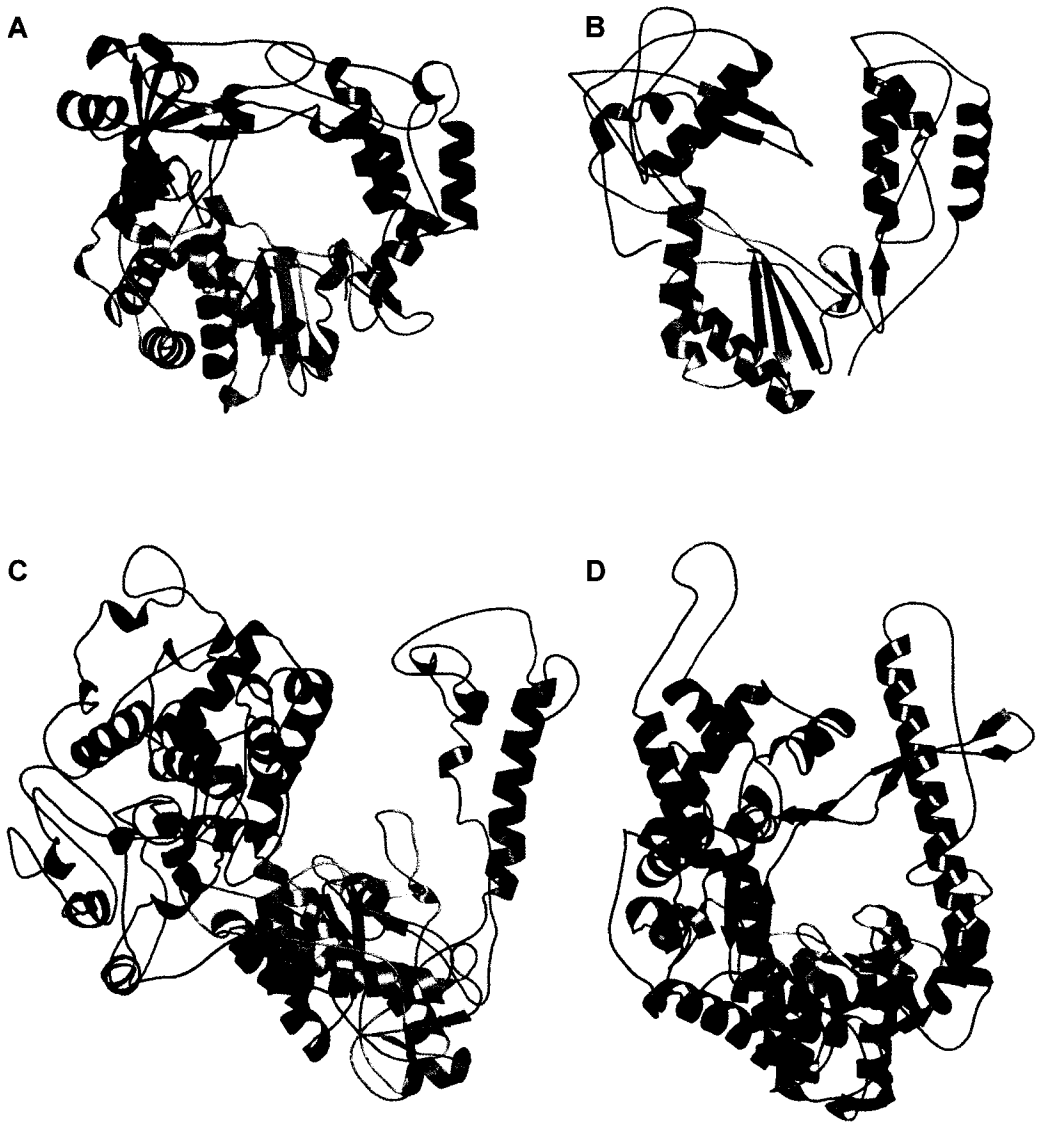


Figure 4.2 The Four Different Polymerase Classes
Front view of (**A**) poliovirus RNA-dependent RNA polymerase, (**B**) HIV-1 reverse transcriptase RNA-dependent DNA polymerase domain (1RTD; Huang *et al.*, 1998), (**C**) Taq DNA-dependent DNA polymerase (1BGX; Murali *et al.*, 1998), (**D**) bacteriophage T7 DNA-dependent RNA polymerase (1CEZ; Cheetham *et al.*, 1999). The fingers, palm, thumb domains are colored red, grey and blue, respectively.

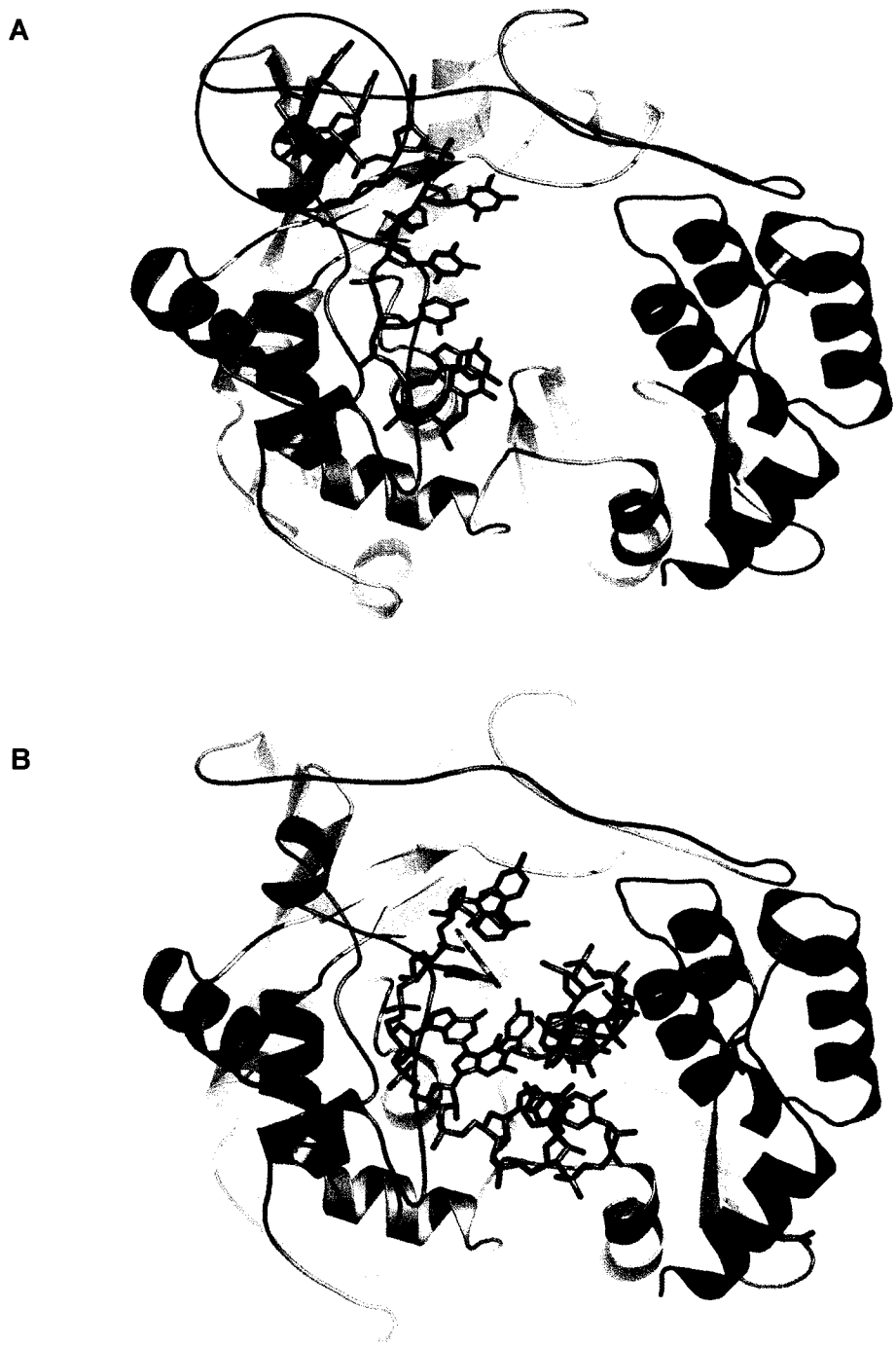


Figure 4.3 Models of Nucleic Acid bound to Poliovirus 3D^{pol}
(**A**) Position of the DNA template strand from the T7 RNA polymerase elongation complex (1MSW; Yin and Steitz, 2002) after active site superimposition. 3D^{pol} is colored according to structural motifs described in Fig. 2.7. The path of the template collides with a small helix of the pinky finger near the putative template entry channel. (**B**) Position of the duplex nucleic acid from superposition of the FMDV-RNA complex (Ferrer-Orta *et al.*, 2004) onto PV 3D^{pol}.

including HIV reverse transcriptase. There are three RdRP structures complexed with nucleic acid, all showing a similar template binding to the groove in the fingers domain and crossing the front of the enzyme. The HCV (genotype 1b, strain J4) (O'Farrell *et al.*, 2003) and bacteriophage ϕ6 polymerases (Butcher and Grimes *et al.*, 2001) are both complexed with single-stranded pentanucleotides whereas the FMDV polymerase is complexed with a template-primer pair composed of decanucleotide able to form a 6-base pair duplex flanked by 4-nucleotide 5' overhangs (Ferrer-Orta *et al.*, 2004). While there is some question as to the quality of the FMDV polymerase structure complexed with nucleic acid, superposition of the PV $3D^{pol}$ coordinates onto this complex yields insight into its binding with duplex nucleic acid (see figure 4.3B). Extension of the template 3' end reveals close contact with the R455D mutation on the thumb. Mutation of the positively charged arginine to a negatively charged aspartate likely accounts for the 55 % loss in activity we observed with the $3D^{pol}$-L446D-R455D mutant. In agreement with this, Schultz and colleagues saw that the single mutations of arginine 455 and 456 to aspartates were significantly more deleterious for enzymatic activity and RNA binding compared to the mutation of leucine 446 to an alanine, although all three mutant viruses failed to form plaques in tissue culture analysis (Hobsen *et al.*, 2001).

Due to the high degree of structural homology, the template-primer from FMDV can easily be accomodated into both the picornaviral and caliciviral polymerase structures. The architecture of the highly conserved palm, which contains the active site, and the four-helix bundle of the thumb domains of these

polymerases are almost indistinguishable. However, the finger domains do show small variations of the same theme, particularly with the topology of an α-helix located on the outside of the pinky (see figure 4.1). The axis of this α-helix from the FMDV polymerase is perpendicular to that in the other picornaviral structures from poliovirus and human rhinovirus; however, it is consistent with the architecture of the caliciviruses and flaviviruses. An interesting paradox is observed when the template-primer pair is modeled onto the bacteriophage ϕ6, reovirus λ3, and flaviviral polymerase structures (Butcher and Grimes et al., 2001; Tao et al., 2002; Ago *et al.*, 1999; Bressanelli *et al.*, 1999; Lesburg *et al.*, 1999; O'Farrell *et al.*, 2003). These enzymes have an extra C-terminal domain (colored yellow in figure 4.1) that covers the front of the thumb domain and occludes the active site, thereby blocking the path of the elongating RNA product.

THE C-TERMINAL PRIMING PLATFORM

Comparisons of the various RdRP structures and their biochemical properties reveal that the extra globular C-terminal domain may facilitate *de novo* (primer-independent) initiation. The picornaviral and caliciviral polymerases do not have the C-terminal priming platform, and as a consequence, they require either oligonucleotide or protein primers for initiation of replication. The more open structure of these enzymes permits the accessibility of protein primers (VPg) (Appleby *et al.*, 2005), which is the exclusive mechanism of initiation used by picornaviral polymerases *in vivo (*Paul, 2002). While there is yeast two-hybrid data indicating VPg binding to the "back" side of the polymerase (Lyle et al.,

2002b), a new report presents the structure of FMDV polymerase complexed with VPg binding to the "front" of the enzyme through interactions with the RNA binding groove and the thumb (coordinates not yet released) (Ferrer-Orta *et al.*, 2006b).

The correlation between *de novo* initiation and the presence of a large C-terminal thumb extension was proposed in the report describing the bacteriophage ϕ6 polymerase structure (Butcher and Grimes *et al.*, 2001). These authors suggest that Tyr630 from this domain, which is conserved as a tyrosine or tryptophan in this family of viruses, provides a platform on which an initiation complex can be constructed via π-stacking interactions with a nucleotide situated at a priming site (site P). A second NTP at the catalytic site (site C) can then form a phosphodiester bond with the 3' OH of the priming nucleotide. Continued nucleotide polymerization either results in the production of small abortive RNA products or a conformational change of the priming platform that permits elongation of the nascent RNA and template. Similar to the ϕ6 priming platform, the N-terminal domain of bacteriophage T7 initiation complex blocks the path of elongating RNA transcript and also requires a significant conformational change to facilitate the transformation from initiation to elongation (see figure 4.4) (Cheetham and Steitz, 1999; Cheetham et al., 1999; Tahirov et al., 2002; Yin and Steitz, 2002).

Similar structural molecular mechanisms facilitating *de novo* initiation have since been described with several other RdRP structures. The λ3

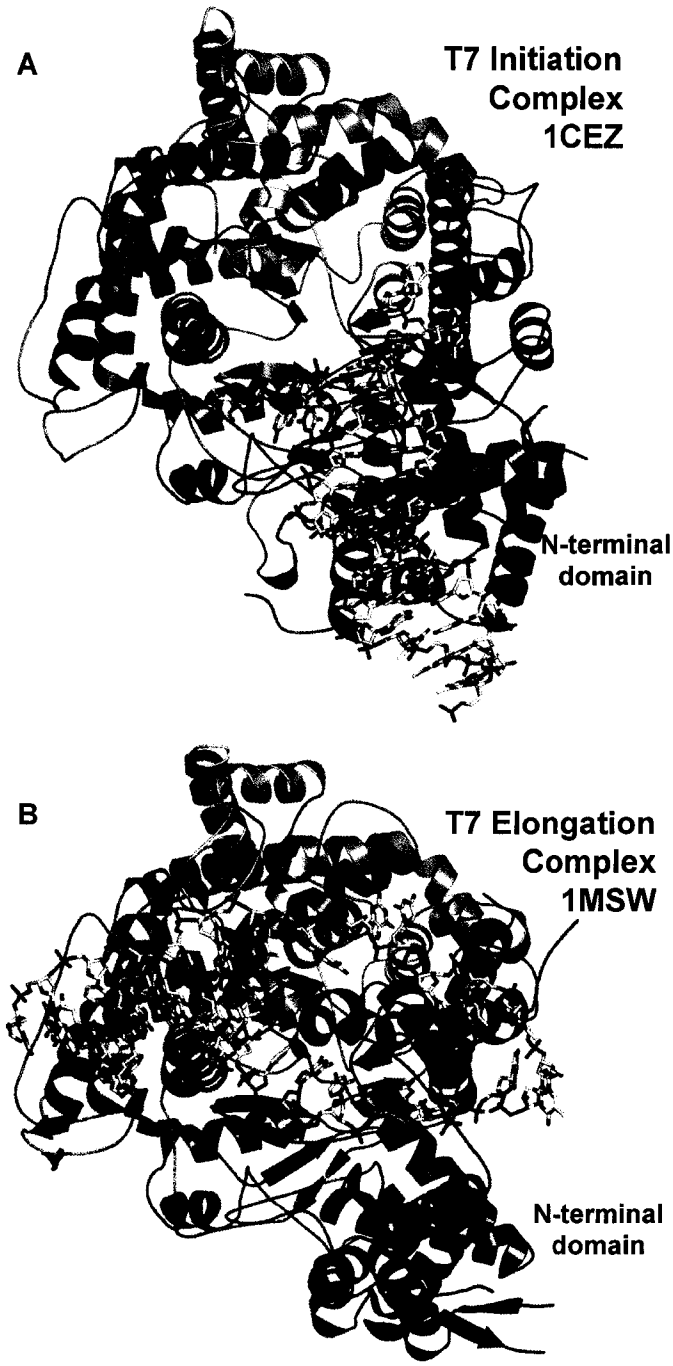


Figure 4.4 Bacteriophage T7 Polymerase Conformational Change from Initiation to Elongation

Bacteriophage T7 polymerase colored according the homologous $3D^{pol}$ structural motifs from figure 2.7, and the N-terminal domain is colored brown. Refolding of the N-terminal domain (brown) of the (**A**) initiation complex (1CEZ; Cheetham *et al.*, 1999) forms a channel that permits egress of RNA from the active site in the (**B**) elongation complex (1MSW; Yin and Steitz, 2002). The core polymerase structures do not significantly change conformation and superposition of the active site was used to generate these fixed views.

polymerase from reovirus has a unique priming loop that protrudes from the palm domain (Tao et al, 2002). After initiation, the loop retracts towards the palm to fit into the minor groove of the product duplex. Although the flaviviral polymerases are typically truncated at the C-terminus to increase the solubility, these structures show the presence of an extra domain extending from the thumb domains that occlude the active site. The priming platform of HCV has a β-hairpin (residues 443-454) that supports the stacking of initiating ribonucleotides, and when this region is deleted from this polymerase and the analogous residues from bacteriophage phi6, these enzymes require primers and are unable to undergo *de novo* replication (Hong et al, 2001; Laurila et al, 2002). Researchers working with the flaviviral dengue polymerase found that *de novo* initiation is inhibited at higher temperatures, whereas elongation phase is relatively thermostable at physiological temperatures (Ackermann and Padmanabhan, 2001). Although there is no published dengue polymerase structure, the authors of this study suggested that this RdRP can exist in a closed or open state, and the latter is favored at higher temperatures thereby precluding *de novo* initiation. The high B-factors of these platforms and potential hinge points support the notion that conformational reorganization of these platforms are required for elongation. Recently, a report presenting the proposed "open" and "closed" structures of the HCV NS5B polymerase was published (see figure 4.5) (Biswal *et al.*, 2005). Curiously, in both of these structures, the priming platform has not changed conformation to permit egress of the product RNA. Although superposition of these structures shows a minor rigid body rotation of the thumb domain that are

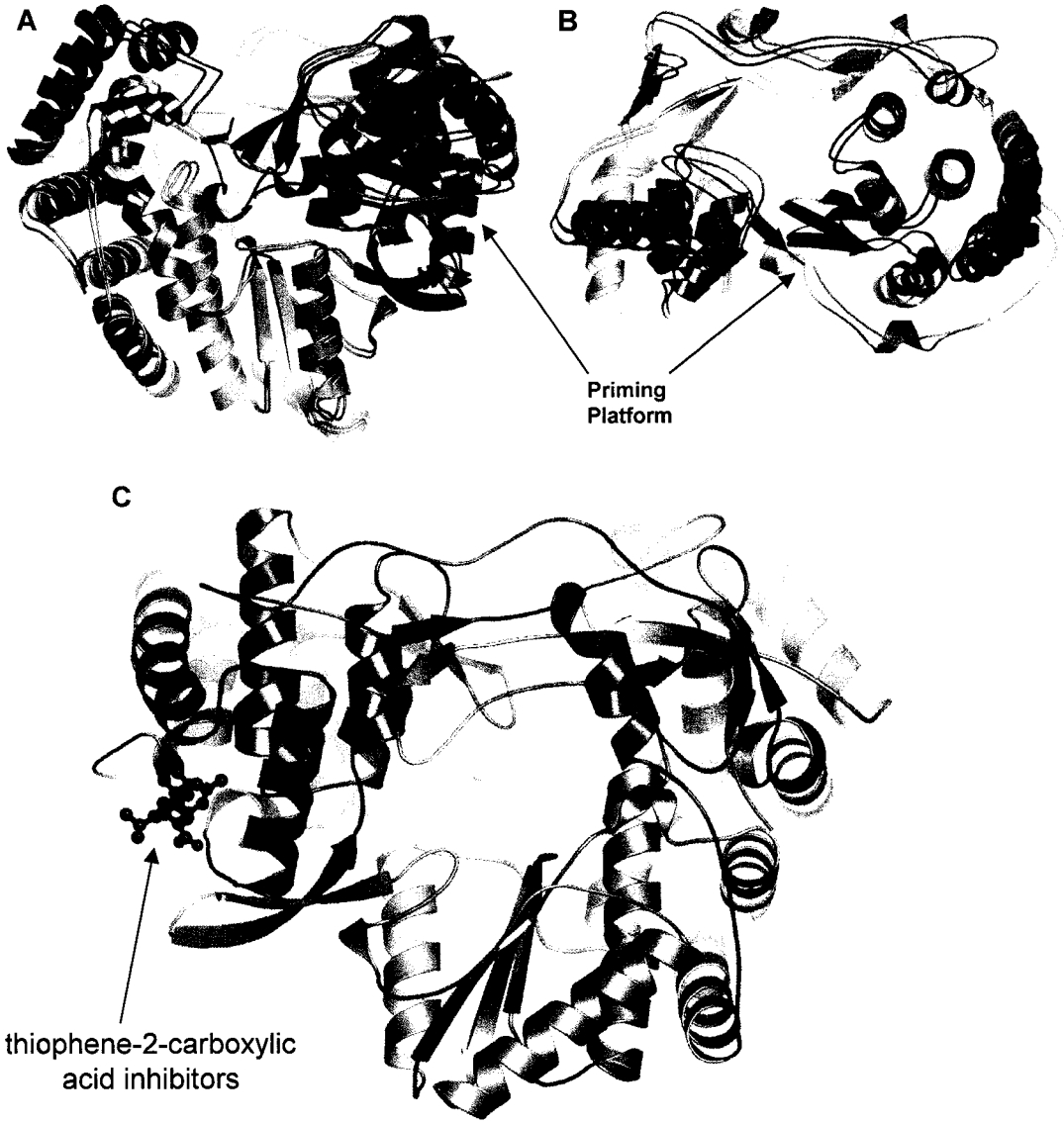


Figure 4.5 HCV NS5B polymerase
(**A**) Front view of HCV NS5B polymerase structures with active site superpositions reported to represent the open and closed conformation of this enzyme (Biswal *et al.*, 2005). However, it should be noted that the priming platform (brown) occludes the active site in both structures. The "core" polymerase is colored according to the structural motifs as in figure 2.7. (**B**) Top view of panel (A). (**C**) Back view of HCV NS5B polymerase structure complexed with a thiophene-2-carboxylic acid inhibitor bound to the thumb and priming platform (Biswal *et al.*, 2005).

likely due to crystal packing constraints, it is unlikely that either of these structures represent an elongation-competent structure, as the priming platform still occludes the active site. These authors also present structures of thiophene 2-carboxylic acid inhibitors bound to a region at the interface of the priming platform and thumb, which appear to function by preventing conformational changes of this region (see figure 4.5C) (Biswal et al., 2005).

It has been known that rGTP stimulates flaviviral polymerase *in vitro* activity (Lohmann *et al.*, 1999; Kao et al., 1999, 2001; Nomaguchi et al., 2003; Yi et al., 2003). However, recent evidence demonstrates that this ribonucleotide selectively stimulates *de novo* but represses primer-dependent initiation during *in vitro* RNA synthesis (Ranjith-Kumar et al, 2003). Co-crystallization of the HCV polymerase with rGTP indicated binding in the active site and 30 Å away on the surface of the thumb, in very close proximity to the index fingertip (Bressanelli et al, 2002) (Figure 4.6). Because the bound rGTP interacts with the index fingertip through direct interactions with Arg32 and several water-mediated hydrogen bonds, the binding of this ribonucleotide at this site was proposed to regulate RNA synthesis by modulating interactions between the fingers and thumb domains. However, subsequent mutational analysis of this binding site showed that *in vitro de novo* initiation is not affected, although these mutations either severely impaired or completely abolished the viability of subgenomic HCV RNAs *in vivo* (Cai et al, 2005). Perhaps the role of this GTP-binding site is to reinforce the interaction between the thumb and fingers domain for various steps of replication at the temperatures used in these assays (and *in vivo*). It should also

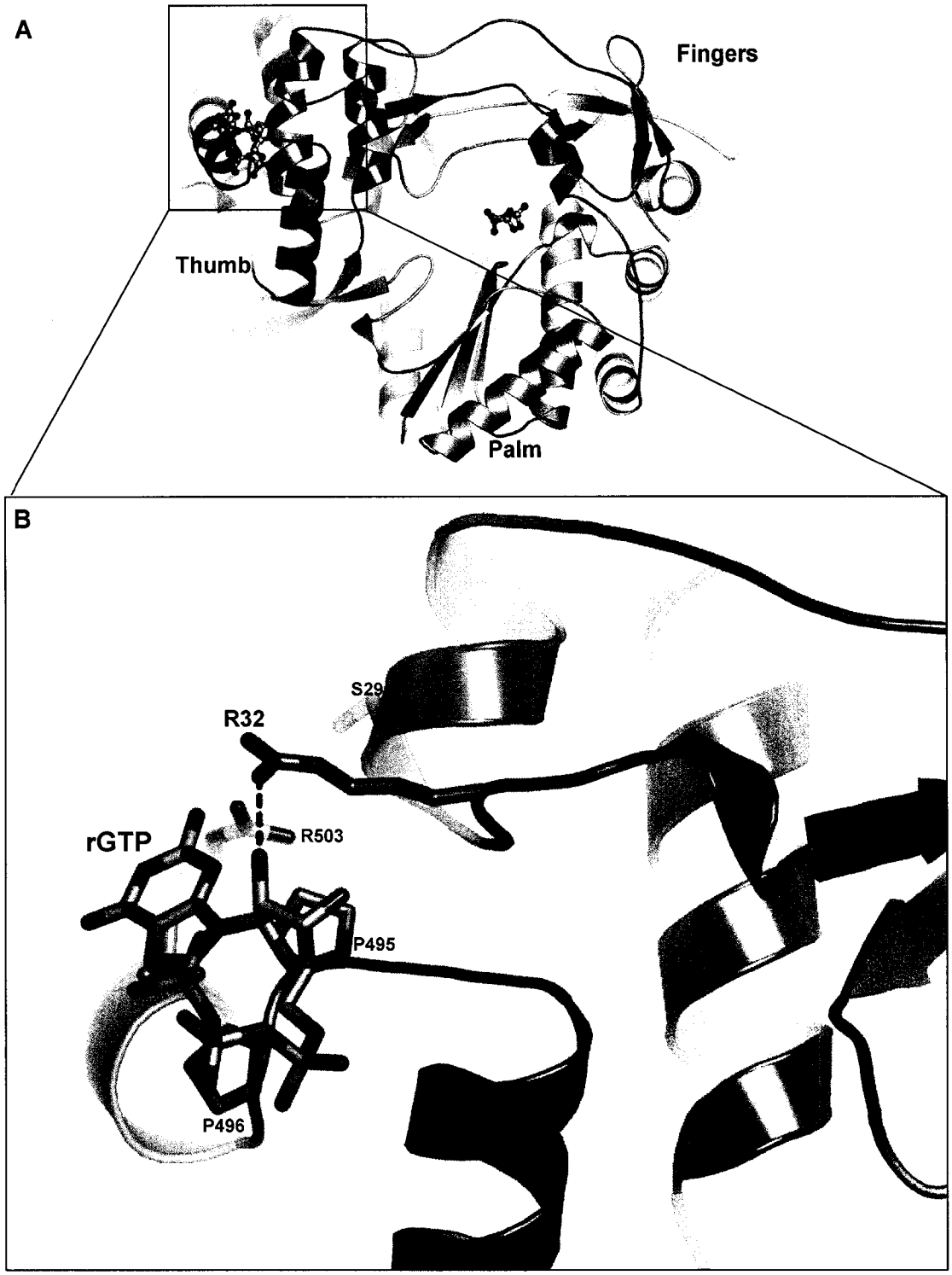


Figure 4.6 HCV NS5B complexed with rGTP
(**A**) Back view of HCV NS5B polymerase complexed with rGTP. The binding site is specific for rGTP and interacts with both the thumb-tip and index fingertip. (**B**) Close up view of panel (A) highlighting several of the interactions between the polymerase and GTP , as described by Bressanelli et al., 2002

be noted that the NS5B construct used for these ribonucleotide co-crystal structures has a 55-residue C-terminal truncation therefore removing the 21-residue membrane-anchoring region and the 34-amino-acid linker connecting the thumb to the membrane anchor. Structural comparison of this structure with constructs with only a 21-residue C-terminal truncation show differences in the conformation of the priming platform. It is possible that removal of the 34-amino-acid linker prevents proper folding of the priming platform, thereby precluding native *de novo* interactions with substrates.

The more recent structure of the bovine viral diarrhea viral polymerase complexed with rGTP has potentially established the role of this ribonucleotide in *de novo* initiation (Choi et al, 2004). This structure shows rGTP bound ~6 Å away from the catalytic center at the priming site (site P) through interactions with the priming platform, reminiscent of the bacteriophage ϕ6 polymerase. Only the rGTP ribonucleotide binds to this P site in the BVDV polymerase with interactions that appear to specify guanine binding. Mutation of the residues responsible for interacting with rGTP completely abolishes de novo synthesis, whereas primer-dependent initiation is reduced by only 2 - 10 fold (Lai et al., 1999).

N-TERMINAL REGIONS

Including PV $3D^{pol}$, the crystal structures of five picornaviral polymerases have been solved and they all show N-terminal glycines buried in shallow pockets on the back of the fingers domains (see figures 4.7 & 4.8). Although the

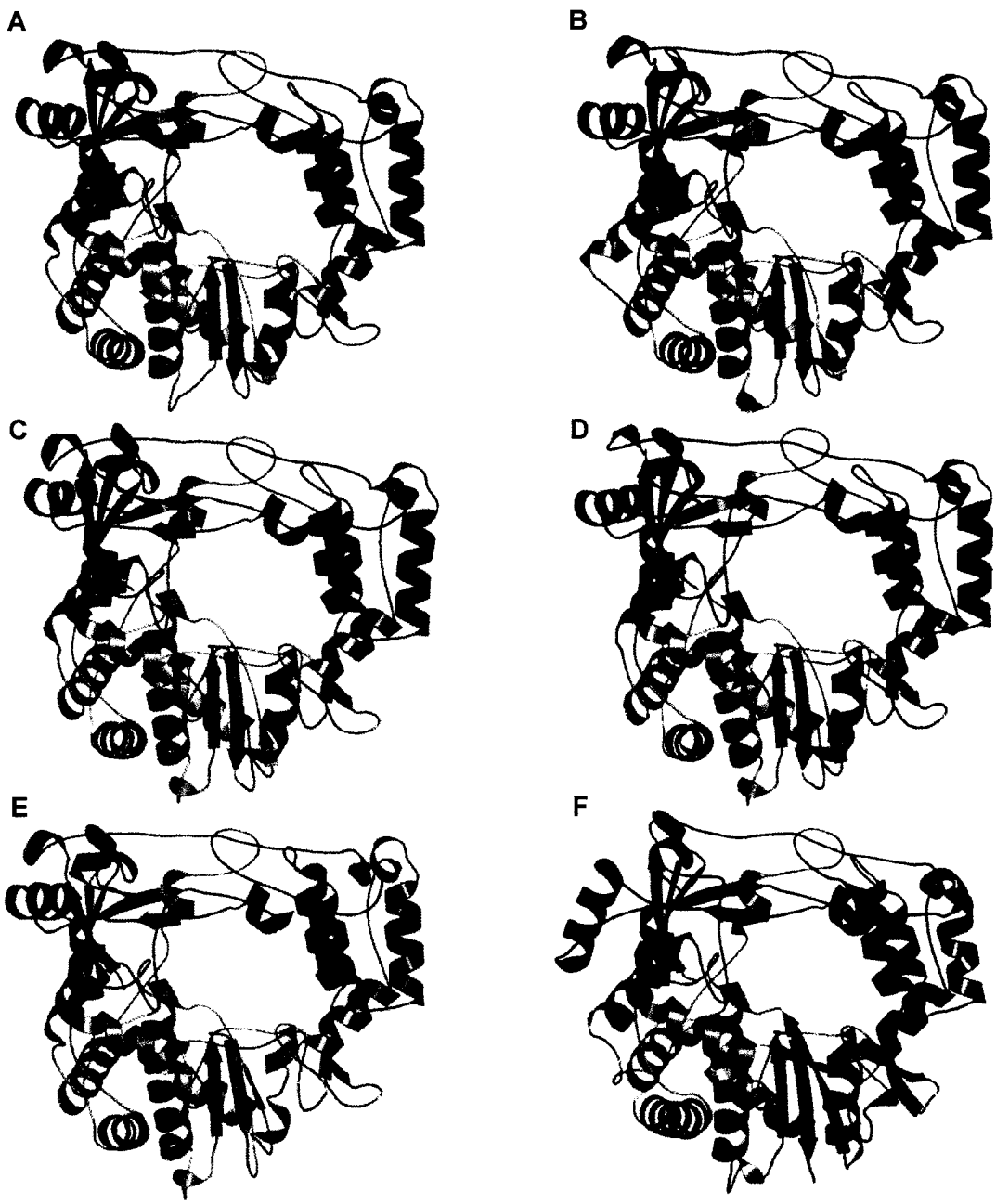


Figure 4.7 Polymerase Structures from the *Picornaviridae* Family
Front view of picornaviral polymerases colored according to the right hand analogy with the fingers, palm and thumb colored red, grey, and blue respectively. Included are the polymerase structures from (**A**) poliovirus (**B**) coxsackie virus (O. Peersen, *personal communication*), (**C**) rhinovirus (serotype 14), (**D**) rhinovirus (serotype 1b), (**E**) rhinovirus (serotype 16) (Love *et al.*, 2004), (**F**) foot-and-mouth disease virus (Ferrer-Orta *et al.*, 2004).

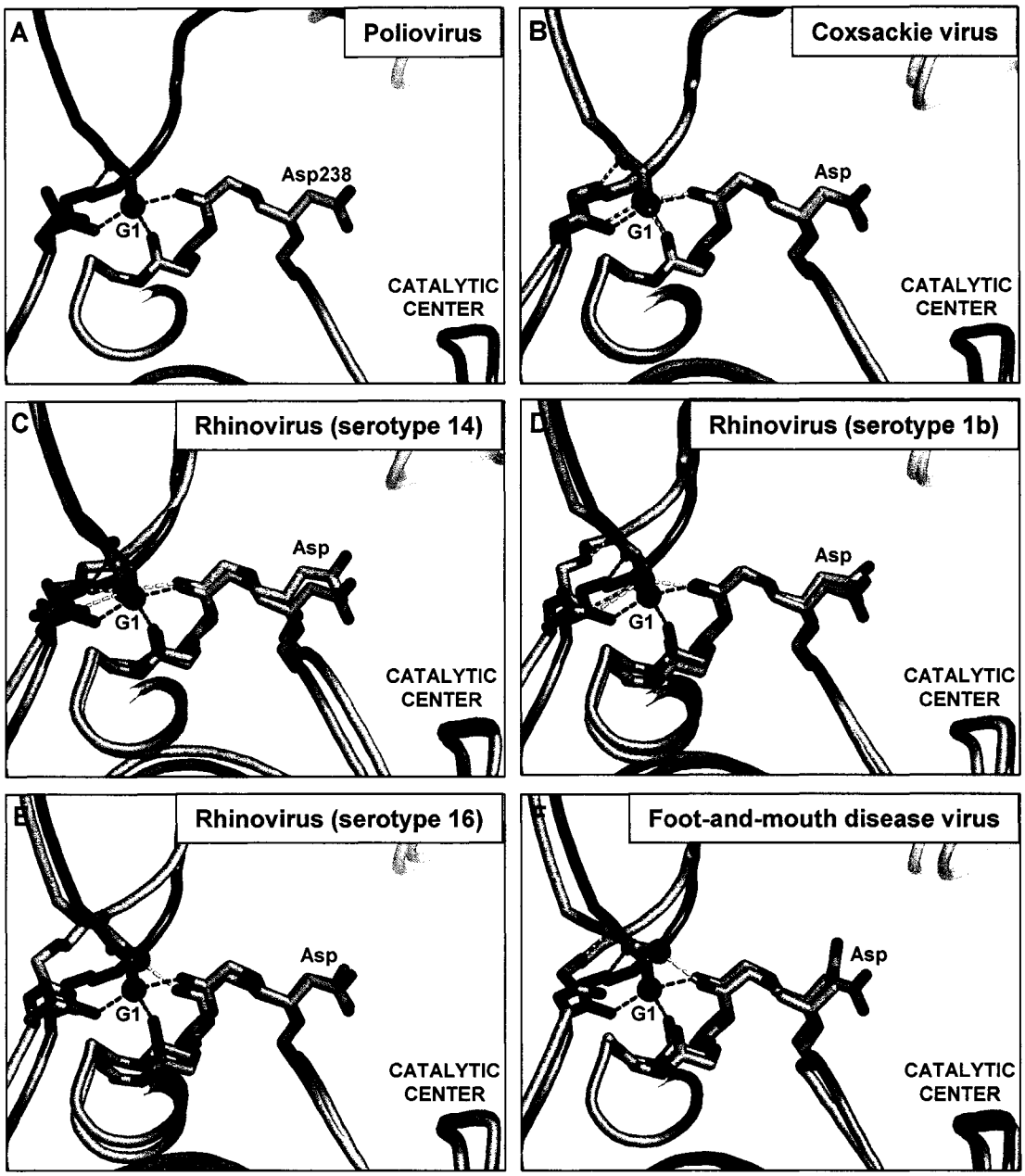


Figure 4.8 Structure Comparison of Picornaviral Polymerase N-termini
Structural comparison of picornaviral polymerase N-termini. (**A**) poliovirus polymerase is shown in all panels colored acording to the structural motifs described in Figure 2.7, and potential hydrogen bonds are colored magenta. The (**B**) coxsackie virus polymerase (O. Peersen, *personal communication*), (**C**) rhinovirus (serotype 14) polymerase, (**D**) rhinovirus (serotype 1b) polymerase, (**E**) rhinovirus (serotype 16) polymerase (Love *et al.*, 2004), and (**F**) foot-and-mouth disease virus polymerase (Ferrer-Orta *et al.*, 2004) are all colored orange with potential hydrogen bonds colored yellow.

sequences of these polymerases are quite different, the structural homology is excellent (see table 4). The N-terminus, its binding pocket, and the aspartate 238 (or homologous aspartate) are all conserved, clearly suggesting a common mechanism is involved with activating the polymerase following cleavage from its precursors (discussed in chapters II and III). While the FMDV and rhinovirus polymerase structures were published around the same time as that of poliovirus, no mention was made by these authors of potential allosteric positioning of the aspartate involved with ribonucleotide 2' hydroxyl interactions. However, Grace Campagnola and Mark Wygant from our laboratory (Peersen Lab) recently solved the coxsackie polymerase structure and were able to show that its N-terminus is indeed buried and that Gly1 is very sensitive to mutations. It is likely that similar molecular mechanisms for the allosteric activation of picornaviral polymerases exist to keep precursor proteins from functioning as polymerases.

Table 4 Picornaviral structure and sequence similarity

polymerase comparison with PV $3D^{pol}$	r.m.s.d.	seq. identity
coxsackie virus	0.7 Å	74.2 %
rhinovirus (serotype 14)	1.0 Å	64.3 %
rhinovirus (serotype 1b)	1.0 Å	57.5 %
rhinovirus (serotype 16)	1.1 Å	56.4 %
FMDV	1.8 Å	30.9 %

r.m.s.d. and seq. identity data obtained from combinatorial extension web server (Shindyalov and Bourne, 1998)

The reovirus and bovine viral diarrhea virus polymerase structures both have extra domains attached to the N-terminus (colored yellow in figure 4.1 and brown in figures 4.4 & 4.5). This domain from the reovirus polymerase, consisting of the first 380 amino acids, covers the back of the thumb and wraps around the top of the fingers domain. Although the authors suggest that these interactions anchor and augment the fingers-thumb connection, the function of

this domain remains unclear. The BVDV polymerase N-terminal domain consists of the first 138 amino acids, of which the first 70 were deleted and the next 71-91 were disordered. The remaining portion interacts with both the fingers and the top of the thumb in a similar fashion as the homologous domain from reovirus. A positively charged β-hairpin from the BVDV N-terminal domain partially occludes the template channel, suggesting this region may be involved in removing RNA secondary structure or in helicase activity (Choi *et al.*, 2004). Structural and bioinformatic analysis suggests that this region may be involved with helicase activity of all RdRPs (Bruenn, 2003). Telomerases are reverse transcriptases that have unique N-terminal domains that bind RNA (Weinrich *et al.*, 1997; Bryan *et al.*, 2000; Lai *et al.*, 2001). The RNA binding may function as an 'anchor site' that binds telomeric DNA upstream of the telomere 3' end, conferring the property of repeat-addition processivity on telomerase (Lue et al., 2005). However, it has also been proposed that the N-terminal regions of other polymerases may be involved in binding proteins required for *in vivo* viral replication (Shirako et al, 2000).

CHAPTER V

CRYSTALLIZATION AND ANALYSIS OF $3D^{POL}$ COMPLEXED WITH RIBONUCLEOTIDES

RNA replication by $3D^{pol}$ requires binding sites for the template as well as the incoming nucleoside triphosphate, and understanding the mechanisms in which these ligands are bound for catalysis aids in the development of polymerase specific anti-viral drugs. While we have been unable to obtain structures of $3D^{pol}$ complexed with template RNA, we have solved co-crystal structures of $3D^{pol}$ bound with nucleotides and divalent cations. These structures confirm the role of Asp238 in interacting with and selecting for the 2' hydroxyls of ribonucleotides and provide insight into the mechanism of catalysis.

Early studies on the poliovirus $3D^{pol}$ indicated a gradual loss of polymerase activity over time, which is accelerated by incubation at physiological temperatures. Interestingly, it was found that incubation of the enzyme with high concentrations of NTPs resulted in a dramatic protection against heat inactivation and activity loss. The minimum concentration of nucleotide that $3D^{pol}$ must be incubated with to provide thermal stabilization at 42°C is ~2 mM, with maximal protection occurring at concentrations around 5 mM (Richards *et al*, 1992). The Cameron lab found that the concentration of ATP sufficient to decrease the

observed rate of inactivation to 50 % of the maximum value ($K_{0.5(NTP)}$) exhibited a temperature dependence, ranging from 17 μM at 0°C to 84 μM at 30°C (Arnold & Cameron, 2000). Based on the structures of $3D^{pol}$-ribonucleotide complexes we have obtained, we believe this thermal protection is due to NTP-facilitated bridging interactions between the palm and fingers domains that stabilize the structure (Figure 5.1).

We performed biochemical activity assays to confirm this and found that divalent cations stimulate the thermal protection provided by ribonucleotide binding. However, there are currently no reports that quantify the structural effect of rNTP binding on protecting $3D^{pol}$ from thermal inactivation. An extensive study on the thermal stability of $3D^{pol}$ is discussed in chapter VII where we use circular dichroism to show that the melting temperature (T_m) for the full-length native polymerase is approximately 39°C. However, because of the high absorbance of ribonucleotides at the required concentrations, circular dichroism was unsuitable for quantifying any rNTP-mediated increase in the T_m. Rebecca Albertini is developing an 2-anilino-6-naphthalene sulfonate (ANS) binding assay that relies on the change in quantum yield of the fluorophore upon binding within hydrophobic areas. Exposure of hydrophobic cavities occurs concomitantly with partial unfolding of the protein, and because the excitation wavelength used for these studies (360 nm) is far from the absorbance maximum of rNTPs (~260 nm), she was able to quantitate the increase in thermal stability upon binding ribonucleotides.

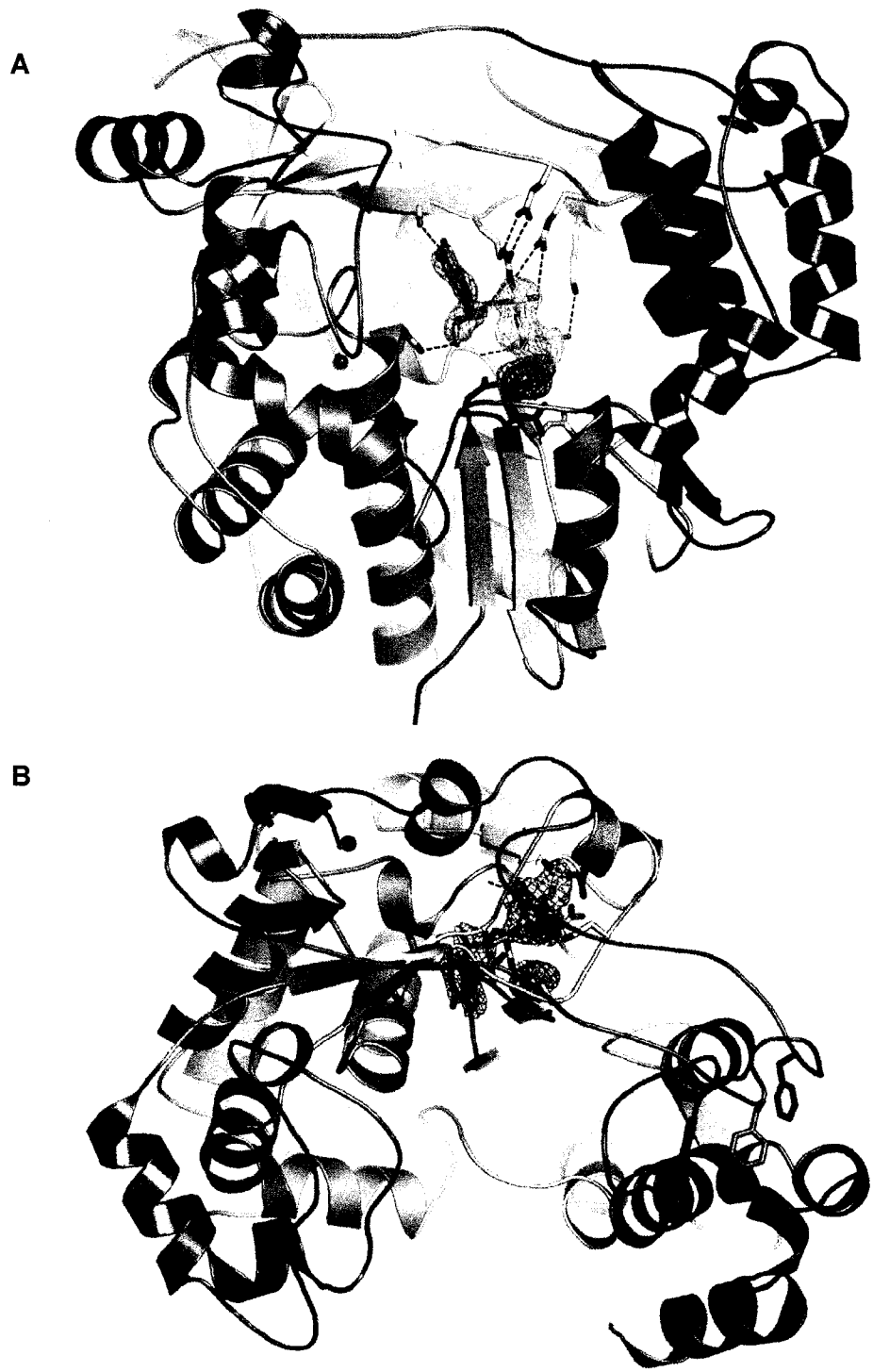


Figure 5.1 Overview of $3D^{pol}$ Structure Complexed with rGTP and Mn^{2+}
(**A**) Front view and (**B**) top view of $3D^{pol}$ complexed with rGTP and Mn^{2+} near the catalytic center which is colored magenta. The polymerase is colored according to the structural motifs described in Figure 2.7, and the residues involved with binding ribonucleotides are depicted with sticks. A 2.5 Å resolution 2Fo-Fc map (green) contoured at 1.6σ around rGTP and a 2.5 Å resolution anomalous map (blue; = 1.54 Å) contoured at 3.9σ around Mn^{2+}.

Results

Crystallization

Successful co-crystal structures of $3D^{pol}$ with all four rNTPs were obtained by either soaking crystals in 10 mM rNTP or by co-crystallization of the polymerase with a final rNTP concentration of 10 mM in the crystallization drop (Figures 5.2 & 5.3 & 5.4; Tables 5-1 & 5-2). We found that nucleotide is also required in the cryo-preservant as the ligand can readily diffuse out of its binding site before freezing. Electron density maps of $3D^{pol}$-rNTP co-crystal structures reveal superior occupancy for rGTP, moderate density for both rCTP and rUTP, and very poor density for rATP (Figure 5.3). This is consistent with reported K_d values, which may reflect the cellular availability of each ribonucleotide (Table 5-3). It should be noted that the GTP ribonucleotide had low solubility under these conditions, in agreement with observations from other laboratories. Crystallization experiments were thus performed in the presence of saturating concentrations of rGTP, and care was taken to remove any "globules" of insoluble rGTP prior to freezing. While these experiments were performed in the presence of either Mg^{+2} or Mn^{+2}, only electron density for the larger manganese cations was definitely observed (Figure 5.3). Quality rATP electron density maps have

Nucleotide	K_d (μM)	Cell conc. (μM)
ATP	134 ± 18	2102
UTP	98 ± 2	253
GTP	3.8 ± 0.7	468
CTP	19 ± 3	91

Table 5-3 Ribonucleotide K_ds and cellular concentrations
ribonucleotide binding constants from Arnold and Cameron, 2003 and cellular concentrations from Traut et al., 1994.

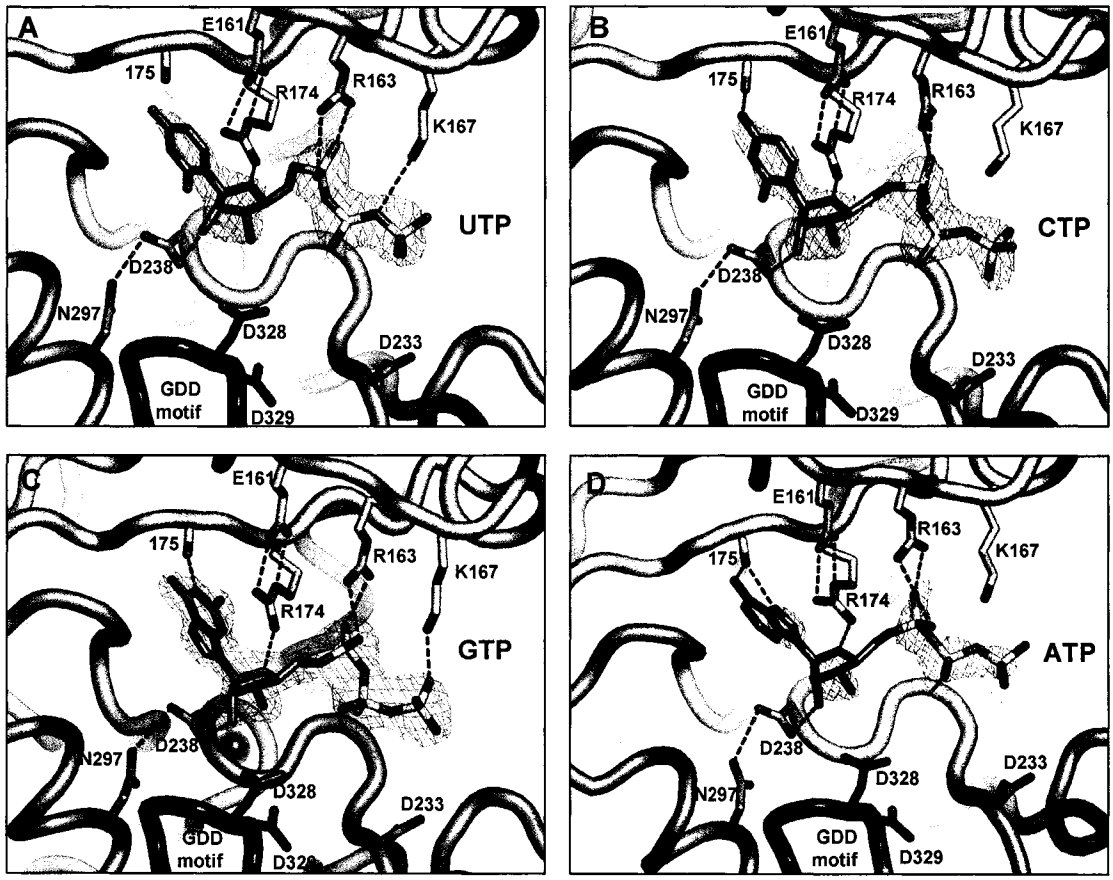


Figure 5.2 Structures of 3D^{pol} Complexed with rNTPs in presence of Mg^{2+}
Structures of 3D^{pol} complexed with the ribonucleotides in the interrogation site (preinsertion site) in presence of Mg^{2+} (**A**) 2.35 Å resolution 2Fo-Fc map contoured at 1.5σ around rUTP. (**B**) 2.25 Å resolution 2Fo-Fc map contoured at 1.5σ around rCTP. (**C**) 2.35 Å resolution 2Fo-Fc map contoured at 1.5σ around rGTP. (**D**) 2.6 Å resolution 2Fo-Fc map contoured at 1.5σ around rATP

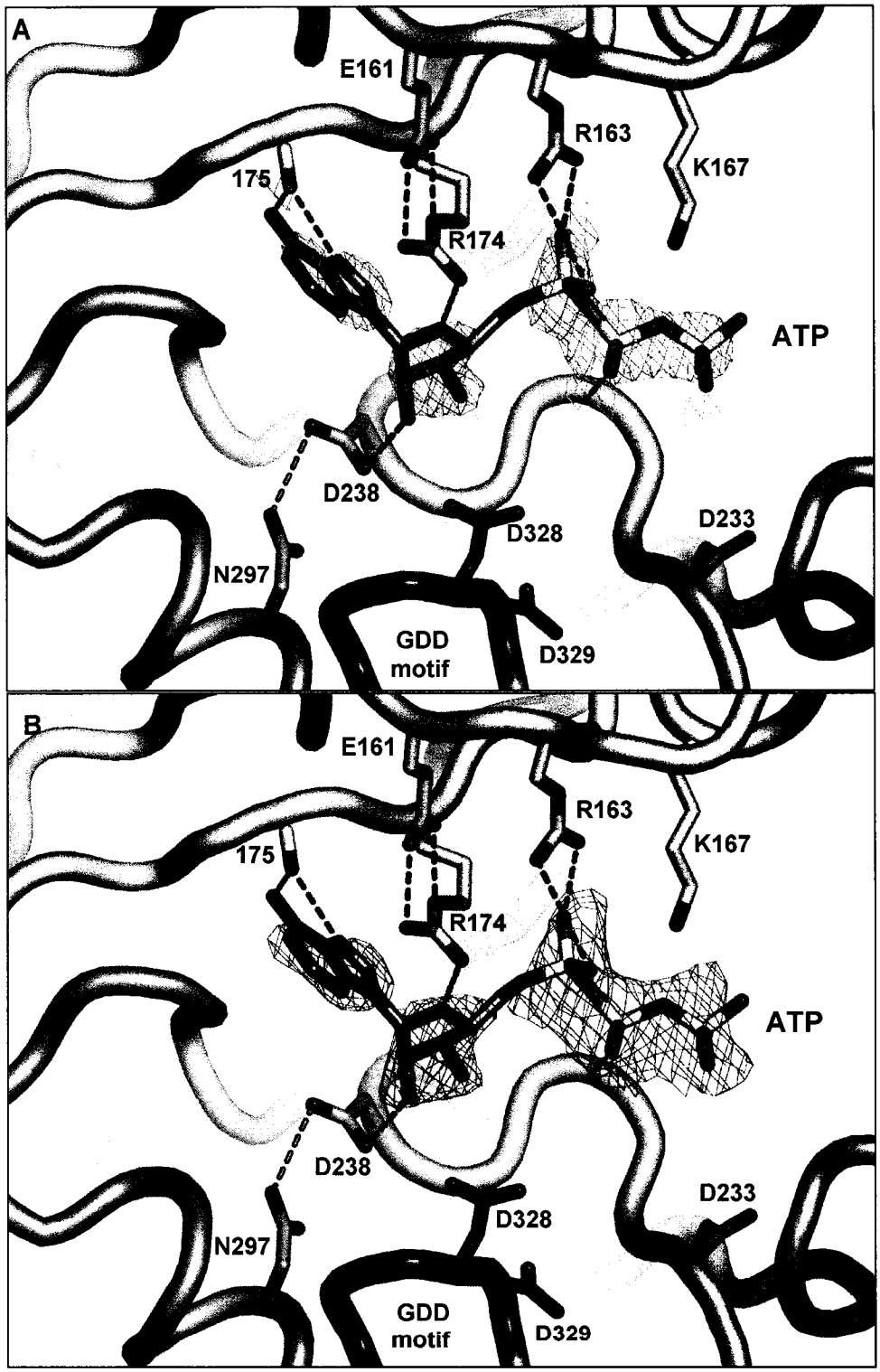


Figure 5.3 Structures of $3D^{pol}$ Complexed with ATP in the presence of Mg^{2+}
Structures of $3D^{pol}$ complexed with the ATP in the interrogation site (preinsertion site) in the presence of Mg^{2+}. (**A**) A 2.6 Å resolution 2Fo-Fc electron density map (green) contoured at 1.2σ, and a (**B**) Fo-Fc electron density map (blue) contoured at 1.8σ are shown around the bound ATP molecule. Many of the residues involved with binding the ribonucleotide and divalent cations are depicted with sticks.

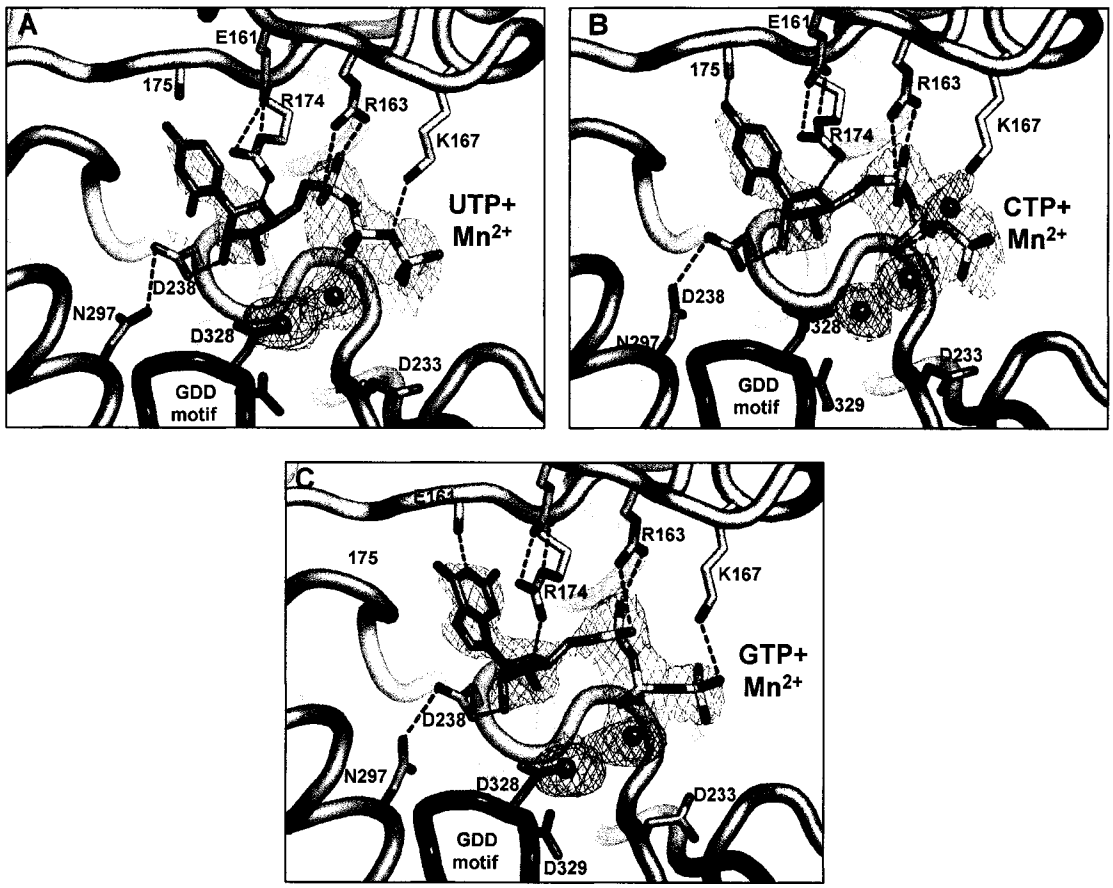


Figure 5.4 Structures of $3D^{pol}$ Complexed with rNTPs and Mn^{2+}
Structures of $3D^{pol}$ complexed with the ribonucleotides in the interrogation site (preinsertion site) with Mn^{2+} (**A**) 2.6 Å resolution 2Fo-Fc map (green) contoured at 1.5σ around rUTP and a 2.6 Å resolution anomalous map (blue; λ = 1.54 Å) contoured at 3.5 σ Mn^{2+}. (**B**) 2.5 Å resolution 2Fo-Fc map (green) contoured at 1.5σ around rCTP and a 2.2 Å resolution anomalous map (blue; = 1.9 Å) at 10σ around Mn^{2+}. (**C**) 2.5 Å resolution 2Fo-Fc map (green) contoured at 1.5σ around rGTP and a 2.5 Å resolution anomalous map (blue; = 1.54 Å) contoured at 3.9σ around Mn^{2+}.

	$3D^{pol}$+UTP	$3D^{pol}$+CTP	$3D^{pol}$+GTP	$3D^{pol}$+ATP
PDB Code	N/A	N/A	1RA7	N/A
Space Group	$P6_5$	$P6_5$	$P6_5$	$P6_5$
Unit Cell (Å)	*a*=*b*=129.0 *c*=113.1	*a*=*b*=127.7 *c*=113.0	*a*=*b*= 128.0 *c*= 113.3	*a*=*b*=128.5 *c*=112.9
X-Ray Source	Cu-Kα 1.54 Å	ALS 0.98	Cu-Kα 1.54 A	Cu-Kα 1.54 Å
Resolution Limits	30 – 2.35 (2.43 – 2.45)	30 – 2.25 (2.33 – 2.25)	30 – 2.35 (2.43 – 2.35)	30 – 2.6 (2.69 – 2.60)
Reflections				
Total Collected	313203	372805	215638	168973
Unique	43793	49540	43126	31801
Redundancy	7.2 (7.1)	7.5 (6.8)	5 (3.9)	5.31 (5.02)
I/σ [1]	16.6 (4.6)	13.1 (3.6)	16.2 (2.1)	11.9 (3.7)
Completeness (%)	98.5 (97.7)	99.7 (99.5)	98.6 (96.2)	97.6 (96.3)
R_{merge} (%)	7.7 (43.9)	8.5 (46.2)	8.2 (55.7)	11.5 (45.8)
Refinement				
Resolution Range	30 – 2.35	30 – 2.25	30 – 2.35	30 – 2.6
R	22.9	23.7	22.4	22.6
R-free (10 % of data)	26.6	26.0	26.2	25.4
Ramachandran statistics [2]				
Favored	378	373	379	368
Allowed	30	36	30	39
Generous	2	1	1	3
Disfavored	1	1	1	1

Table 5-1 Data collection and refinement statistics for $3D^{pol}$-ribonucleotide-Mg^{2+} complexes

Data collection and refinement statistics for the $3D^{pol}$-ribonucleotide-Mg^{2+} complexes. Data in parentheses are for the highest resolution shell, and Ramachandran statistics do not include glycine and proline residues

	$3D^{pol}$+UTP+Mn^{2+}	$3D^{pol}$+CTP+Mn^{2+}	$3D^{pol}$+GTP+Mn^{2+}
PDB Code	N/A	N/A	N/A
Space Group	$P6_5$	$P6_5$	$P6_5$
Unit Cell (Å)	*a*=*b*=128.1 *c*=112.8	*a*=*b*=128.9 *c*=113.0	*a*=*b*=128.2 *c*=112.6
X-Ray Source	Cu-Kα – 1.54 Å	Cu-Kα – 1.54 Å	Cu-Kα – 1.54 Å
Resolution Limits	30 – 2.6 (2.69 – 2.6)	30 – 2.5 (2.59 – 2.5)	30 – 2.5 (2.59 – 2.5)
Reflections			
Total Collected	276952	302975	136298
Unique	32149	36470	36244
Redundancy	8.6 (8.5)	8.3 (8.3)	3.8 (3.7)
I/σ [1]	15.6 (4.7)	15.5 (4.7)	9.9 (2.2)
Completeness (%)	99.3 (98.8)	98.8 (98.5)	99.8 (100)
R_{merge} (%)	9.0 (45.8)	8.2 (44.6)	8.1 (47.6
Refinement			
Resolution Range	30 – 2.6	30 – 2.5	30 – 2.5
R	23.0	23.1	23.0
R-free (10 % of data)	26.3	26.3	25.5
Ramachandran statistics [2]			
Favored	363	360	366
Allowed	44	48	42
Generous	4	2	2
Disfavored	0	1	1

Table 5-2 Data collection and refinement statistics for $3D^{pol}$-ribonucleotide-Mn^{2+} complexes

Data collection and refinement statistics for the $3D^{pol}$-ribonucleotide-Mn^{2+} complexes. Data in parentheses are for the highest resolution shell, and Ramachandran statistics do not include glycine and proline residues

been the most difficult to obtain and therefore only a $3D^{pol}$-rATP-Mg^{+2} structure is presented (Figures 5.2 & 5.3 and Table 5-1).

Ribonucleotide-Divalent Metal Mediated Thermal Stabilization

Although these data are all preliminary, our results confirm the thermal protection of $3D^{pol}$ conferred by excess ribonucleotides. In addition, we show that this protection is stimulated in the presence of stoichiometric concentrations of Mg^{2+} (Mn^{2+} not yet tested).

By performing a poly-A/oligo-dT extension assay to measure ^{32}P-UMP incorporation into a product RNA strand, we observe an inhibition of activity with increasing concentrations of ATP, CTP and GTP (Figure 5.5). This may be due to chelation of necessary divalent metal by the triphosphate moietys of these ribonucleotides. When Mg^{2+} is stoichiometrically added with relatively high concentrations of ribonucleotides (5 - 25 mM), there is a dramatic stimulation of activity until this concentration reaches 50 mM. (Figure 5.5). At the highest concentrations of CTP, GTP, and Mg^{2+} (50 mM), polymerase activity is then inhibited likely due to competition of the active site with radiolabeled UTP (ATP not tested; Figure 5.5).

When $3D^{pol}$ is incubated at various temperatures for 20 minutes prior to elongation in primer extension assays, the polymerase retains activity until this pre-incubation approaches physiological temperatures (Figure 5.6A; blue bars), the T_m of this protein as discussed in chapter VII. As discussed above, the addition of 2.5 mM ATP to the incubated $3D^{pol}$ has an inhibitory effect (~25%

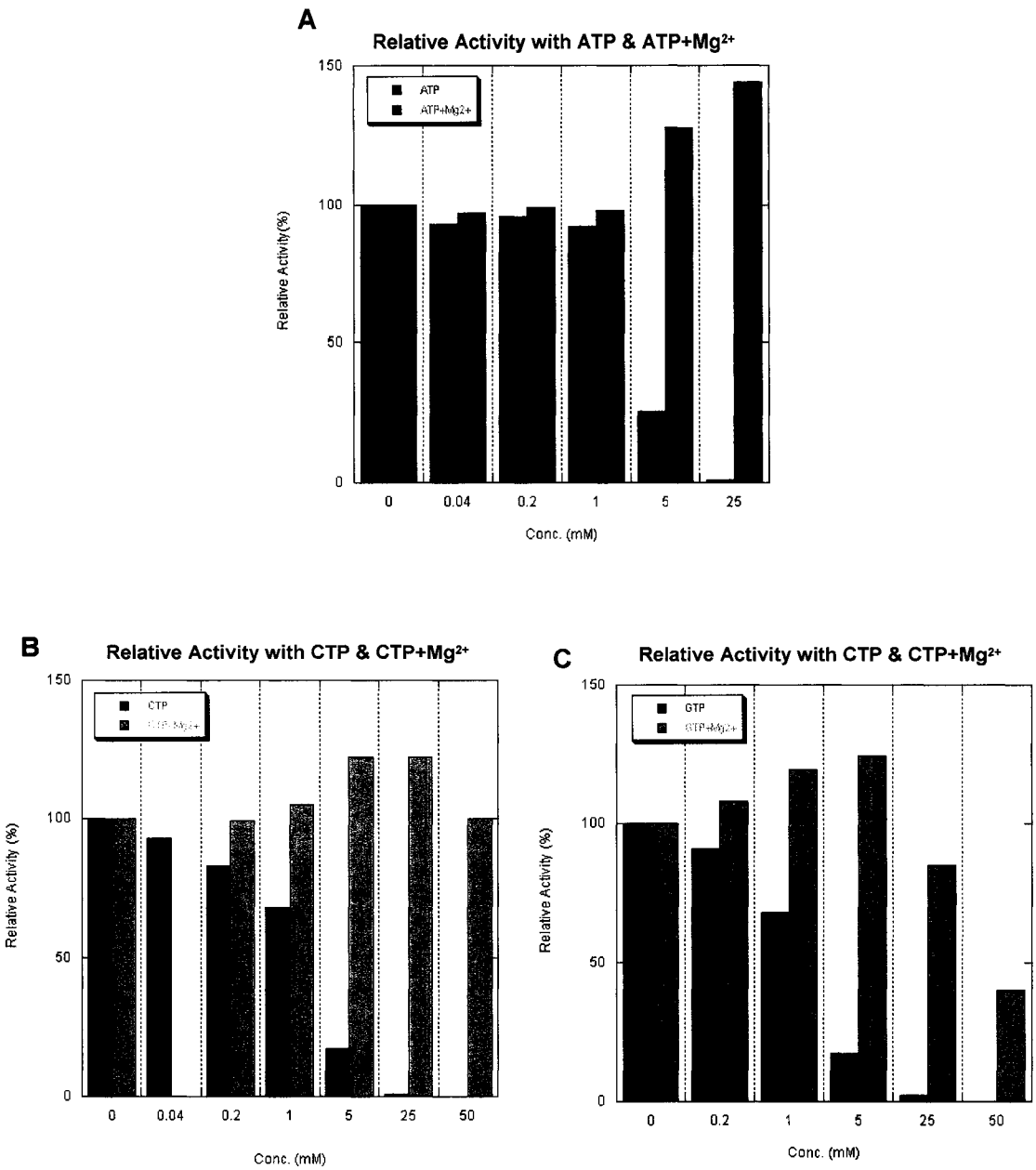


Figure 5.5 Stimulation and Inhibition of Ribonucleotides vs. Ribonucleotides+Divalent Metals

Plot showing inhibition of polymerase activity with increasing amounts of (**A**) ATP, (**B**) CTP or (**C**) GTP. When stoichiometric amounts of Mg^{2+} are added with the ribonucleotides, there is a stimulation of activity until this concentration reached 50 mM. The GTP+Mg^{2+} becomes inhibitory at 25 mM.

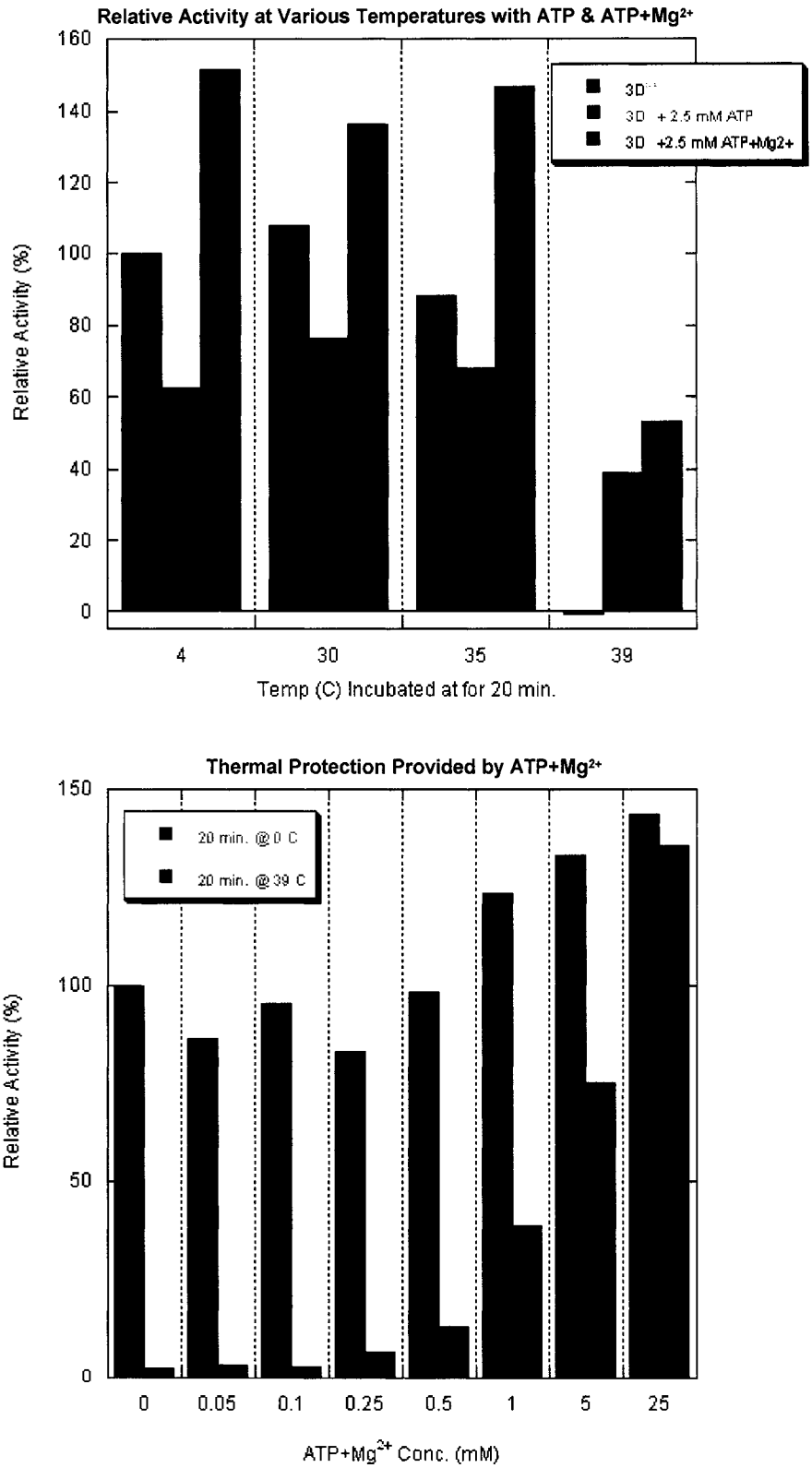


Figure 5.6 $3D^{pol}$ **Thermal Protection provided by ribonucleotides and** Mg^{2+}
(**A**) Relative activity of $3D^{pol}$ after pre-incubation for 20 minutes at various temperatures in the absence of ligand (blue bars, with 2.5 mM ATP (red bars), or with 2.5 mM ATP+Mg^{2+} (green bars). (**B**) Relative activity of $3D^{pol}$ after pre-incubation on ice (light blue bars) and at 39°C (grey bars) with various concentrations of ATP+Mg^{2+}.

decrease in activity) at the lower temperatures; however, this ribonucleotide clearly protects the enzyme from thermal inactivation at 39°C as activity is retained at 30°C (Figure 5.6A; red bars). The addition of 2.5 mM Mg^{2+} to these incubations further stimulates activity at all temperatures compared to $3D^{pol}$ alone or with the addition of only ATP (Figure 5.6A). We then tested the concentration dependence of ATP+Mg^{2+} necessary to provide thermal stability of $3D^{pol}$ during 20-minute incubations at 39°C (Figure 5.6B). In agreement with data described above, ATP+Mg^{2+} stimulates polymerase activity when incubated on ice, and more importantly these data show that low millimolar concentrations of ribonucleotide-divalent metal ion substantially protect the polymerase from thermal inactivation. Rebecca Albertini has obtained preliminary data showing a clear increase in the T_m values that is dependent on rNTP-Mg^{2+} concentrations. Using ANS binding experiments (data not shown), the T_m of unliganded $3D^{pol}$ is just under 40°C, which increases to approximately 42 - 43 °C in the presence of low CTP+Mg^{2+}concentrations (2 - 5 mM), and 43.5 - 45 °C at higher ribonucleotide-Mg^{2+} concentrations (7.5 - 20 mM).

Discussion

All four rNTPs bind polio $3D^{pol}$ in the same site, which is a pre-catalytic site that corresponds to the "interrogation site" described in the bacteriophage Φ6 polymerase structure (Butcher and Grimes *et al*, 2001) and the "pre-insertion" site in bacteriophage T7 polymerase (Figure 5.7B; Yin and Steitz, 2002). The phosphate groups are *not* correctly positioned above the active site aspartate residues involved in binding the catalytically required magnesium or manganese metal ions (Asp328, Asp329 and Asp233). Rather, the three phosphates are trailing out through the rNTP entry tunnel, where they interact with conserved basic residues in a structure that appears to select for a complete triphosphate moiety. These $3D^{pol}$-rNTP-divalent metal ioncomplexes could be obtained either by co-crystallization or by soaking of $3D^{pol}$ crystals, indicating that NTP binding is not accompanied by large conformational changes in the protein that would break the crystal lattice.

The Ribonucleotide-Binding Pocket

The co-crystal structures of $3D^{pol}$ in complex with all four ribonucleotides show that a number of hydrogen bond and electrostatic interactions are responsible for binding and positioning the substrates. The rNTPs are all bound with their 2' OH group making the anticipated hydrogen bond with Asp238 (Figures 5.2, 5.3, 5.4), and when Asp238 is mutated to an alanine, there is no crystallographic evidence of nucleotide binding, thus highlighting the pivotal importance of this residue. The nucleosides are positioned such that the base

Figure 5.7 Structures of Bacteriophage T7 Polymerase showing Transiocation Structures of the bacteriophage T7 polymerase active site showing the various steps of translocation. The "O" helix, which is structurally analogous with the $3D^{pol}$ ring finger is colored yellow, the catalytic center is magenta. Carbon atoms from the template nucleic acid are colored cyan, the daughter strand is yellow, and incoming nucleotides are brown. (**A**) Catalytically inactive "open" conformation of elongation phase (1MSW; Yin and Steitz, 2002). (**B**) Inactive "open" conformation of elongation with ATP bound at the pre-insertion site (1S0V; Yin and Steitz, 2004). (**C**) (**D**) Catalytically active "closed" polymerase conformation due to conformational change of the "O" helix (1S76 & 1S77; Temiakov *et al.*, 2004). Bridging interactions between the bacteriophage T7 polymerase "O" helix (ring finger) and palm aspartate carboxylates, mediated by Mg^{2+} and pyrophosphate, are responsible for open and closed conformations of the polymerase and therefore translocation

moieties interact with the guanidinium group of Arg174, from the ring finger, in a cation–π stacking interaction similar to that observed in single-stranded nucleic acid-binding proteins (Theobald *et al*, 2003). The Arg174 guanidinium group is further locked in place by a charge interaction with Glu161 from the other end of the ring finger, and it makes a hydrogen bond with the ribose sugar ring oxygen. The triphosphate moieties are involved in ionic interactions with Lys167 and Arg163 of the ring finger that drop down from the roof of the rNTP entry tunnel, and the β phosphate is hydrogen bonded to the backbone amide of residue 236 in the palm domain.

Ribonucleoside Conformation

The energetically most stable conformation of the ribonucleotide five membered ribose ring is intrinsically non-planer. The resulting sugar pucker enables substituents at the four ring carbon atoms to distance themselves maximally. Typically the C2' and C3' atoms of the ribose ring are displaced on different sides of the ring with one of these exhibiting a larger deviation than the other (Neidle, 2002). In all of the $3D^{pol}$-ribonucleotide complex structures, the observed pucker is of the C3'-*endo* type, where the C3' shows the largest deviation from the sugar plane towards the C4'-C5' bond and the C2' deviates away from this bond in the *exo* direction (Figure 5.8) (other types discussed in Neidle, 2002). Although purines often show a preference for C2'-*endo* puckers in solid state structures, NMR studies suggest that there is rapid interconversion in solution with an energy barrier between 1.5 – 5 kcal/mol (Neidle, 2002; Harvey and Prabhakaran, 1986; Olson and Sussman, 1982). The C3'-*endo* ribose pucker

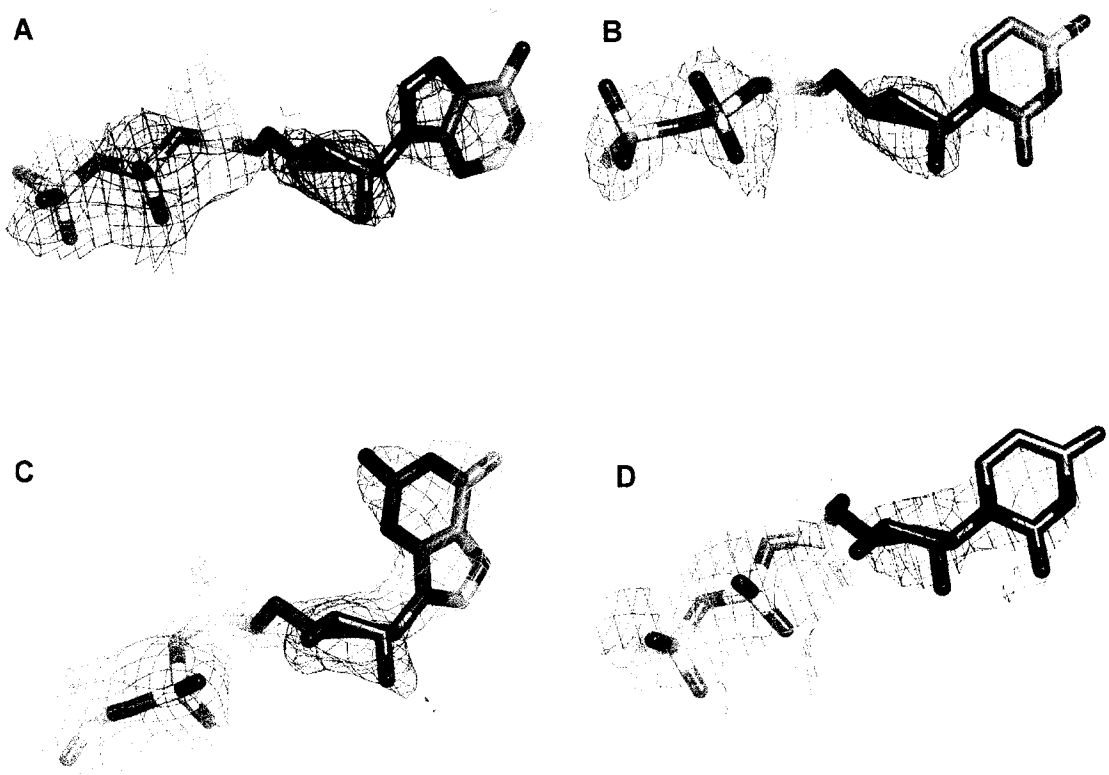


Figure 5.8 Ribonucleotide Conformations
Structures of ribonucleotides as they are bound to $3D^{pol}$ at the interrogation site (preinsertion site). (**A**) 2.6 Å resolution 2Fo-Fc map (green) contoured at 1.2σ and Fo-Fc map (blue) contoured at 1.8σ around rATP. (**B**) 2.25 Å resolution 2Fo-Fc map at 1.4σ contoured around rCTP. (**C**) 2.35 Å resolution 1500 K composite SA-omit map at 1.6σ contoured around rGTP. (**D**) 2.6 Å resolution 2Fo-Fc map (green) contoured at 1.4σ around rUTP

we observe in all $3D^{pol}$-rNTP structures may be dictated by the hydrogen bond interactions of the 2' hydroxyl with Asp238 and the O4' with Arg174, an interaction that positions the sugar moiety of the ribonucleotide.

Rotation about the glycosidic bond linking the ribose sugar and base is predicted to populate two predominant conformations (Neidle, 2002). The existence of two low-energy minima results in an *anti* conformation, where the base-pairing face points away from the sugar, and a *syn* conformation, where the base-pairing face is directed towards the ribose ring. The s*yn* and *anti* conformations of pyrimidine bases result in repositioning of a single exocyclic O2 atom. Although the electron density for these bases is not outstanding, these ribonucleosides appear to be in the *anti* conformation (Figure 5.8), and refinement of these nucleosides in the *syn* conformation yields significant negative |Fo-Fc| difference electron density around the exocyclic O2 atom of the base. We also observe rATP in the *anti* conformation whereas the rGTP base sits over the sugar in the *syn* conformation. These findings are in agreement with calculated energy differences of DNA nucleotide conformations (Table 5-4). The guanosine nucleotide is thought to favor the *syn* conformation because of favorable electrostatic interactions between the exocyclic N2 atom of the base with the 5' phosphate atom (Neidle, 2002). However, the structures of

Nucleotides	Energy difference (kcal / mol)
A(anti) > A(syn)	0.3
T(anti) > T(syn)	1.7
G(syn) > G(anti)	3.3
C(anti) > C(syn)	1.8

Table 5-4 *syn-anti* Energy Differences
Table from Neidle, 2002 showing energy differences (kcal / mol) between glycosidic angle conformations for B-form DNA nucleotides, calculated with the AMBER program.

$3D^{pol}$ complexed with the four ribonucleotides show that this interaction is precluded due to steric interference from the guanidinium group of Arg174 (Figures 5.2 & 5.3). While it is possible that the *syn* guanosine conformation is stabilized by the interaction between the guanine's N1 atom and the carbonyl group of residue 175, this does not explain the observed conformation of the adenine base. This base has the potential for the same interaction as the guanosine, but the adenine is found stabilized in the *anti* conformation by potential hydrogen bond interactions between the carbonyl group of residue 175, and the N7 and exocyclic N6 atoms of the adenine base.

Divalent Metal Ion Analysis

Crystallization of $3D^{pol}$ with ribonucleotides in the presence of various divalent cations has not yielded discernable electron density for Mg^{2+}, however, we are able to confirm the presence of two Mn^{2+} cations in the active site by anomalous difference density maps (blue maps in Figure 5.4). The apparent lack of Mg^{2+} binding may be due to its scant scattering power, which is roughly equal to that of a water molecule. The relatively electron dense Mn^{2+} on the other hand results in large |2Fo-Fc| electron density peaks, and its anomalous absorption edge (~1.9 Å) yields large peaks in anomalous difference electron density maps calculated with 1.54 Å and 1.9 Å wavelength data (Figure 5.4)

The observed binding of the divalent metal ions is consistent with the established two-metal mechanism of catalysis (Brautigam and Steitz, 1998; Steitz, 1998; Sosunov *et al.*, 2003, 2005). This model involves the binding of two

divalent cations termed metal ions A and B, which have distinct roles. Metal ion A (bottom left metal ion in Figure 5.4) is thought to function as a Lewis acid that activates the 3' OH group of the primer to an O^- for nucleophilic attack on the α-phosphate of the incoming rNTP. Metal ion B (upper right metal ion in Figure 5.4) is thought to interact with and stabilize the negatively charged triphosphate moiety in a catalytically competent conformation. *In vivo*, the *ad hoc* utilization of metal ion B is thought to be supplied to the active site in a complex with the triphosphate moiety of incoming rNTPs. Our structures show that these metal ions are coordinated by the triphosphate moiety and the three conserved aspartates of the catalytic site (Asp328, Asp329, Asp233). Occasionally we observed a third anomalous peak that may be from a Mn^{2+} ion with partial occupancy coordinated to the triphosphate moiety (Figure 5.4B). While the purpose of this divalent cation in unclear, a third metal located in the same position is also observed in the HCV NS5B polymerase complexed with UTP and Mn^{2+} (Bressanelli et al., 2002), and the bacteriophage ϕ polymerase structure complexed with template-primer and Mg^{2+}/Mn^{2+} (Butcher and Grimes et al., 2001)

The structures of $3D^{pol}$ complexed with the various ribonucleotides and divalent cations show that Mg^{2+} and Mn^{2+} do not significantly affect the orientation of the rNTPs. However, in the presence of Mn^{2+}, superposition of the $3D^{pol}$ active site shows that the side chain of aspartate 233 is "swung" away from the catalytic center and the polypeptide backbone of this region is pulled in towards the active site by 1.1 Å (Figure 5.9). The flexibility of this conserved

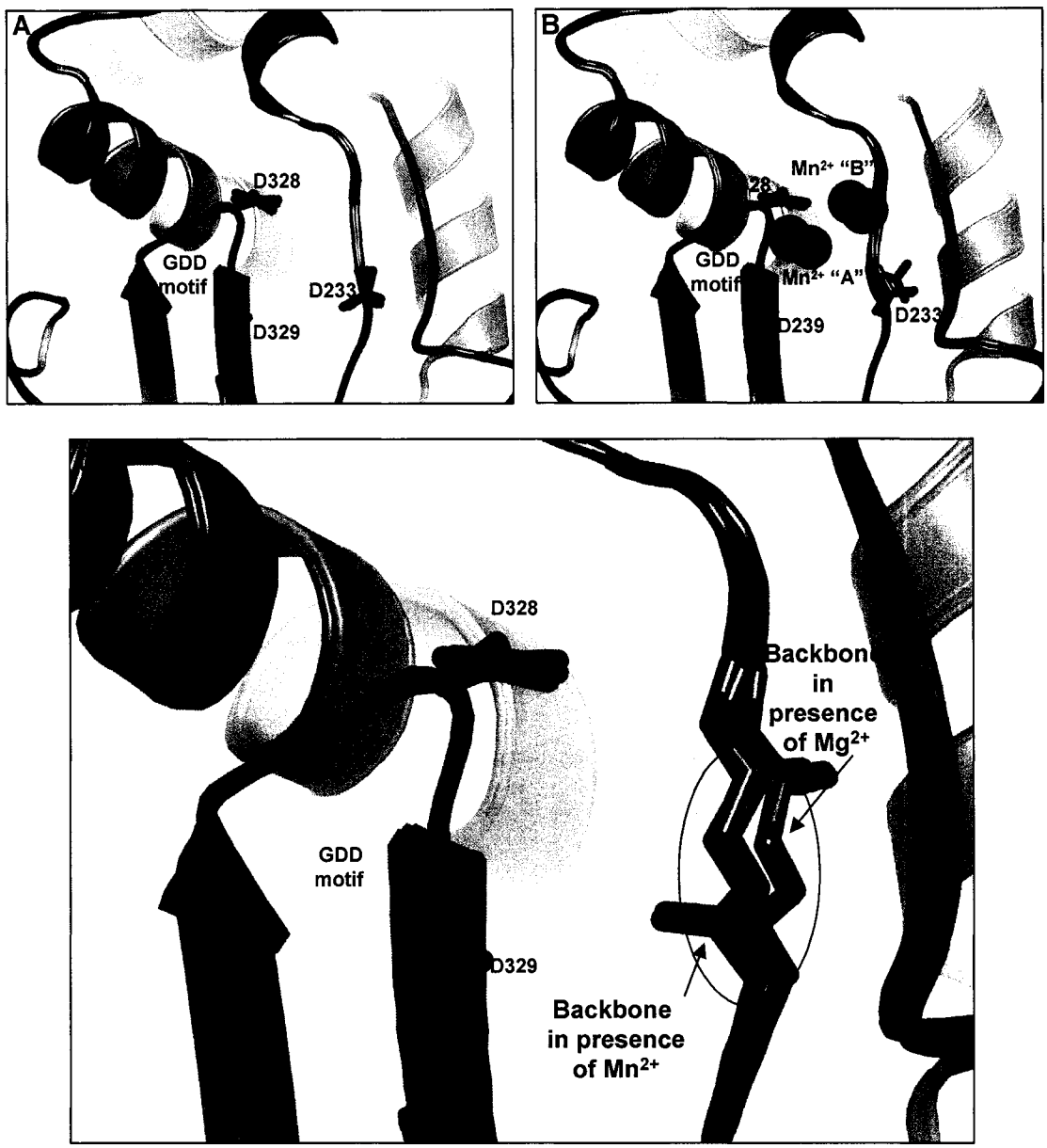


Figure 5.9 3D^{pol} Palm topology in the presence of Mg^{2+} or Mn^{2+}
(**A**) Four 3D^{pol} structures in the presence of Mg^{2+}, and (**B**) Three 3D^{pol} structures complexed with Mn^{2+} all superimposed by active sites (**C**) Structures of 3D^{pol} complexed with Mn^{2+} (brown palm) superimposed onto the active site of structures complexed in the presence of Mg^{2+} (grey palm) to highlight the 1.1 Å movement of the polypeptide backbone towards the GDD motif in the presence of Mn^{2+}.

region is further highlighted by several other RdRP structures complexed with either divalent or trivalent (Sm^{3+} & Lu^{3+} for phasing purposes) cations in different crystal forms (Ng *et al.*, 2002; Love *et al.*, 2004).

While both Mg^{+2} and Mn^{+2} support polymerase activity *in vitro*, there are differences in fidelity and kinetics. Many studies report that nucleotide specificity is most stringent when Mg^{+2} rather that Mn^{+2} is the metal cofactor (Arnold *et al*, 2004; Ide *et al*, 1993; Sousa & Padilla, 1995; Tabor & Richardson, 1989). Fidelity measurements using Mn^{+2} as the cofactor instead of Mg^{+2} have indicated differences in the biochemical properties of the enzyme, permitting the increased use of DNA templates and nucleotides with incorrect sugar conformations or incorrect bases (Arnold *et al*, 1999). The most common explanation typically lies with the ability of each metal, with different ionic radii, to properly coordinate substrates for formation of a phosphodiester bond. Because the rate of nucleotide incorporation is approximately ten times faster when Mn^{2+} is the cofactor, it has been proposed that an accelerated phosphoryl transfer step in the presence of this divalent metal results in a reduction of nucleotide interrogation and thus fidelity (Arnold *et al*, 2004). In a related model, it has been suggested that the higher affinity Mn^{+2} has for the active site and triphosphate moieties may alter the kinetics of the reaction such that non-cognate and incorrectly positioned nucleotides lacking the full complement of native interactions may be incorporated at a higher rate (Castro et al., 2004). In the HCV polymerase, the apparent K_d values for the binding of Mg^{2+} and Mn^{2+} are 3.1 and 0.3 mM,

respectively (Bougie et al., 2003). Although we do not observe binding of Mg^{2+}, our structures do not allow us to either support or discount any of these models.

Ribonucleotide / Divalent Metal Thermal Protection

Earlier studies found that incubation of the enzyme with high concentrations of NTPs resulted in a protection against thermal inactivation and activity loss (Richards et al., 1992; Arnold and Cameron 2000). Biophysical studies of the HCV NS5B polymerase show that structural stability of this enzyme is increased in the presence of Mg^{2+} (Benzaghou et al., 2004). Although all ribonucleotide/divalent metal ion-mediated thermal stabilization data we have collected are preliminary at this point, we confirm the thermal protection provided by ribonucleotides and show this is enhanced in the presence of divalent metal ions. Based on analysis of the crystal structures we have obtained, we believe this thermal protection is due to NTP-metal facilitated bridging interactions between the palm and fingers domains that stabilize the polymerase structure. The importance of these interactions are underscored by evidence suggesting that electrostatic bridging interactions between the bacteriophage T7 polymerase "O" helix (ring finger) and palm aspartate carboxylates, mediated by Mg^{2+} and pyrophosphate, are responsible for open and closed conformations of the polymerase and therefore translocation (Figure 5.7; Yin and Steitz 2004).

Rebecca's ANS-binding thermal denaturation studies show an increase in structural stability, by virtue of an increase in the T_m with increasing concentrations of ribonucleotide-divalent metal ion. An increase in protein

thermostability is common following ligand binding, and the denaturation temperatures (T_ms) often increase concomitantly with ligand concentration (Celej et al., 2005). As discussed in Chapter VII, we believe this is due to rNTP-metal ion-facilitated bridging interactions between the palm and the marginally stable fingers domain, resulting in stabilization of the structure

Conclusions

The existing RdRP structures complexed with ribonucleotides, and inspection of all available electron density maps from these structures, suggests that poliovirus $3D^{pol}$ provides unique conditions for the binding and crystallization of complexes with ribonucleotides. To date, there is no report presenting the structures of an RdRP complexed with all four ribonucleotides. In addition, Aaron J. Sholders has been able to obtain excellent structures of $3D^{pol}$ complexed with the antiviral nucleoside analogue ribavirin, which has not yet been reported by other investigators.

The structures of $3D^{pol}$ complexed with ribonucleotides show that extensive hydrogen-bonding, cation–π stacking, and electrostatic interactions facilitate binding at the "interrogation" or "pre-insertion" site. These interactions involve Asp238 and backbone amine of residue 236, which are both from the palm, and several conserved basic residues from the ring finger including Arg174, Arg163, and Lys167. The rNTP-mediated bridging interactions thus serve to link the palm and fingers domains. The structures bound with Mn^{2+} suggest divalent cations enhance this domain linkage through electrostatic interactions between the conserved aspartates of the palm domain and the triphosphate moiety of the ribonucleotides. These observations provide a structural explanation for the increase in thermal stability of the protein in the presence of these ligands.

CHAPTER VI

ANALYSIS AND CRYSTALLIZATION OF THE PINKY MUTANTS

In all known polymerase-nucleic acid structures, entering single-stranded template binds the fingers domain before interacting with the active site (Cramer 2002). This interaction typically involves an approximate 90° bend between the helical axes of entering and exiting nucleic acid (Cramer 2002). There are several structures available of the DNA-dependent RNA polymerase from bacteriophage T7 in complex with a number of nucleic acid templates and template-product pairs. A comparison of these structures with the $3D^{pol}$ structure provides insight into the potential function of the poliovirus $3D^{pol}$ fingers domain. Located at the top of the $3D^{pol}$ pinky there is asmall helix (residues 118-123) followed by a large loop structure (residues 124-149) that appears to be quite flexible by virtue of weak electron density and high B-factors (Figure 6.1A, Table 6-1). This loop is most likely involved in RNA binding based on a structural alignment of the $3D^{pol}$ active site with that of

(Sub)domain	B-factor (Å^2)
Thumb	48.0
Palm	42.9
Index finger	58.1
Middle finger	51.1
Ring finger	55.2
Pinky finger	75.0
Pinky loop (96-149)	78.4
Pinky base (180-190)	57.0

Table 6-1 Average (Sub)domain B-factors
Average B-factors of native $3D^{pol}$ structure (PDB code 1RA6). All atoms are included in the average, and the pinky is subdivided into its two polypeptide regions.

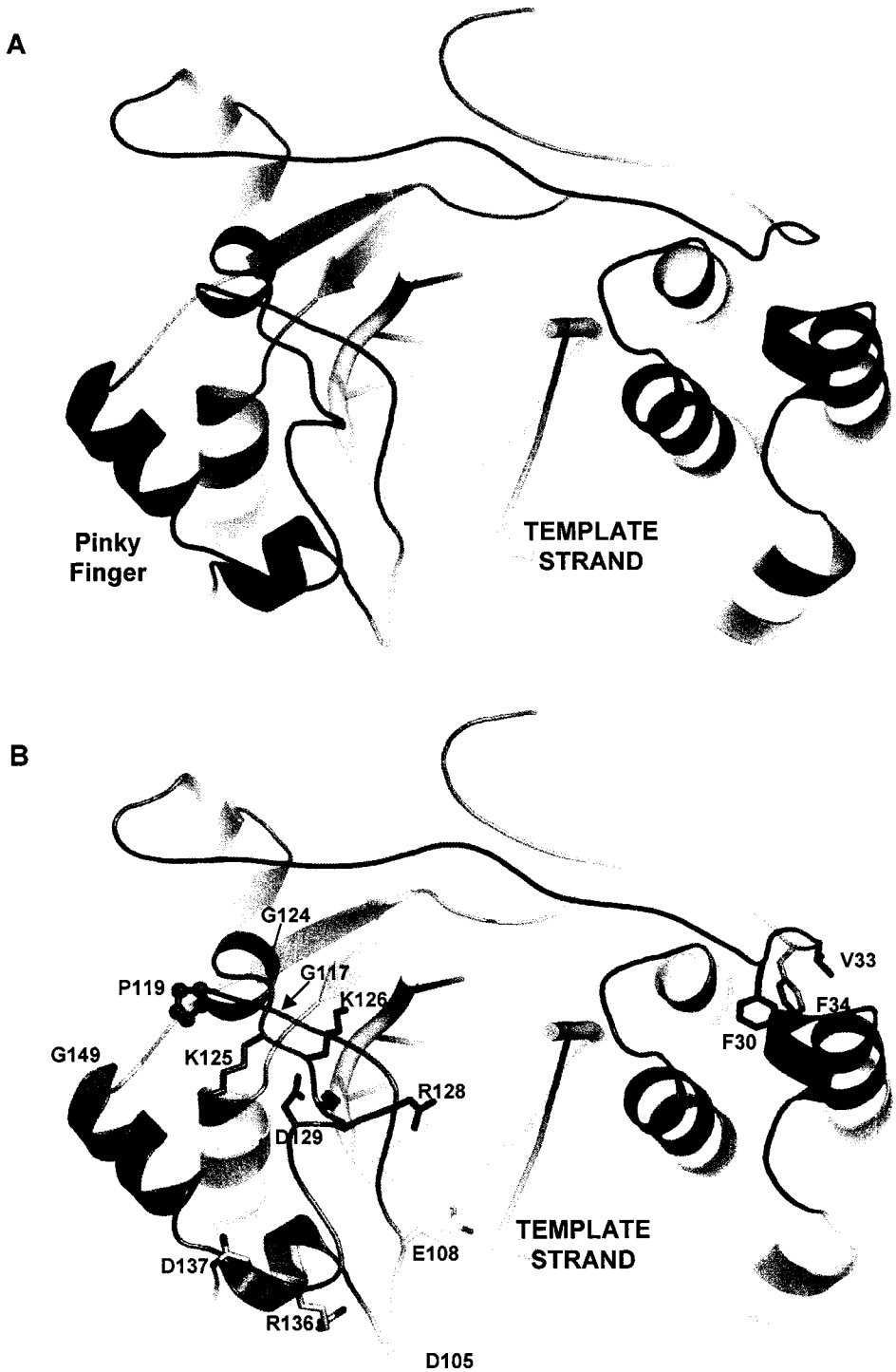


Figure 6.1 The 3D^{pol} Pinky Finger.
Top view of 3D^{pol} modeled with nucleic acid template-primer from FMDV (**A**) 3D^{pol} colored according to B-factors with warm (red) colors for backbone residues representing high values and cold (blue) colors for residues with low values (8 bins). (**B**) 3D^{pol} is colored with a blue thumb, red fingers and pink pinky finger. Mutation of residues that destroy viral infectivity are colored purple and residues with moderate effects on plaque formation are colored wheat (data from Diamond and Kirkegaard, 1994). Glycines and Pro119 of the putative isomerization motif are colored cyan.

bacteriophage T7 RNA polymerase (Figure 6.2). In this superimposition with the initiation complex (PDB code 1CEZ), the pinky loop roughly superimposes on the T7 polymerase specificity loop (residues 740-770), which makes sequence-specific contacts with the T7 promoter in the initiation complex (Cheetham et al., 1999a) and non-specific contacts in the T7 elongation complex (Temiakov et al., 2000; Yin and Steitz, 2002). Karla Kirkegaard and colleagues have shown by clustered charged-to-alanine analysis that the basic residues predicted to interact with template on the inside of $3D^{pol}$'s pinky are absolutely required for plaque production following transfection of full-length poliovirus cDNA clones (purple residues in Figure 6.1B) (Diamond and Kirkegaard, 1994).

A superposition of the $3D^{pol}$ active site with the T7 polymerase active site in that enzyme's elongation complex structure (Yin & Steitz, 2002) shows that the template strand of the T7 complex roughly threads through a cleft in the $3D^{pol}$ fingers domain between the index and pinky fingers that may serve as a template entry channel (Figures 2.7 & 4.3). This short helix in the putative template entry channel contains a *cis* proline (Pro119) that is 100% conserved among picornaviral polymerases and is flanked by glycines 117 and 124 that are also 100% conserved (Figures 6.1B & 6.3). This sequence is consistent with an emerging conserved motif used for facilitating alternate conformations of loops via proline *cis/trans* isomerization (Andreotti, 2003).

Prolines are unique among the 20 naturally occurring amino acids because isomerization about the peptide bond (dihedral angle ω) yields both *cis* and *trans* conformations that are nearly isoenergetic. The *cis* isomer of the remaining 19

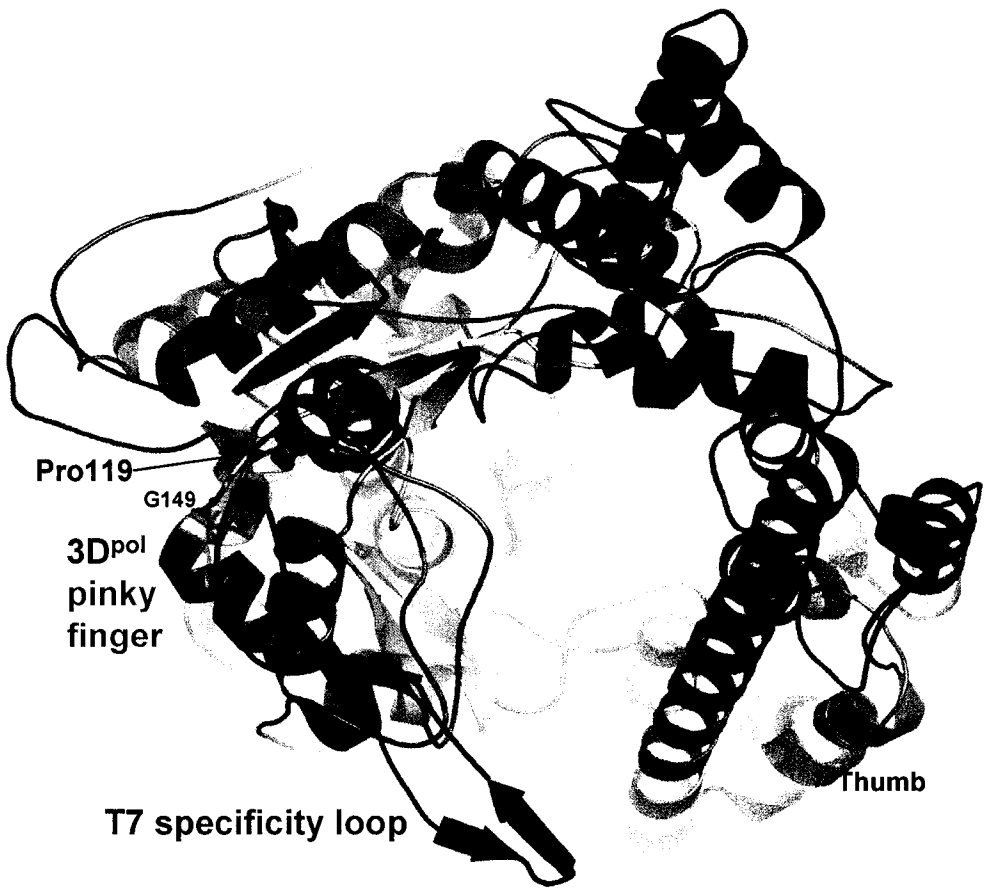


Figure 6.2 3D^{pol} Superposition with Bacteriophage T7 RNA PolymeraseInitiation Structure.
Superposition of 3D^{pol} with the T7 polymerase structure (brown) from the initiation complex (1CEZ; Cheetham *et al.*, 1999) showing the structural alignment of the 3D^{pol} pinky finger (pink) with the T7 specificity loop (red). Comparison of the poliovirus 3D^{pol} and bacteriophage T7 polymerases are based on superimposing the C_α atoms of their 'motif C' structures. This motif contains the pair of β–strands in the core of the conserved palm that present the β–turn GDD motif in the active site.

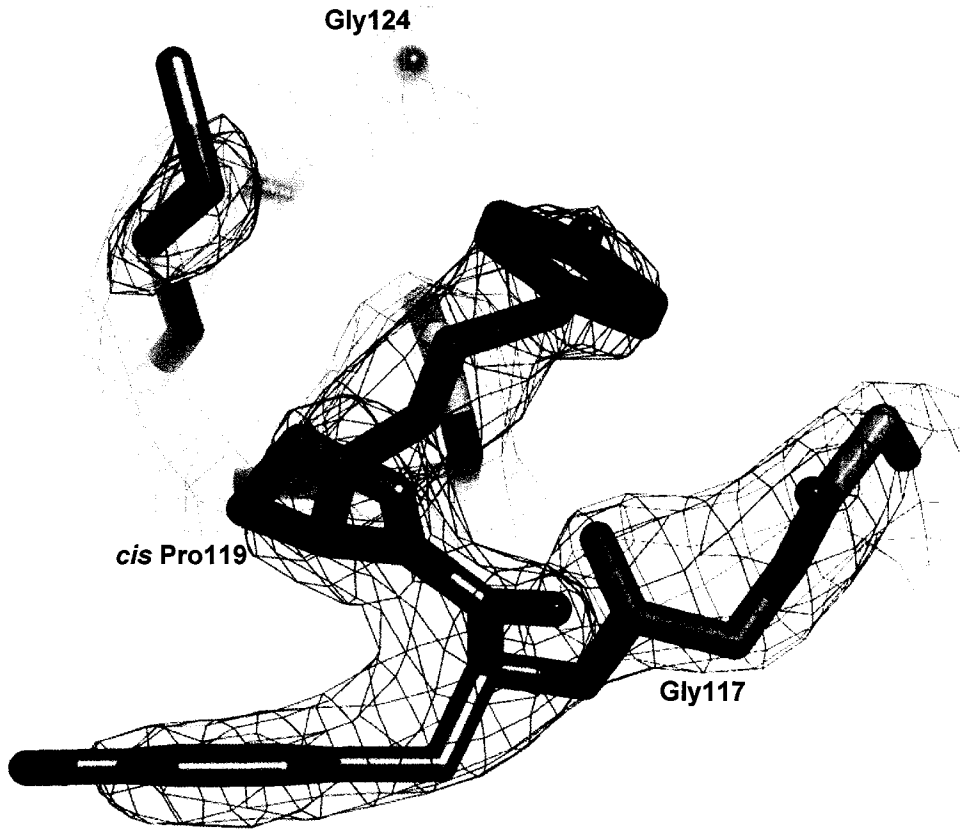


Figure 6.3 *cis* Proline 119 flanked by Glycines 117 and 124
Electron density map contoured around the putative proline isomerization motif on the pinky finger (pink). Glycines 117 and 124, which flank the *cis* proline, are colored cyan. The 2 Å resolution 2Fo-Fc electron density map is contoured at 1.6 σ.

amino acids is not observed due to unfavorable interactions between the C_1^{α} and H_1^{α} atoms with the C_2^{α} and H_2^{α} atoms, and from more favorable electrostatic interactions between O_1 and C_2 in the trans form (Figure 6.4). Because of the cyclic nature of prolines, unfavorable steric interactions between the C_1^{α} and C_2^{α} atoms with the C_1^{α} or C^{δ} atoms bonded to the proline amide nitrogen are energetically comparable. However, the isomerization has a relatively large energy barrier of 14-24 kcal/mol due to the partial double bond character of the peptide bond (Kang and Choi, 2004; Wedemeyer *et al.*, 2002 and references therein). Because of the existence of two distinct energy minima separated by this energy barrier, regions of a fully folded protein containing a proline can populate two discrete conformations.

It had previously been shown that upon incubation of $3D^{pol}$ with primer-template, there is a slow isomerization ($t_{1/2} \sim 12$ s) to an elongation-competent state that has a half-life of ~2-8 hours (Arnold & Cameron, 2000). This slow conversion to the elongation-competent state is consistent with proline isomerization relaxation rates which are on the order of tens to hundreds of seconds at 25° C (Brandts et al., 1975; Grathwohl & Wuthrich, 1981; Schmid, 1993; Reimer et al., 1998). While the exact structural role of Pro119 in the $3D^{pol}$ elongation complex remains to be determined, our data show that this residue is essential for activity and suggest that its ability to isomerize may be a required structural feature. We therefore hypothesize that *cis-trans* isomerization of Pro119 results in conformational changes in the pinky finger that lock the enzyme into the stable elongation-competent mode.

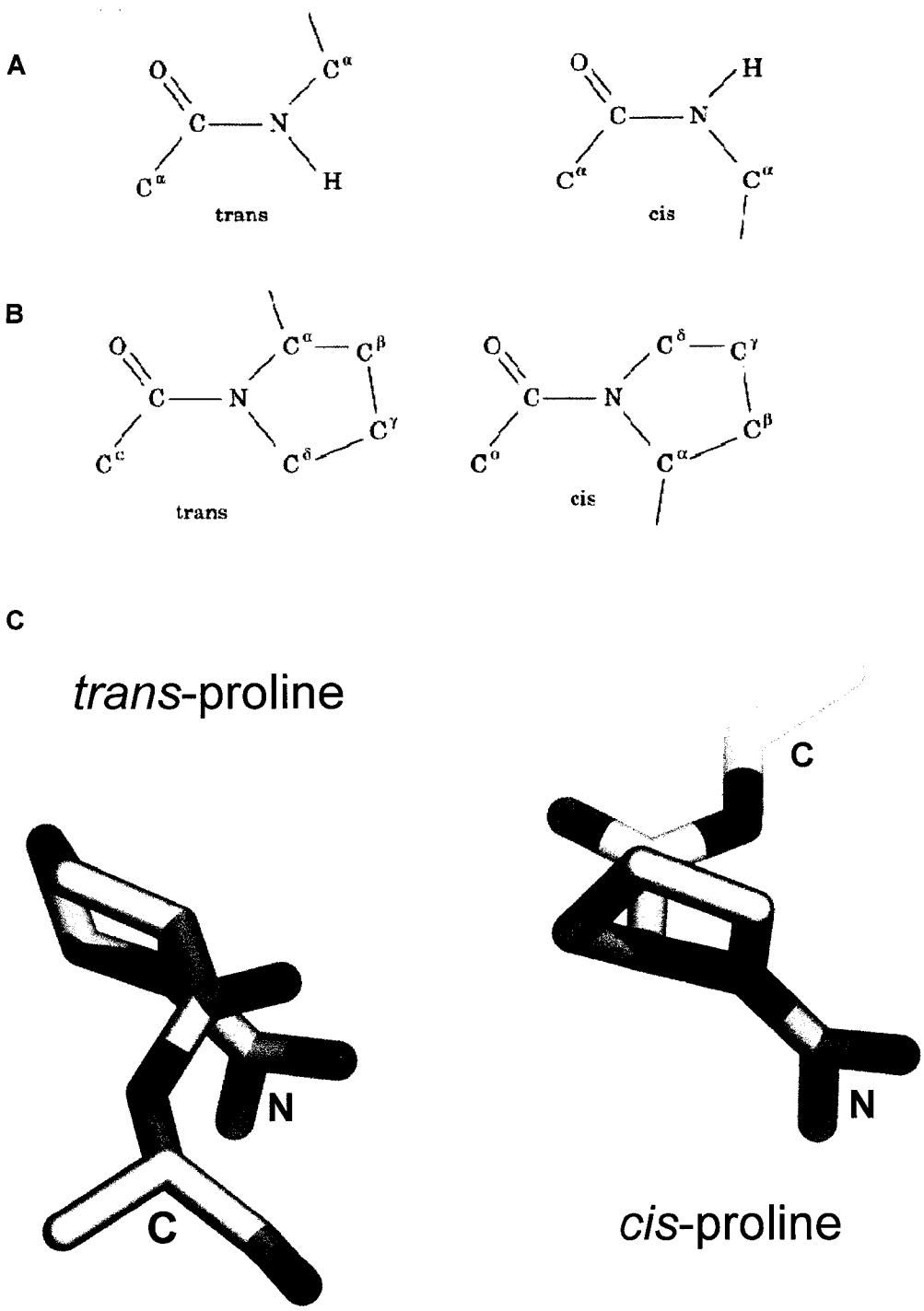


Figure 6.4 Proline Isomerization
Non-proline (**A**) and proline (**B**) peptide fragments in *trans* and *cis* conformations. (fig. adapted from Wedemeyer *et al.*, 2002). (**C**) Three-dimensional representation of the *trans* and *cis* proline conformations. The cis proline is orientated as Pro119 in Fig. 6.3. The N and C-terminal ends of each peptide are labeled and colored orange.

Results

To ascertain the importance of the putative proline isomerization motif, we began by mutating proline119 to both an alanine and flexible glycine and found that both mutations completely abolish elongation of a poly(A)/oligo(dT) substrate (Table 6-2). While these mutants would not crystallize, circular dichroism studies indicated identical secondary structure composition to that of the native polymerase (data not shown). Using the fluorescence polarization assay described in appendix 1, RNA binding constants (K_ds) of these mutants were determined and compared to the native construct. The $3D^{pol}$-P119A and $3D^{pol}$-P119G mutants had very similar K_ds, which both indicated a 1.25 fold decrease in affinity when compared to the native polymerase (Table 6-2).

Table 6-2 Pinky Mutant Activity and RNA binding K_ds

Protein	% Activity	RNA binding K_ds
$3D^{pol}$ L446D-R455D	100 ± 4	1.75 ± 0.05 X 10^{-6} M
Mutants below are in addition to L445d-R455D		
$3D^{pol}$ P119A	-0.1 ± 0.2	2.28 ± 0.07 X 10^{-6} M
$3D^{pol}$ P119G	-0.1 ± 0.2	2.2 ± 0.1 X 10^{-6} M
$3D^{pol}$ G117A	0.8 ± 0.3	1.81 ± 0.06 X 10^{-6} M
$3D^{pol}$ G124A	2.4 ± 0.4	0.85 ± 0.02 X 10^{-6} M
$3D^{pol}$ G149A	0.3 ± 0.3	1.83 ± 0.07 X 10^{-6} M

The glycines that flank Pro119 are important residues of the putative isomerization motif by virtue of the backbone flexibility of these residues. Mutation of glycines 117 and 124 to alanines almost completely abolishes polymerase activity. The $3D^{pol}$-Gly117Ala construct displayed aRNA binding K_d similar to the native polymerase, and it crystallized, yielding an interesting 2.2 Å resolution structure showing disorder of the polypeptides flanking this region (Table 6-2 & 6-3, Figure 6.6). The $3D^{pol}$-Gly124Ala mutant presented solubility

problems during concentration of the protein after the final gel filtration column (in 200 mM NaCl). Although the protein solution remained clear, it stopped flowing through the Centricon-30, reaching a solubility limit of ~50 μM. Because this construct contains the L446A-R455D mutation, which normally yields highly soluble protein, we suspected structural defects of the $3D^{pol}$-G124A mutant and this was confirmed by CD wavelength scan and thermal melt analysis (Figure 6.5). Despite only a two-fold increase in the RNA binding Kd (Table 6-2), it came as no surprise that this mutant would not crystallize.

Glycine 149 lies on the "back" side of the pinky finger and appears to be a hinge point if this entire finger changes conformation. This glycine has weak electron density and high B-factors, compared to other amino acids in this region, suggesting it may be flexible. Although mutating this residue to an alanine results in only a slight decrease in the RNA binding affinity, this construct displayed absolutely no polymerase activity (Table 6-2). Approximately one month after setting up crystallization trays, a single cluster of very "sick"-looking crystals grew and were teased apart, enabling a low-resolution (2.9 Å) view of this structure. However, initial electron density maps did not reveal any significant conformational change and therefore this structure was not extensively refined.

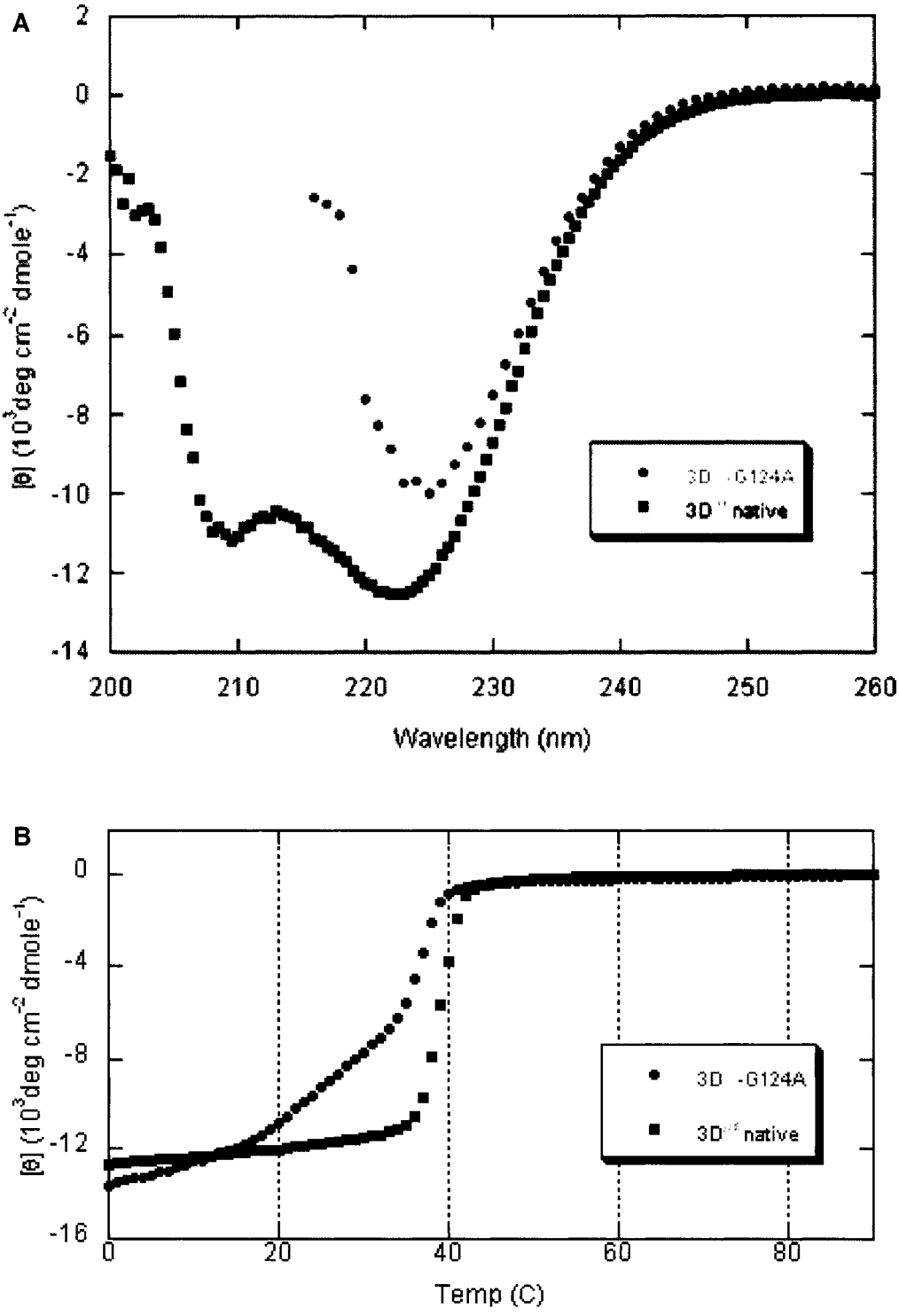


Figure 6.5 CD data for 3D^{pol}-G124A

(**A**) Wavelength and (**B**) thermal denaturation CD data for native 3D^{pol} (blue points) and the 3D^{pol}-G124A mutant (red points). The wavelength spectra were taken at 20°C and the thermal denaturation spectra were taken at 222 nm wavelength. The wavelength data collected for the G124A mutant was unreliable <215 nm and therefore is not shown.

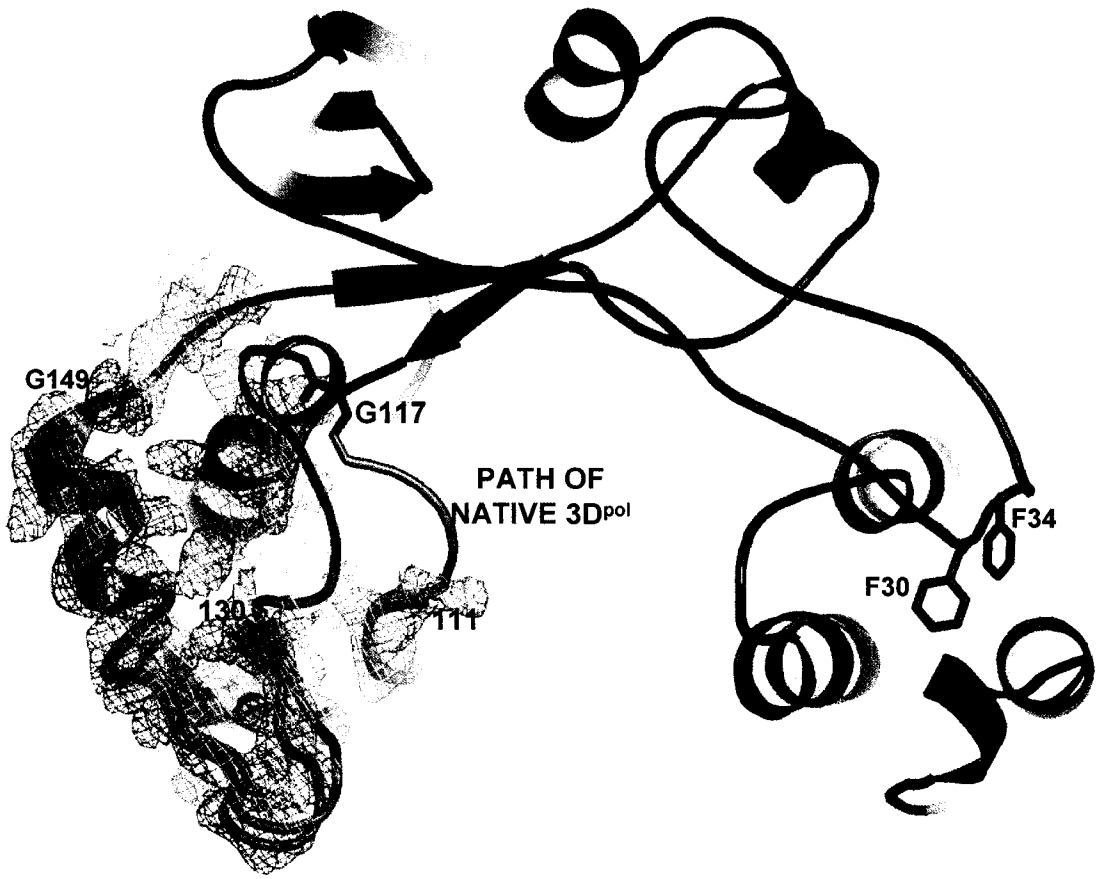


Figure 6.6 $3D^{pol}$-G117A Structure
2.2 Å resolution Electron density map (yellow) of the G117A mutant pinky region. Residues 112-129 are disordered so the backbone path of the native structure is shown in green with the Gly117 residue colored cyan. The cartoon structure of the G117A mutant is colored with a blue thumb and red fingers with the pinky finger colored pink.

	$3D^{pol}$-G117A
PDB Code	N/A
Space Group	$P6_5$
Unit Cell (Å)	$a=b=128.5$ $c=113.8$
X-Ray Source	ALS – 1.1 Å
Resolution Limits	30 – 2.20 (2.28 – 2.20)
Reflections	
Total Collected	369988
Unique	54060
Redundancy	6.8 (6.7)
I/σ [1]	12.4 (3.2)
Completeness (%)	100 (100)
R_{merge} (%)	6.6 (41.7)
Refinement	
Resolution Range	30 – 2.30
R	22.2
R-free (10 % of data)	24.7
Ramachandran statistics [2]	
Favored	361
Allowed	31
Generous	1
Disfavored	1

Table 6-3 Data collection and refinement statistics for $3D^{pol}$-G117A.
Data collection and refinement statistics for the $3D^{pol}$-G117A mutant. Data in parentheses are for the highest resolution shell, and Ramachandran statistics do not include glycine and proline residues

Discussion

3D^{pol}-P119 mutants

Mutating proline119 to either an alanine or glycine completely abolished polymerase activity, and unfortunately crystallization of these mutants has been unsuccessful. However, circular dichroism studies show negligible changes to the secondary structure composition when compared to the native polymerase (data not shown). It is possible that *trans* amino acid isomers at position 119 alter the topology in that region such that the pinky finger interferes with crystal lattice contacts found in the $P6_5$ space group. Interestingly, an extensive structural review of a handful of proteins that use proline isomerization as molecular switches indicates that the *trans* conformer typically leads to loops with more extended, solvent-exposed conformations (Andreotti, 2003). The *cis* isomer typically leads to loops that are tucked or bent down towards the protein leading to more extensive interactions that serve to stabilize the slightly less stable isomer. The 3D^{pol} *cis*-proline 119 is also sandwiched in a π stack between two highly conserved tyrosines, which may further stabilize this conformation (Figure 6.3). This is consistent with data correlating the presence of aromatic amino acids preceding *cis*-prolyl bonds in proteins (Stewart et al., 1990; MacArthur and Thornton, 1991) and peptides (Grathwohl & Wuthrich, 1981; Yao et al., 1994; Reimer et al., 1998). The introduction of template nucleic acid could then provide the necessary interactions required to stabilize the *trans* Pro119, and a more extended pinky finger topology. Thus, Pro119 is essential for activity and

the total loss of activity suggests that the *cis* peptide bond conformation may be a required structural feature.

3D^{pol}-G117A mutant

Crystals grown of the 3D^{pol}-Gly117Ala mutant revealed disorder of the structure flanking this mutation (amino acids 112-129) (Figure 6.6). While an alanine side chain would not pose any steric clashes in the native structure, the backbone flexibility of a glycine appears to be required for function as this mutant polymerase is catalytically dead. Procheck analysis (Laskowski et al., 1993) of the native structure indicates that glycine 117 has the highest G-factor of all glycines, meaning that the backbone dihedrals are in an unfavorable conformation even for this flexible residue. Perhaps the role of Gly117 is not just to provide flexibility to accommodate the isomerization of Pro119, but to act as a spring where the energy stored in the unfavorable conformation decreases the energy barrier of proline isomerization.

3D^{pol}-G124A mutant

The low solubility (~50 μM) of the 3D^{pol}-G124A mutant suggested conformational differences from the native structure since this construct contains the L446A-R455D mutation, which normally yields highly soluble protein. This was confirmed by wavelength and thermal melt CD analysis (Figure 6.5). The wavelength scan (260 – 200 nm) performed at 20°C clearly indicated an altered secondary structure when compared to the native protein with high errors in the far-UV region (215 - 200 nm). However, the signal at 222 nm was consistent, and

monitoring this wavelength during a thermal melt indicated this mutant was very unstable (Figure 6.5). It is therefore unclear why the $3D^{pol}$-G124A mutant displayed such high RNA binding affinities (Table 6-2). It should be noted that the binding curve was not sigmoidal, and the amplitude of binding was not as high, possibly due to precipitated protein at the high concentrations used in this assay (see appendix 1). However, curve fitting of the data yielded a K_d suggesting 2-fold higher RNA affinity compared to the native protein (Table 6-2). It is possible that the $3D^{pol}$-G124A results in a more exposed/extended pinky finger conformation such that the RNA binding domain is more accessible. Unfortunately, this is only speculation since this mutant would not crystallize. Because high NaCl concentrations are inhibitory in activity assays, we typically employ a three-step dilution scheme to all proteins to systematically lower the concentration of both the salt and protein. The low solubility of this mutant required a customized dilution scheme and thus may account for the ~2 % activity.

$3D^{pol}$-G149A mutant

Glycine149 appears to be located at a hinge point if the entire pinky finger changes conformation as the specificity loop of T7 polymerase does during the conformational transition from initiation to elongation (Figures 6.1, 6.2, 6.6) (Tahirov et al., 2002; Yin and Steitz, 2002; Yin and Steitz, 2004). The B-factors and electron density of this region suggest that this glycine forms a boundary between flexible and inflexible residues (Figure 6.1). Inserting an Ile after residue 149, thereby shifting the register of the α-helix on the outside of the pinky,

completely inactivates $3D^{pol}$ activity and destroys its ability to function as an infectious cDNA during *in vitro* transcription assays (Burns et al, 1989). Residues 149 and 150 were also mutated to alanines, resulting in the complete elimination of plaques after viral transfection into HeLa cells (Diamond and Kirkegaard, 1994). We likewise found that mutating Gly149 to an alanine abolishes all activity, although there is only a slight decrease in its RNA binding affinity (Table 6.2). As discussed above, a low-resolution structure was obtained from a crystallite teased from a small unhealthy looking cluster of crystals that took a full month to grow. Initial electron density maps did not reveal significant conformational differences from the native structure, and therefore this structure was not extensively refined. Because of the odd nature of crystallization, we believe that only a small subset of molecules able to temporarily adopt the native-like conformation crystallized, thereby trapping this conformation in the crystal lattice.

Conclusions

We propose that *cis-trans* isomerization of Pro119 may be responsible for both the conversion to a stable elongation complex and subsequent processivity by "clamping" down onto substrate. In the absence of crystal structures showing both conformations, studying isomerization *per se* is difficult. While proline isomerization has long been recognized as the rate-limiting step in the folding of many proteins, nuclear magnetic resonance (NMR) spectroscopy is the ideal technique to study isomerization in a fully folded protein (Andreotti, 2003), which is beyond the scope of this study.

Structural analysis suggests that essential prolines often form tight turns on the surface of proteins and are found in the *cis* conformation (Wedemeyer et al., 2002). Prolines with the demonstrated ability to isomerize in solution are often found on or near small loops in close proximity with flexible glycines which facilitate localized conformational changes of these regions (Andreotti, 2003). Analysis of the six available Picornaviridae polymerase structures all show that proline 119 (Pro120 in FMDV) is in the *cis* conformation (Figure 6.7). These *cis* prolines are all flanked on both sides with glycines on structurally homologous loops that protrude into the putative RNA template entry channel. The conformations of these loops are extremely homologous. Although the Norwalk virus and rabbit hemorrhagic disease virus are in a completely different family (*Caliciviridae*), the polymerase structures from these viruses show similar loop conformations (Figure 6.8). Whereas the Norwalk virus polymerase proline 122 adopts the *cis* isomer, the rabbit hemorrhagic disease virus polymerase contains

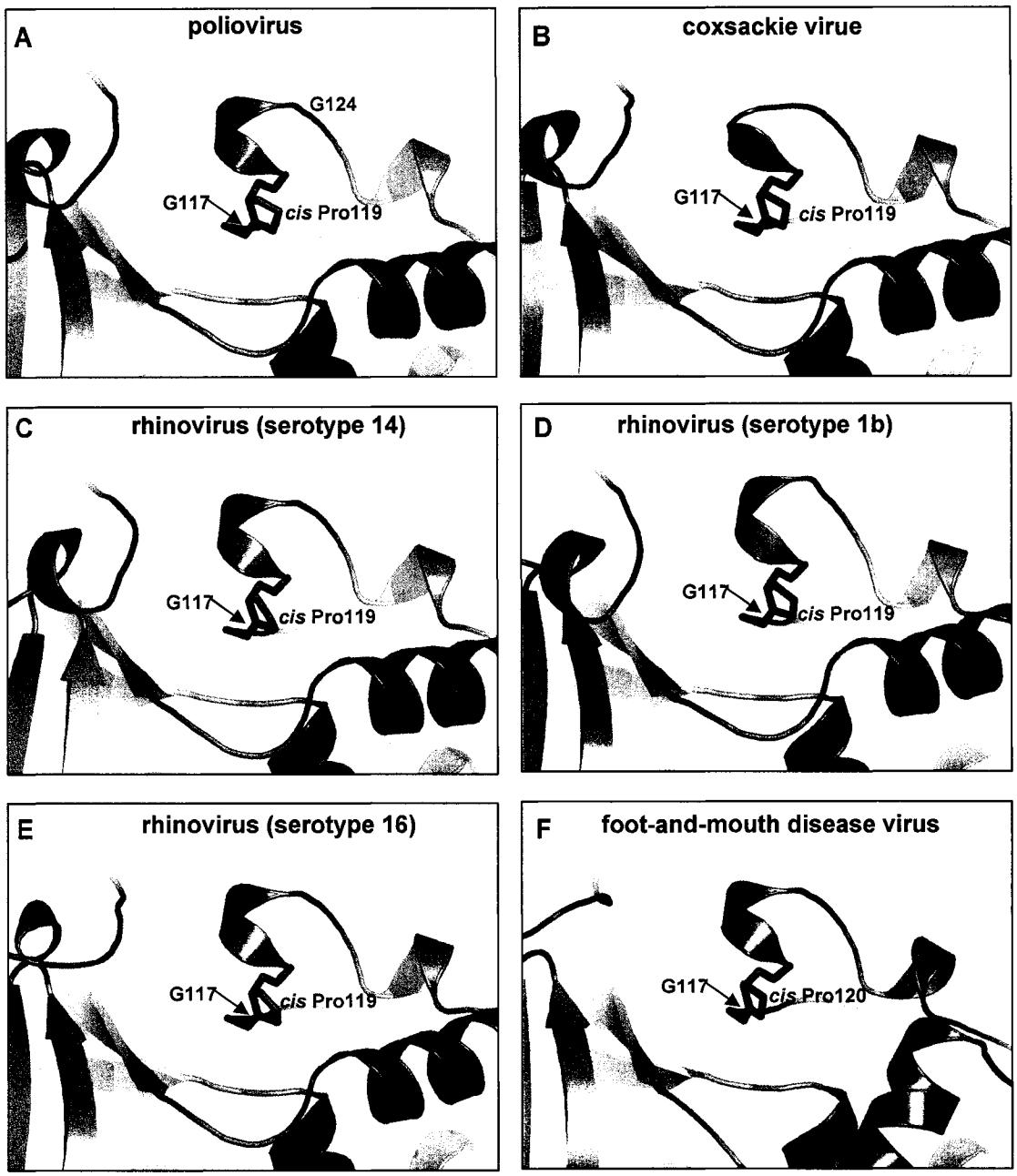


Figure 6.7 Structural Comparison of cis-prolines among Picornaviral Polymerases

Structures of the putative template entry channel from the six available picornaviral polymerases of (**A**) poliovirus, (**B**) coxsackie virus (**C**) rhinovirus (serotype 14), (**D**) rhinovirus (serotype 1b), (**E**) rhinovirus (serotype 16) (Love *et al.*, 2004), (**F**) FMDV (Ferrer-Orta *et al.*, 2004). Colored in cyan are conserved *cis* prolines flanked by glycines. These fixed views are from superpositions of the full polymerase

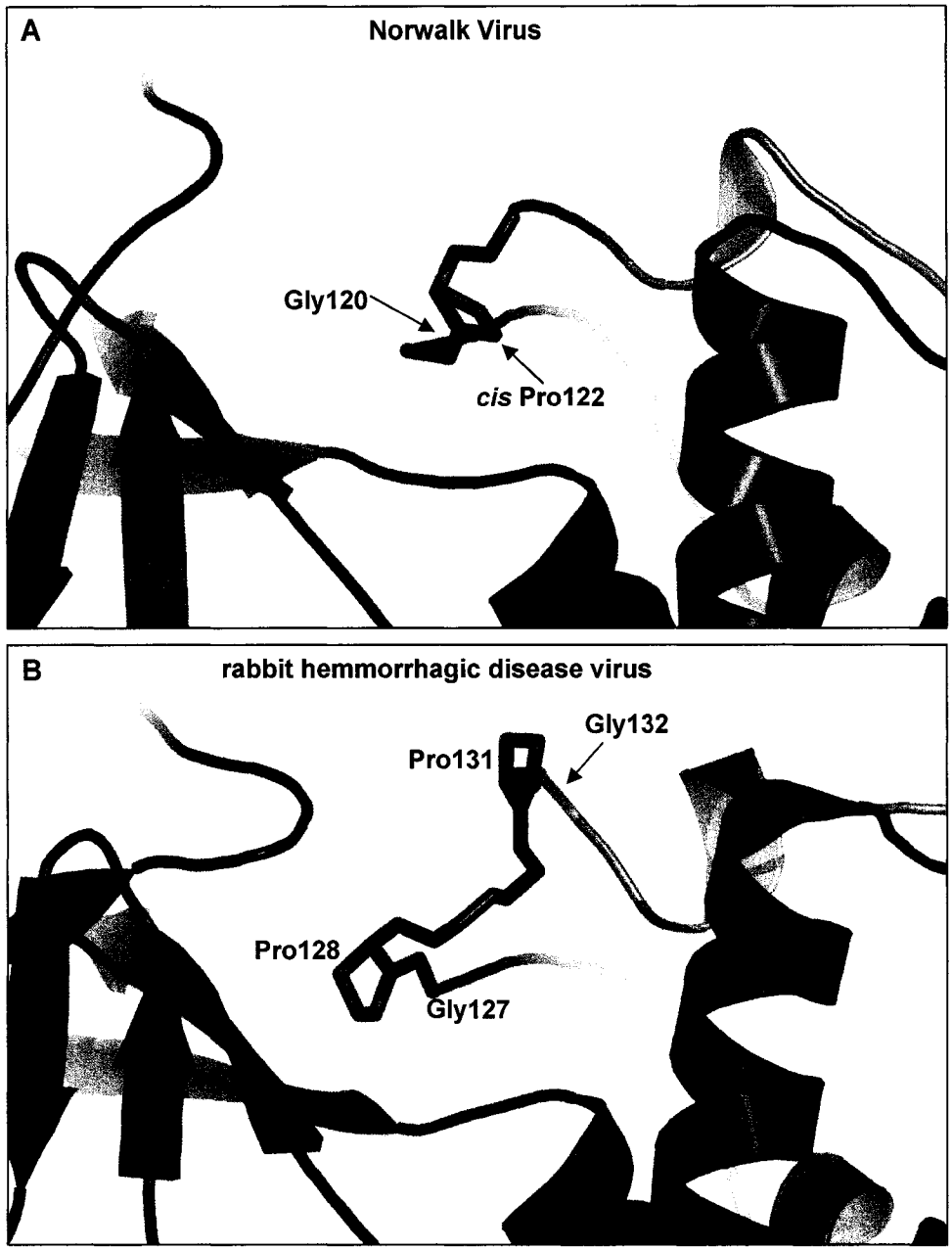


Figure 6.8 Structural Comparison of Prolines among Caliciviral Polymerases
Views of the putative template entry channel of the Caliciviridae polymerase structures from (**A**) Norwalk virus (Ng *et al.*, 2004) and (**B**) rabbit hemorrhagic disease virus (Ng *et al.*, 2002). The fingers are colored red, and the prolines and glycines that compose the putative isomerization motif are colored cyan.

two *trans* prolines Pro131 and Pro 128, and glycine residues are found in close proximity in both structures. Although the architecture of the Flaviviridae polymerases are quite different, these structures are proline-rich on the inside of the pinky fingers.

It is possible that these prolines are simply required for protein folding by promoting the association of residues which in turn steer and stabilize native folding pathways (Lewis et al., 1971). Or perhaps the proline-containing protuberance functions in helicase activity, which $3D^{pol}$ is known to possess (Cho et al., 1993). However there are many emerging examples of proteins that use proline isomerization as intrinsic molecular switches including: the bacteriophage MS2 coat protein (Valegard et al., 1990), Candida rugosa lipase (Grochulski et al., 1993; 1994), the transforming growth factor β-binding domain from human fibrillin-1 (Yuan et al., 1998), the SH2 domain from Interleukin-2 tyrosine kinase (Mallis et al., 2002), rabbit muscle adenylate kinase (Sheng et al., 1997) and the gene-3-protein of filamentous phage fd (Eckert et al., 2005). The high energy barrier of isomerization can also be lowered by prolyl *cis/trans* isomerases through favorable interactions with proline-containing motifs. These ubiquitous enzymes compose three different families that are involved in many aspects of proper cell function. Pin1, the prototypical representative from the parvulin family, has been shown to be important for the cellular progression through mitosis (Lu et al., 1996; Yaffe et al., 1997), pre-mRNA 3'-end formation by binding to the CTD of RNA polymerase II (Morris et al., 1999), and the regulation of Alzheimer's disease (Lu et al., 1999; Liou et al., 2003; Lu et al.,

2003). This growing family of prolyl *cis/trans* isomerases accomplish many diverse biological processes through the modulation of signal transduction (Wulf et al., 2005) in a phosphorylation-dependent manner (Shen et al., 1998; Yaffe et al., 1997).

The cyclophilins represent another well studied family of proline isomerases. Interestingly, cyclophilin A is found inside HIV-1 virons through interactions with the viral capsid protein (Thali et al., 1994; Luban et al., 1993). Mutations that prevent the inclusion of this isomerase into the HIV viron result in less infectious virus (Frake et al., 1994; Braaten et al., 1996a). There is a growing body of evidence that suggests proline isomerization of the HIV-1 capsid may be required for proper viral infection (Saphire et al., 2002; Gitti et al., 1996; Gamble et al., 1996; Howard et al., 2003; Bosco et al., 2002). Although PV $3D^{pol}$ has not been shown to interact with prolyl *cis/trans* isomerases, it does not preclude the possibility that proline 119 may isomerize in solution.

Comparison to other polymerase-nucleic acid complexes and mutational analysis of the pinky show the involvement of this region in binding template RNA. The high B-factors of this finger and the glycine-to-alanine mutations suggest the pinky loop, consisting of residues 117-149, are dynamic. The time scale for proline isomerization is consistent with the slow $3D^{pol}$ isomerization ($t_{1/2}$ ~ 12 s) to an elongation-competent state that has a half-life of ~2-8 hours (Arnold & Cameron, 2000). As mentioned above, the energy barrier for this conformational change may be further reduced by the "spring" energy stored in the native Gly117. While the exact structural role of Pro119 in the $3D^{pol}$

elongation complex remains to be determined, our data show that this residue is essential for activity and suggest that its ability to isomerize may be a required structural feature.

CHAPTER VII

THERMAL ANALYSIS OF THE FINGERS DOMAIN

An emerging structural feature of the viral RNA-dependent RNA polymerases is the presence of an "enclosed" active site (discussed in chapter IV). The enclosed active site is characterized by a direct connection between the fingers and thumb domains, leading to a large tunnel through the molecule. The poliovirus $3D^{pol}$ fingers domain has its own hydrophobic core that includes residues from the index, middle, and ring fingers and extends from the base of the fingers, across the top of the NTP entry tunnel, and over to the thumb. The $3D^{pol}$ structures shows that the index finger (residues 1-68, green in figure 2.7) is largely responsible for this domain linkage. In the original wild-type PV $3D^{pol}$ structure where the fingers domain was almost entirely disordered due to steric clashes in the crystal lattice, electron density was present for the index finger-tip (residues 12-35) showing that residues 22-35 are anchored to the thumb in an identical fashion to that of the full-length structure, highlighting the resilience of this interaction. This interaction appears to be dominated by the insertion of Phe30 and Phe34 into the hydrophobic core at the top of the thumb (see figure 7.1). Mutating these phenylalanines to alanines destroys both *in vitro* polymerase activity and *in vivo* viral infectivity (Hobson *et al.*, 2001). Likewise, a single

Figure 7.1 Structure of 3D^{pol} highlighting the Interdomain linkage between the Fingers and Thumb Domains

Back views of PV 3D^{pol} colored according to structural motifs described in Fig. 2.6. (**A**) Surface representation of the thumb showing a hydrophobic canyon in which Phe30 and Phe34 from the index fingertip are inserted. (**B**) Cartoon representation showing many of the hydrophobic and electrostatic interaction between the fingers and thumb domains.

mutation of Leu30 from the same region of HCV NS5B polymerase to the polar residues Arg or Ser was sufficient to inactivate polymerase activity (Labonte *et al.*, 2002) (Figure 7.7B). The ring finger (yellow in figures 2.7 & 7.1) interacts with both the thumb and index finger and may provide additional stability to the enclosed structure. More importantly however, these interactions correctly position the ring finger over the active site, allowing a number of highly conserved residues to interact with incoming rNTPs (discussed in chapter V).

Several investigators have observed an increase in the thermal stability of $3D^{pol}$ in the presence of ribonucleotides (Richards *et al.*, 1992; Arnold & Cameron, 2000). An increase in protein thermostability is common following ligand binding, and the denaturation temperatures (T_ms) often increase concomitantly with ligand concentration due to the coupling of binding with unfolding equilibrium (Celej *et al.*, 2003). Our laboratory has been able to measure a ~4 degree (C) increase in T_m in the presence of ribonucleotides (discussed in chapter V), and based on crystal structure analysis, we attribute this to rNTP-facilitated bridging interactions between the palm and the marginally stable fingers domain that result in stabilization of the structure.

The notion that the fingers are marginally stable stems from the original partial wild-type $3D^{pol}$ structure, which showed that the fingers could be selectively disordered while leaving the palm and thumb intact. The structure of a 68-residue N-terminal deletion mutant, whose fingers were also disordered, further showed that the entire index finger could be deleted without affecting the folding of the palm and thumb. Analysis of the complete structure reveals

average B-factors in the fingers are significantly higher than those in the palm and thumb (Figure 6.1 & Table 6-1), which is consistent with regions known to be involved with dynamic processes.

Here we use circular dichroism to show that an intact index finger is important for the thermal stability of the poliovirus polymerase. Perturbations of the index finger that prevent interdomain interactions between the fingers and thumb domains result in dramatic changes in the thermal melting profile and *in vitro* activity. These data suggest that the fingers domain of the poliovirus polymerase is inherently unstable when compared to the palm and thumb, and that the enclosed domain structure plays a significant role in enhancing this stability. The discovery and characterization of efficacious non-nucleoside inhibitors that target the interface between the fingers and thumb of the HCV polymerase, thereby preventing native interaction between these domains, highlights the relevance of this study (Tomei *et al.*, 2003; Di Marco *et al.*, 2005).

Results

Full-length versus Δ-68 N-terminus Truncation

In the course of examining the structural changes dependent on mutations at the N-terminus, we found a significantly different thermal melting profile when comparing full-length $3D^{pol}$ to the 68-residue N-terminal truncation (Figure 7.2). All thermal melting data were acquired by circular dichroism at a wavelength of 222 nm. The observed T_m for full-length polymerase was approximately ~39°C with a very sharp cooperative transition. At temperatures above 45°C, the protein precipitated and the sample appeared as large white flocculent "flakes". The same observations were made for mutants containing mutations of the very N-terminus. Interestingly, two distinct melting transitions located at ~33.5°C and ~76°C were observed when 68 residues from the N-terminus are deleted ($3D^{pol}$-Δ68 construct) (Figure 7.2). At temperatures above 75°C, this truncation mutant appeared precipitated with large white flakes in the sample, but at the plateau between 35°C and 65°C, the $3D^{pol}$-Δ68 mutant appeared only slightly turbid. Based on wavelength scans collected at 50 and 60°C, this intermediate appeared to have a radically different secondary structure compared to the scans performed at 20°C (Figure 7.3). Although thorough computer analysis of these changes has not been done, this intermediate clearly has significant β-sheet content. This change in structure is not reversible by cooling the protein from 50°C back down to 20°C (data not shown). Additionally, centrifugal clarification of the $3D^{pol}$-Δ68 truncation after incubation at 55°C resulted in too little protein for any subsequent biophysical analysis.

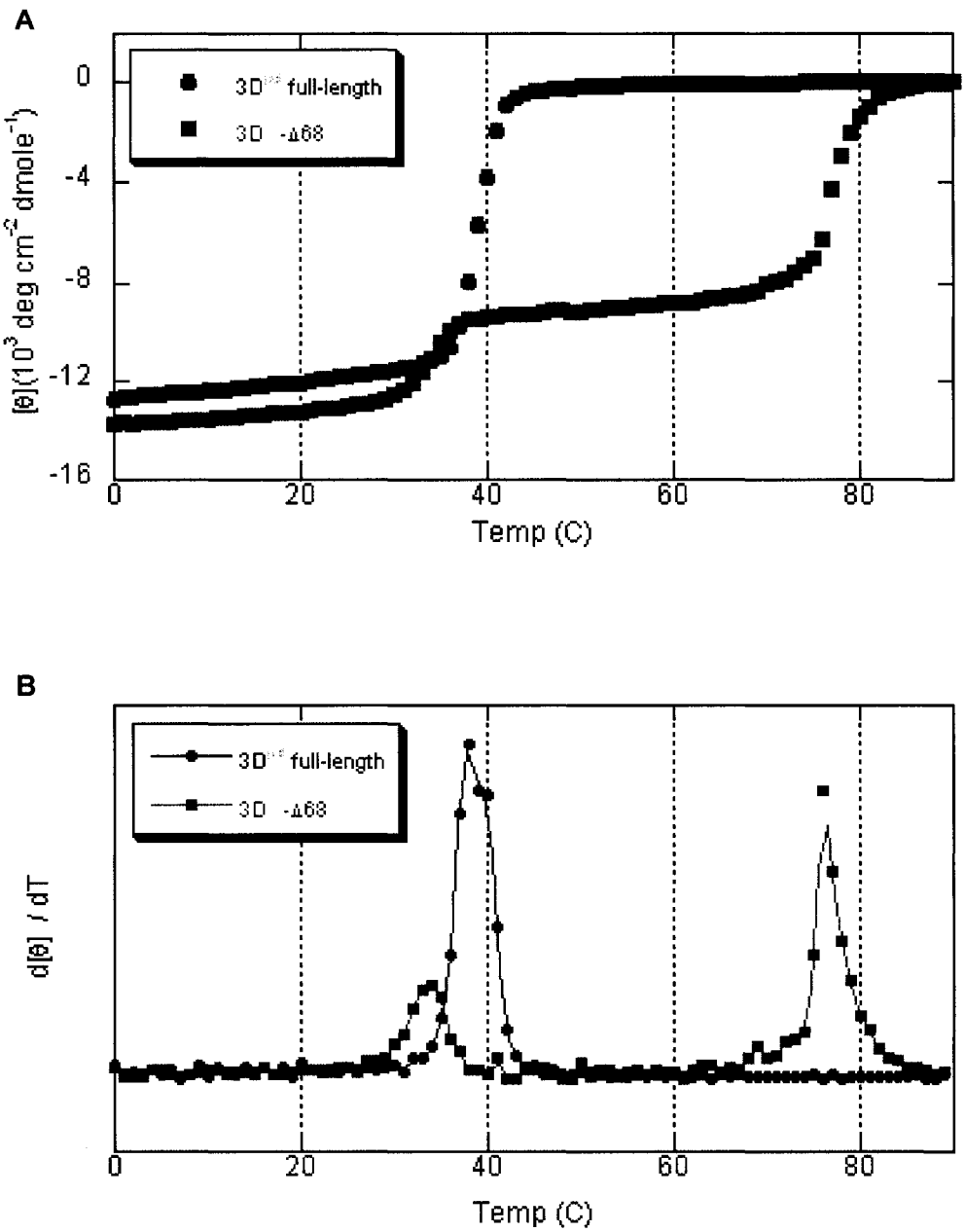


Figure 7.2 CD Thermal Denaturation Data for native $3D^{pol}$ and $3D^{pol}$-Δ68
(**A**) CD thermal denaturation for native $3D^{pol}$ (blue circles) and the $3D^{pol}$-Δ68 mutant (red squares). The thermal denaturation spectra were obtained at a wavelength of 222 nm. (**B**) d[θ] vs. dT plot of data in panel (A) where the interpolation maxima indicate the T_m values.

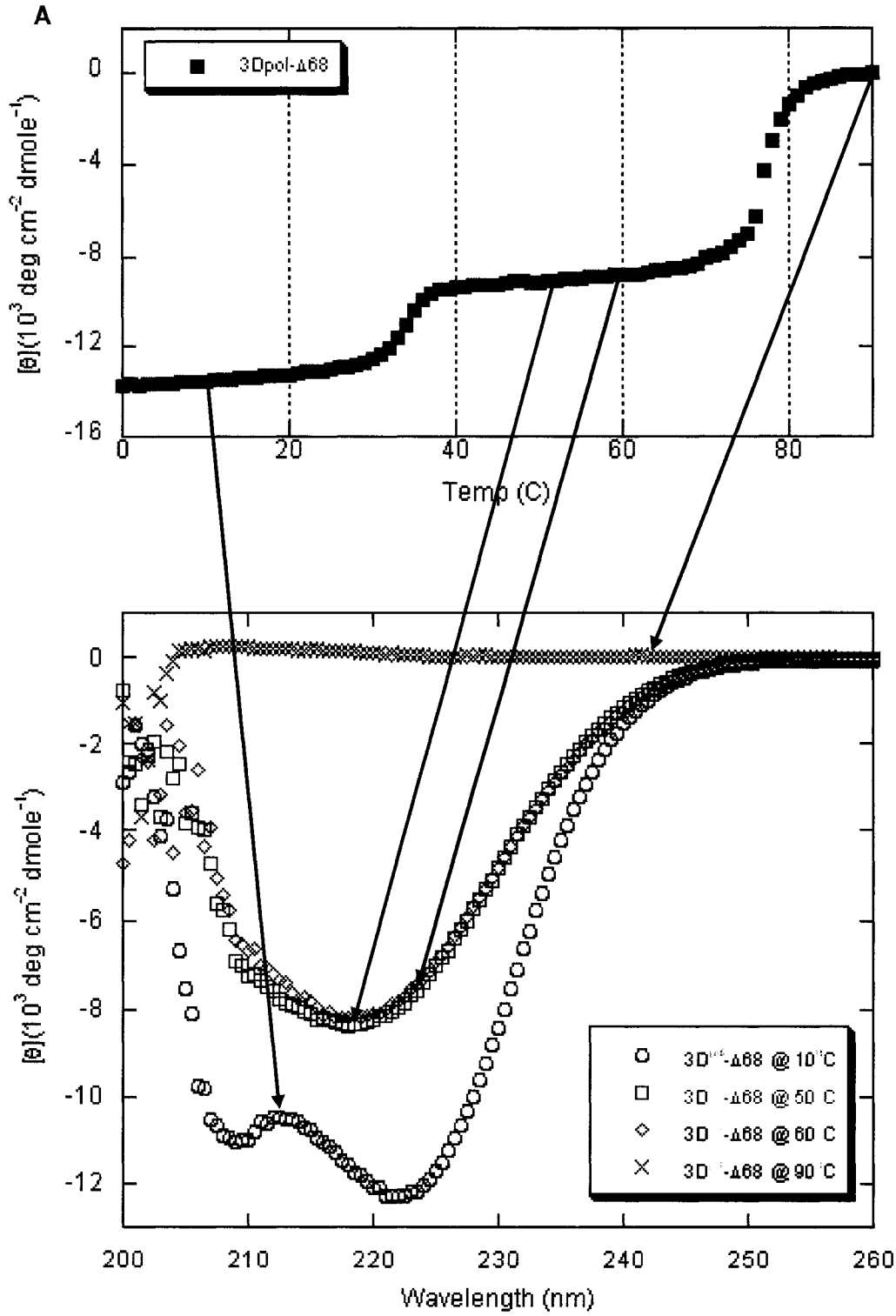


Figure 7.3 CD Data for 3D^{pol}-Δ68

(**A**) CD thermal denaturation for the 3D^{pol}-Δ68 mutant (red points; also figure 7.2) obtained at a wavelength of 222 nm. (**B**) Wavelength CD data of the 3D^{pol}-Δ68 mutant performed at 10°C (open red circles), 50°C (open blue squares), 60°C (open green triangles), and 90°C (black Xs)

In contrast with previously published data (Hobsen *et al.*, 2001), we observe a dramatic decrease in $3D^{pol}$'s ability to bind a duplex RNA construct containing an 8-base pair duplex flanked by 4-nucleotide 5' overhangs when the index finger is removed. By measuring the fluorescence polarization of fluorescein attached to both 5' ends of this construct, we measure more than a 3.5-fold decrease in the binding affinity when the index finger is truncated ($3D^{pol}$-L446D-R455D versus $3D^{pol}$-Δ68-L446D-R455D) (Table 7-1). Mutation of Leu446 to an alanine ($3D^{pol}$-Δ68-L446A-R455D) mitigates the loss of RNA binding affinity with a 2-fold decrease in the K_d (Table 7.1). It should be noted that the aforementioned $3D^{pol}$-Δ68-L446A-R455D construct is the only mutant in this dissertation containing an alanine at position 446. It was included in this analysis since we were able to crystallize this construct (discussed in chapter II), although the structures and CD analysis show no structural differences dependent on amino acid 446 (data not shown). All N-terminal truncation mutants were completely dead in elongation assays (Table 7-1) due to the requirement for the native Gly1 to allosterically position Asp238 in the active site.

$3D^{pol}$-Δ22 N-terminus Truncation

Deletion of the first 22 residues from the N-terminus ($3D^{pol}$-Δ22) results in an enzyme with half of an intact index finger, but preserves the interaction with the thumb domain and a majority of the ring finger. Like the $3D^{pol}$-Δ68 constructs described above, the $3D^{pol}$-Δ22 truncation exhibited a biphasic melting transition (Figure 7.4). The T_m of the first transition is approximately ~37°C and this

Mutation	RNA binding Kd	% Activity	~T_m1	~T_m2
L446D; R455D	1.75 ± 0.05 X 10^{-6} M	100 ± 4	39	N/A
All Mutants below contain the L446D-R455D mutation				
$3D^{pol}$-D22	3.4 ± 0.2 X 10^{-6} M	-0.3 ± 0.7	37	50
$3D^{pol}$-D68-AD	3.8 ± 0.4 X 10^{-6} M	0.0 ± 0.2	34	76
$3D^{pol}$-D68-DD	6.5 ± 0.4 X 10^{-6} M	N/A	34	76
$3D^{pol}$-D68-380	0.15 ± 0.02 X10^{-6} M	-0.3 ± 1.2	37	75
$3D^{pol}$-F30A	N/A	4.3 ± 2.1	N/A	N/A
$3D^{pol}$-F34A	1.70 ± 0.03 X 10^{-6} M	4.3 ± 1.5	N/A	N/A
$3D^{pol}$-F30A-F34A	2.2 ± 0.1 X 10^{-6} M	-0.3 ± 0.7	38	N/A
$3D^{pol}$-F30D-F34D	3.0 ± 0.1 X 10^{-6} M	2.0 ± 1.0	35	57
$3D^{pol}$-403A	1.7 ± 0.1 X 10^{-6} M	2.0 ± 0.2	34.75	44.5
$3D^{pol}$-403D	2.3 ± 0.2 X 10^{-6} M	1.3 ± 0.6	33-34	39-40

Table 7-1 RNA binding K_ds and Polymerase Activity and appriximate T_ms of $3D^{pol}$ index Finger and Thumb Mutants that disrupt the Interdomain Linkage
The RNA binding data was acquired by fluorescence polarization; activity data from poly(A)/oligo d(T) extension assays; T_m data from circular dichroism thermal denaturation studies obtained at a wavelength of 222nm. The experimental methods used to acquire the data is described in Appendix 1.

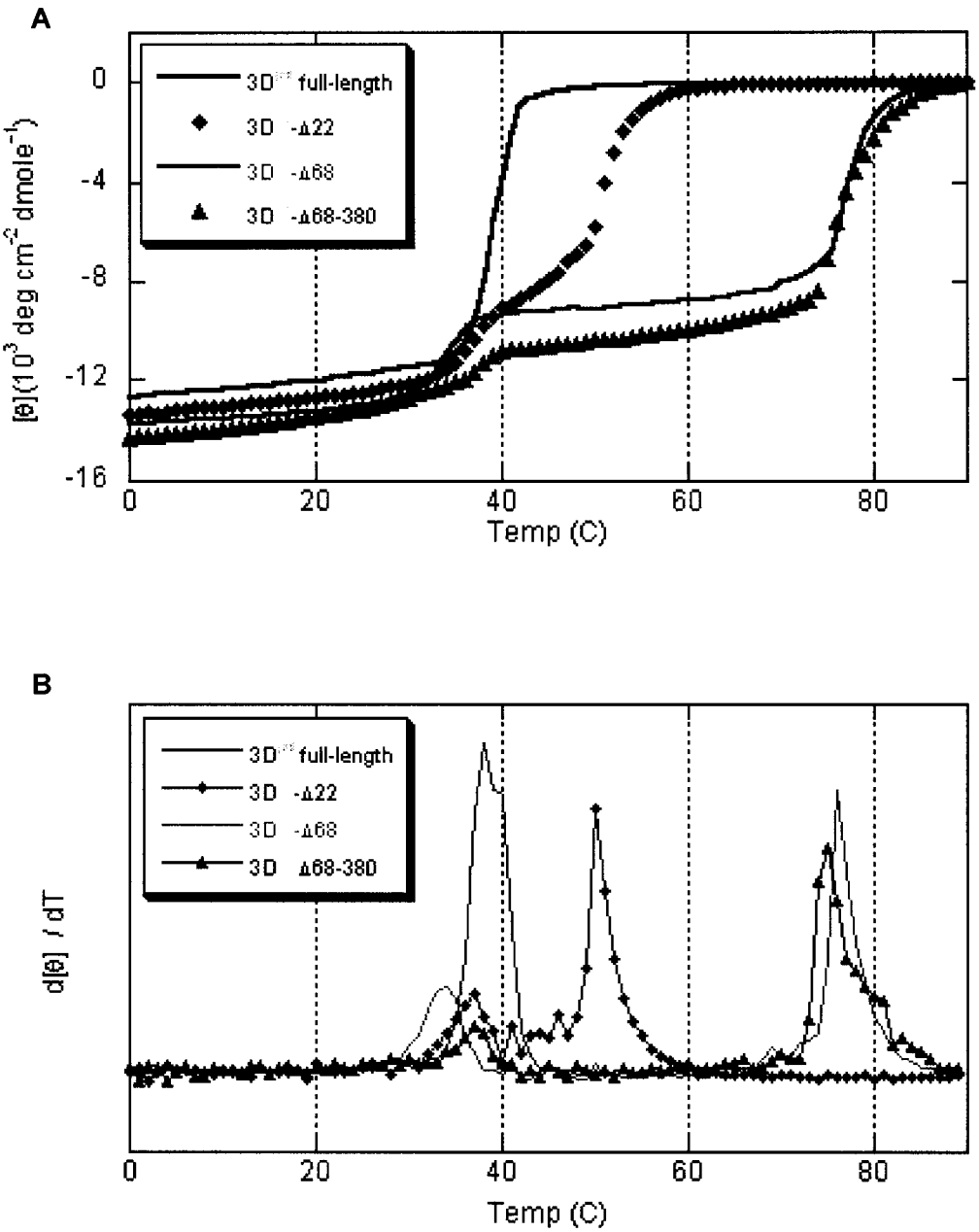


Figure 7.4 CD Thermal Denaturation Data for 3D^{pol} Truncation Mutants
(**A**) CD thermal denaturation for 3D^{pol}- Δ22 (green diamonds) and 3D^{pol}-Δ68-380 (black triangles). The native 3D^{pol} and 3D^{pol}-Δ68 from Fig 7.2 are depicted with a blue and red line respectively. The thermal denaturation spectra were obtained at a wavelength of 222 nm. (**B**) d[θ] vs. dT plot of data in panel (A) where the interpolation maxima indicate the T_m values.

unfolding transition appears to closely mimic that of the full-length protein. However, the intermediate structure of this truncation appears less stable when compared to the 68-residue N-terminal deletion and the second transition has a T_m of approximately ~50°C. Like the Δ68 truncation, deleting the first 22 residues of the protein results in a large 2-fold decrease in RNA binding affinity and enzymatic activity is abolished (Table 7-1)

$3D^{pol}$-Δ68 - 380 Truncation

The presence of biphasic melting transitions of index finger truncations not only indicates that the index finger plays a significant role in the thermal stability of the protein, it also suggests that different domains of the polymerase melt at different temperatures. In the attempt to assign T_ms to distinct domains, we have constructed various truncations of the different domains for CD analysis. Unfortunately, deleting just the thumb from the full-length construct, or cloning just the thumb domain yielded constructs that did not express. However, we were able to express a $3D^{pol}$ construct at low levels with a deletion of both the index finger (residues 1-68) and the thumb (residues 381-461). While we were unable to obtain highly pure crystallographic quantities of this mutant, the CD thermal melt exhibited thermal unfolding transitions similar to that of the $3D^{pol}$-Δ68 construct. The observed T_ms of this construct were approximately ~37°C and ~75°C.

Not surprisingly, deleting the thumb and index finger resulted in a catalytically dead polymerase. However, the extremely low K_d for binding RNA

did come as a surprise, but the quality of this finding is questionable due to the odd behavior of this truncated construct. The binding curve was not sigmoidal when the log of the protein concentration was plotted on the X-axis (raw data not shown), and it appeared that this mutant was precipitating at higher concentrations.

Mutations of the Index fingertip

To correlate thermal stability with residues involved in the interdomain linkage, we also tested the melting profile of full-length polymerase constructs with mutations located at the index fingertip. We mutated the hydrophobic phenylalanines that insert into the thumb to alanines and aspartic acids with the following constructs: $3D^{pol}$-F30A; $3D^{pol}$-F34A; $3D^{pol}$-F30A-F34A; $3D^{pol}$-F30D-F34D. While single point mutations of phenylalanines 30 and 34 to alanines yielded polymerases with thermal stability almost identical to that of the native polymerase (data not shown), the double phenylalanine to alanine mutant ($3D^{pol}$-F30A-F34A) displayed a single transition with slightly less cooperativity and a T_m approximately 1 degree lower that the native structure (Figure 7.5). Interesting and dramatic differences were observed with the $3D^{pol}$-F30D-F34D mutant protein. Mutation of these phenylalanines to negatively charged aspartate residues strongly disfavors insertion of this region into the thumb, and the result is a biphasic melting profile more reminiscent of the $3D^{pol}$-Δ22 construct with T_ms of ~35°C and ~57°C. Despite the complete absence of activity in elongation assays, the ability of these mutants to bind RNA remained intact with K_ds comparable to that of the native enzyme (Table 7-1).

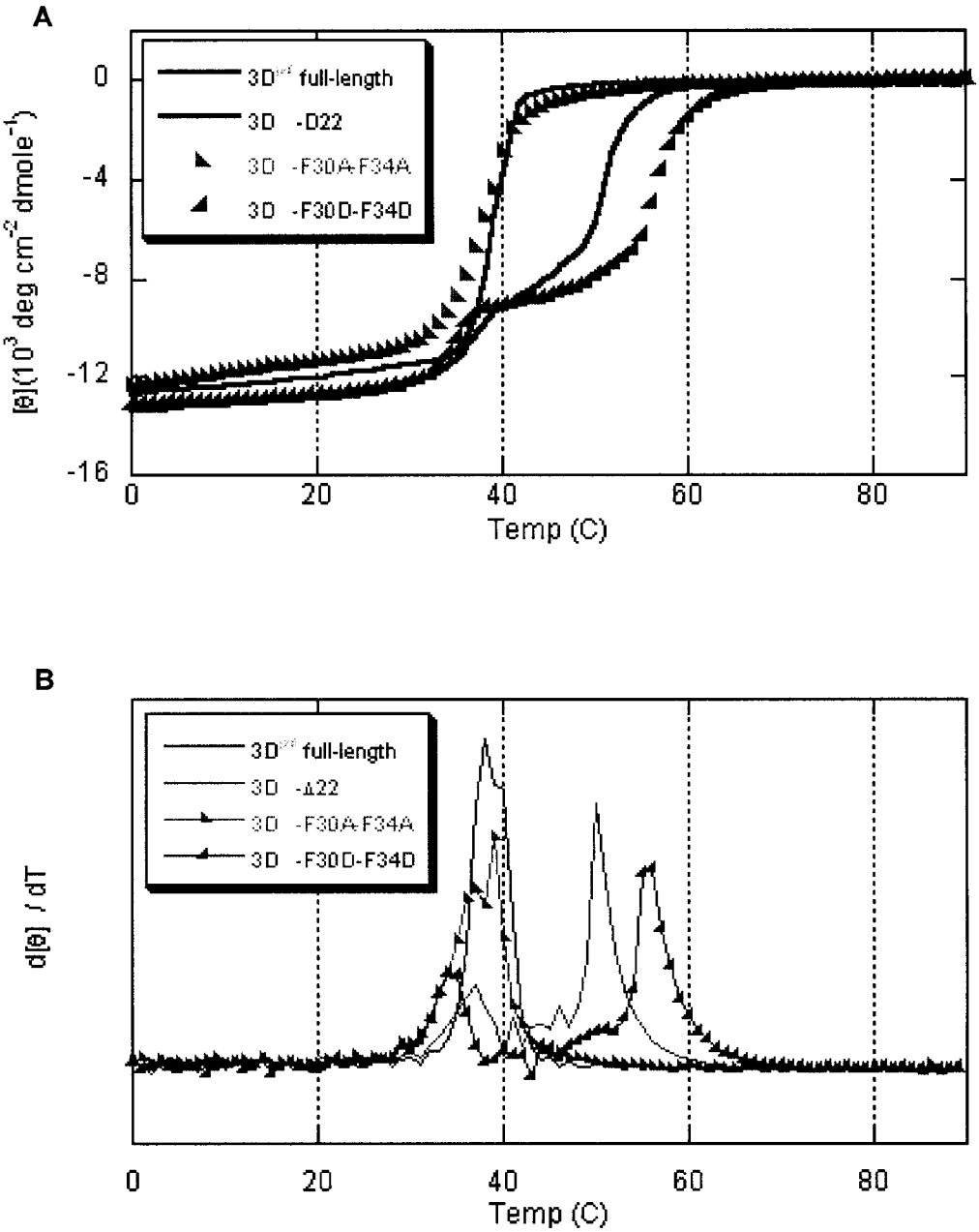


Figure 7.5 CD Thermal Denaturation Data for 3D^{pol} Index Fingertip Mutants
(**A**) CD thermal denaturation for 3D^{pol}-F30A-F34A (orange triangles) and 3D^{pol}-F30D-F34D (brown triangles). The native 3D^{pol} and 3D^{pol}-Δ22 proteins are depicted with a blue and green line respectively. The thermal denaturation spectra were obtained at a wavelength of 222 nm. (**B**) d[θ] vs. dT plot of data in panel (A) where the interpolation maxima indicate the T_m values.

Mutation of the Thumb-tip

Structural comparison of the complete full-length structure (PDB code 1RA6) with that of the $3D^{pol}$-Δ68 structure (PDB code 1RAJ) reveals a relatively large conformational difference at the tip of the thumb (discussed in chapter II) (see figure 2.4). In particular, Trp403 undergoes a significant change in orientation between these two structures which is dependent on the presence or absence of an intact index finger. One caveat of this observation is the fact that Trp403 participates in a crystal lattice contact in the $P3_121$ space group of the $3D^{pol}$-Δ68 structure. Nonetheless, the ability of Trp403 to adopt an alternate conformation demonstrates flexibility at the top of the thumb domain. Tryptophan 403 appears to play an important role in the native structure where it is surrounded by several conserved hydrophobic residues of the index finger (Figure 7.1B). To test the role of this conserved hydrophobic residue in stabilizing the enclosed structure we mutated Trp403 to both an alanine and aspartic acid. The $3D^{pol}$-W403D mutant displayed a biphasic melting transition with T_ms centered around ~35°C and ~45.5°C (Figure 7.6). The $3D^{pol}$-W403A mutant had an even tighter melting profile and surprisingly, the first T_m is lower that the $3D^{pol}$-W403D mutant (Figure 7.6). This result is reproducible with approximate T_ms of between 33 - 34°C and 39 - 40°C for the two transitions (two melts shown for $3D^{pol}$-W403A mutant; Figure 7.6). While both of these Trp403 mutants are almost completely dead catalytically, the W403A construct binds RNA with comparable affinity of the native polymerase (Table 7-1). The W403D mutant shows a slight 1.3-fold loss in RNA binding affinity.

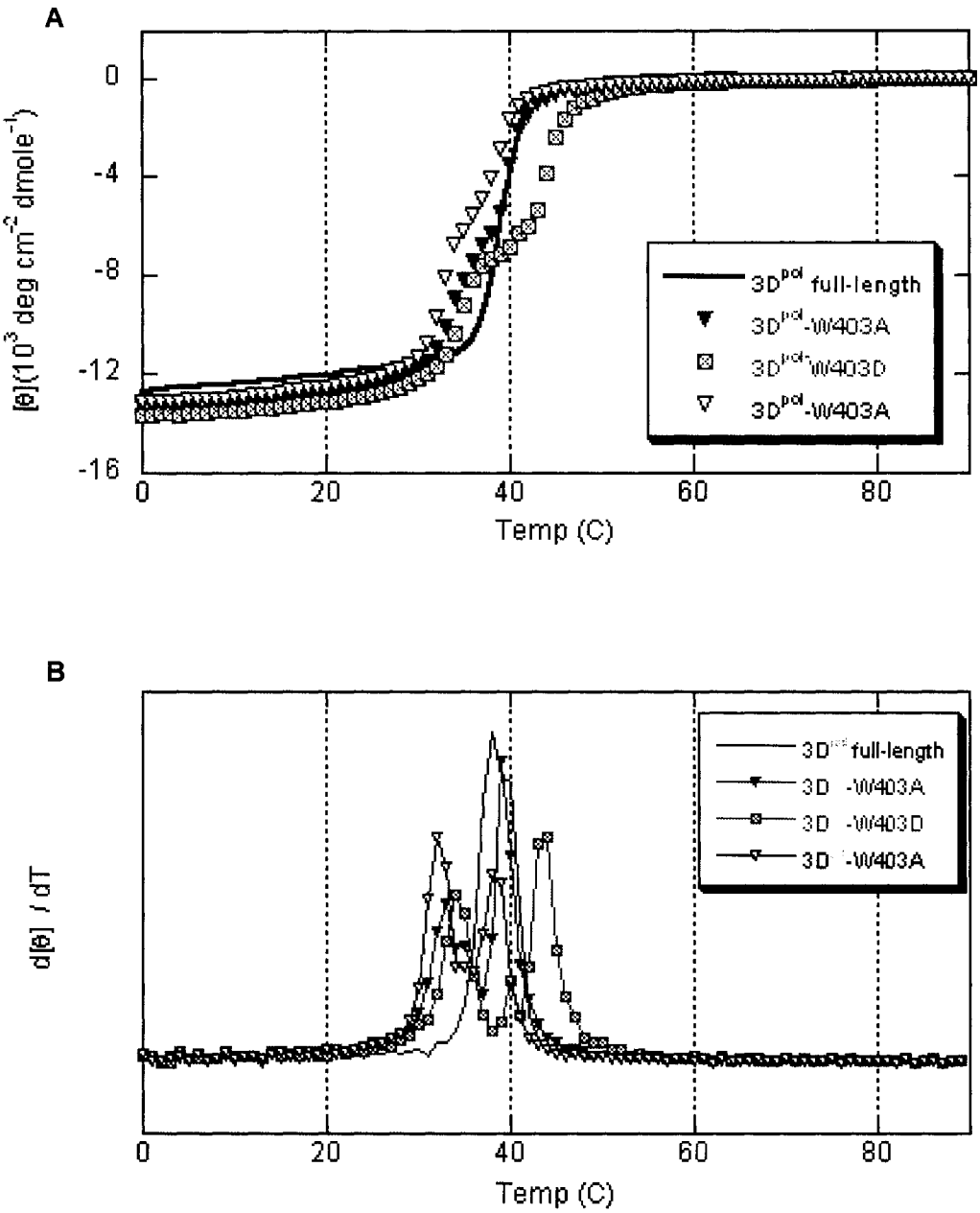


Figure 7.6 CD Thermal Denaturation Data for 3D^{pol}-W403 mutants
(**A**) CD thermal denaturation for 3D^{pol}-W403A (grey triangles) and 3D^{pol}-W403D (purple hatched square). The native 3D^{pol} protein is depicted with a blue line. The thermal denaturation spectra were obtained at a wavelength of 222 nm. (**B**) d[θ] vs. dT plot of data in panel (A) where the interpolation maxima indicate the T_m values.

Discussion

3D^{pol} Truncations

Truncations of the index finger obviously play a significant role in altering the thermal stability of 3D^{pol}. Deleting this entire finger results in two distinct thermal melting transitions flanking an intermediate that appears to be relatively stable between 40°C and 70°C. When the first 22 residues from the N-terminus are deleted from the enzyme, thereby preserving the interaction with the thumb domain and a significant portion of the ring finger, the result is a much tighter biphasic melting profile due to a decrease of the second T_m. Comparing the first transition between these two truncations show that the removal of the entire index finger results in a more significant decrease of the first Tm (drop of ~ 3°C, Figure 7.4, Table 7-1). The unfolding of the 3D^{pol}-Δ22 truncation through the first transition almost mirrors the full-length protein suggesting that residues 23-68 of the index finger partially stabilize the fingers domain. Interpretation of the second transition is more difficult because the protein is precipitating instead of equilibrating between folded and unfolded states. However, we suspect that the exposure of hydrophobic residues from the fingers domain may be responsible for nucleating precipitation of the entire protein. The stable intermediate of the 3D^{pol}-Δ68 structure may exist through a wider temperature range due to the removal of hydrophobic residues near the N-terminus (i.e. Trp5) and the index finger-tip (Phe30, Val33, Phe34 and others). Because the 3D^{pol}-Δ22 truncation has an intact index fingertip containing several hydrophobic amino acids, the unfolding of this region may result in a less stable intermediate. Following this argument suggests

that the unfolding of hydrophobic residues distributed on the index finger may trigger the full-length protein to precipitate following initial unfolding of the fingers domain.

The biphasic nature of the melting profiles described above suggests different domains of the polymerase exhibit differences in thermal stability. Unfortunately, I was only able to obtain a $3D^{pol}$-Δ68-380 construct for CD analysis, which did not contain the thumb or index finger. The melting of this construct vary closely resembles the $3D^{pol}$-Δ68 truncation suggesting that the second transition is an unfolding of the conserved palm domain (Figure 7.4).

Disruption of Finger-Thumb Interactions

It has been shown that mutations of PV and HCV polymerase index fingertip that preclude the full complement of interactions with the thumb have deleterious *in vivo* and *in vitro* consequences (Hobson *et al.*, 2001; Labonte *et al.*, 2002). We have found similar results with respect to *in vitro* activity when Phe30 and Phe34 are mutated (Table 7-1). Although the $3D^{pol}$-F30A mutant was not tested, it is unlikely that this is a result of decreased RNA binding affinities as the $3D^{pol}$-F34A and $3D^{pol}$-F30A-F34A mutants did not have K_ds which were significantly different from the native enzyme (Table 7-2). While the $3D^{pol}$-F30A-F34A mutant appears to cause a modest decrease in thermal stability, mutation of these residues to charged aspartates results in a biphasic melting profile more reminiscent of the $3D^{pol}$-Δ22 construct. The drop in T_m of the first transition is consistent with destabilization of the fingers domain; however, the

increase in T_m of the second transition when compared to the $3D^{pol}$-Δ22 construct is not fully understood. It is possible that mutation of phenylalanines 30 and 34 to hydrophilic aspartates leads to an increase in the solubility. Because the $3D^{pol}$-F30D-F34D construct contains hydrophobic residues from the N-terminus of the protein (residues 2-9), the second thermal transition will not occur at the higher temperatures observed with the $3D^{pol}$-Δ68 constructs.

Tryptophan 403, located at the top of the thumb, appears to function as a hook that aids in latching the index finger to the thumb (Figure 7.1). This residue, which is 100 % conserved as an aromatic hydrophobic side chain, interacts with several amino acids from the index finger that are also highly conserved as hydrophobics among picornaviruses (Ann Palmenberg, personal communication; Hobson, 2000). Mutating Trp403 to either an alanine or aspartate results in tight biphasic melting transitions. Both of these mutations show a decrease of the first T_m, when compared to the native polymerase, suggesting that removal of the "hook" disrupts the interdomain interactions causing a decrease of the thermostability of the fingers domain. Similar to the results of mutating the index fingertip, mutating Trp403 to a charged aspartate appears to increase the solubility of the intermediate such that the second transition of this mutant is higher than the T_m of the native protein.

Conclusions

Polymerases are enzymes with mechanical characteristics defined by their architecture. All RdRP structures to date show a unique topology where the fingers domain is intimately connected to the top of the thumb domain thereby creating an enclosed structure (Figure 4.1). While little is known about the dynamics of these polymerases, all reports presenting RdRP structures suggest that these are unusually rigid (with the exception of the priming platform). This is supported by the structures of several RdRPs complexed with nucleic acid showing negligable conformational changes after binding (O'Farrell *et al.*, 2003; Butcher and Grimes *et al.*, 2001; Ferrer-Orta *et al.*, 2004). In the preceding chapter (VI), we propose that the pinky finger may play a functional and dynamic role in the proper functioning of $3D^{pol}$.

In this chapter, we have shown that the interdomain linkage is important for the thermal stability $3D^{pol}$. Truncations of the index finger and mutations that interfere with interdomain interactions between the fingers and thumb domains result in dramatic changes in the thermal melting profile and *in vitro* activity. These data suggest that the fingers domain of the poliovirus polymerase is inherently unstable when compared to the palm and thumb, and that the enclosed domain structure plays a significant role in enhancing this stability. This is supported with ribonucleotide binding data showing an increase in thermostability of the protein due to bridging interactions between the palm and fingers domain (discussed in chapter V).

The ring finger, which forms the roof of the NTP entry tunnel, may benefit the most from the enclosed active site via the numerous hydrophobic and ionic stabilizing interactions with the index finger (Figure 7.1B). Correct positioning of the ring finger over the active site is important because it allows a number of highly conserved residues to interact with incoming rNTPs (discussed in chapter V; Figures 5.1 & 5.2). In fact, several of the residues from the ring finger known to be important for binding rNTPs are in direct contact with the index finger including E161, R163 (Figures 7.1B, 5.1, 5.2). It is therefore plausible that the index finger anchored to the top of the thumb acts as a structural support platform for the ring finger, and possibly the entire fingers domain. Support for this comes from the recently published crystal structures of the HCV polymerase complexed with non-nucleoside benzimidazole-based allosteric inhibitors bound to the thumb where the A1-loop (index fingertip) normally interacts in the native apostructure (Di Marco *et al.*, 2005) (Figure 7.7). Soaking these drugs into NS5B crystals causes displacement and disorder of the index finger (A1-loop, residues 22-35), which in turn leads to disorder of the ring fingertip (A2-loop in NS5B, residues 148-152) due to the elimination of stabilizing interactions. Interestingly, it was pointed out by the authors of this study that crystals could only be soaked with the drug for a relatively short amount of time, leaving open the possibility that larger domain rearrangements of the fingers could be constrained by crystal packing interactions.

It has been suggested that the NS5B interdomain linkage may coordinate structural changes of the thumb and fingers associated with the various steps of

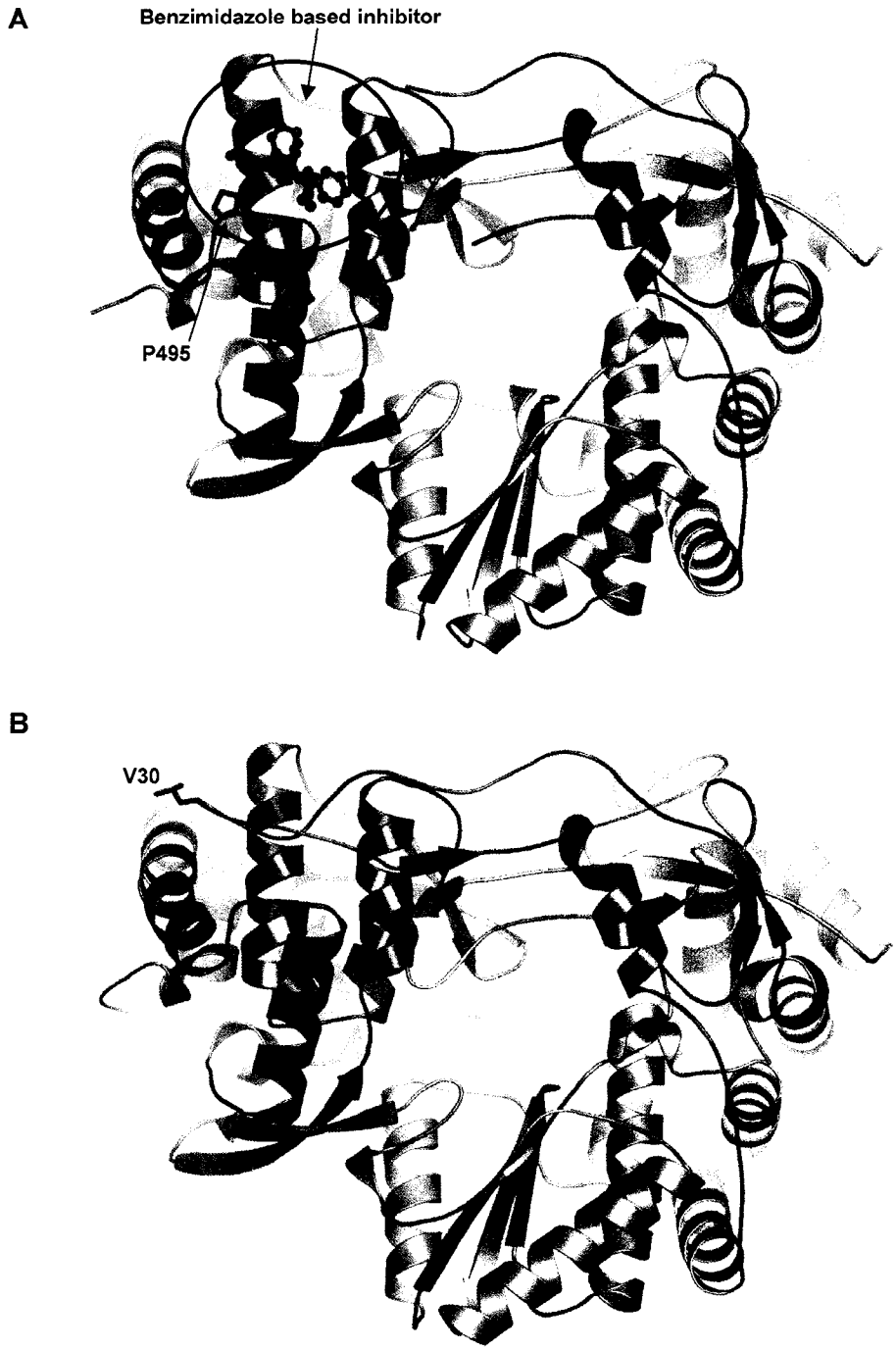


Figure 7.7 Structures of HCV NS5B complexed with Benzimidazole Drugs
Back view of the HCV NS5B polymerase colored according to the PV $3D^{pol}$ structural motifs described in Fig. 2.7 and the priming platform is colored brown.
(**A**) polymerase structures complexed with a benzimidazole-based viral inhibitor bound the the top of the thumb where the index fingertip normally interacts (Di Marco *et al.*, 2005). Drug binding results in disorder of both the index and ring fingertips. (**B**) Apo-structure structure showing the native topology of the index and ring fingertips Biswal *et al.*, 2005). V30 at the tip of the index finger is important for activity and the conformation of the polymerase.

nucleotide polymerization (Labonte *et al.*, 2002). Structures with bound benzimidazole drugs suggest that the interdomain interaction is conformationally flexible to allow access of the inhibitors to the otherwise buried site (Di Marco *et al.*, 2005). While adding increasing amounts of template-primer RNA to the polymerase-drug complex had no affect on inhibition *in vitro*, addition of the allosteric drugs to preformed enzyme-RNA complexes prevented efficacy of the inhibitors even at very high drug concentrations (Tomei *et al.*, 2003). This finding clearly suggests that RNA binding to the polymerase protects this enzyme from the action of the inhibitor. Ligand binding has been correlated with a reduction in conformational flexibility and an increase thermal stability (Celej *et al.*, 2003). Perhaps RNA binding to the fingers stabilizes this domain, similar to the proposed action of ribonucleotides, such that the benzimidazole inhibitor binding site is made inaccessible. Interestingly, rGTP was shown to interfere with the efficacy of the benzimidazole drugs in a concentration-dependent manner (Tomei *et al.*, 2003). The isolation of P495L drug-resistant mutants (Tomei *et al.*, 2003), and the preexistence of HCV polymerase structures bound with rGTP on the thumb near Pro495 (Figure 4.6; discussed in chapter IV; Bressanelli *et al.*, 2002) led investigators to hypothesize competitive overlapping of these binding sites. However the recent NS5B structures bound with these allosteric inhibitors clearly indicate distinct binding sites, and instead it was suggested that the antagonistic effect is due to thumb-rGTP-index finger cross-linking interactions that disfavor opening of the index fingertip (Di Marco *et al.*, 2005).

Leucines 30 and 31 from the HCV NS5B index fingertip only partially fill the pockets occupied by substituents of the benzimidazole inhibitors, leading to the proposal that the sub-optimal anchoring of fingers to the thumb might provide the conformational flexibility required during different steps of RNA synthesis (Di Marco *et al.*, 2005). Elution profiles from heparin-sepharose column chromatography and analytical ultracentrifugation experiments indeed show that mutation of Leu30 results in conformational differences of the fingers domain (Labonte *et al.*, 2002). The six homologous polymerase structures from the *Picornaviridae* family all show more extensive hydrophobic and electrostatic interactions connecting these domains. Although we have shown that residues responsible for the interdomain interaction are required for activity and thermal stability of PV $3D^{pol}$, it remains to be determined if this interaction is in equilibrium between bound and unbound states in solution. The two *Caliciviridae* polymerase structures (Norwalk virus & rabbit hemorrhagic disease virus) show even more fingers-thumb interdomain interactions (Figure 7.8). The thumb domains of these structures have large loops that wrap around the index finger. Although the multitude of electrostatic interactions are not shown in this figure (7.8) for purposes of clarity, the base of this loop actually forms an anti-parallel β-sheet with the index finger. The β-strands contributed from the thumb of both structures contains a tryptophan (W428 in RHDV; W417 in NWV) that is homologous to the poliovirus $3D^{pol}$ Trp403 described above. Although no data exists concerning the requirement of these caliciviral polymerase interdomain interactions, the enclosed structures of viral RdRP is certainly emerging as a

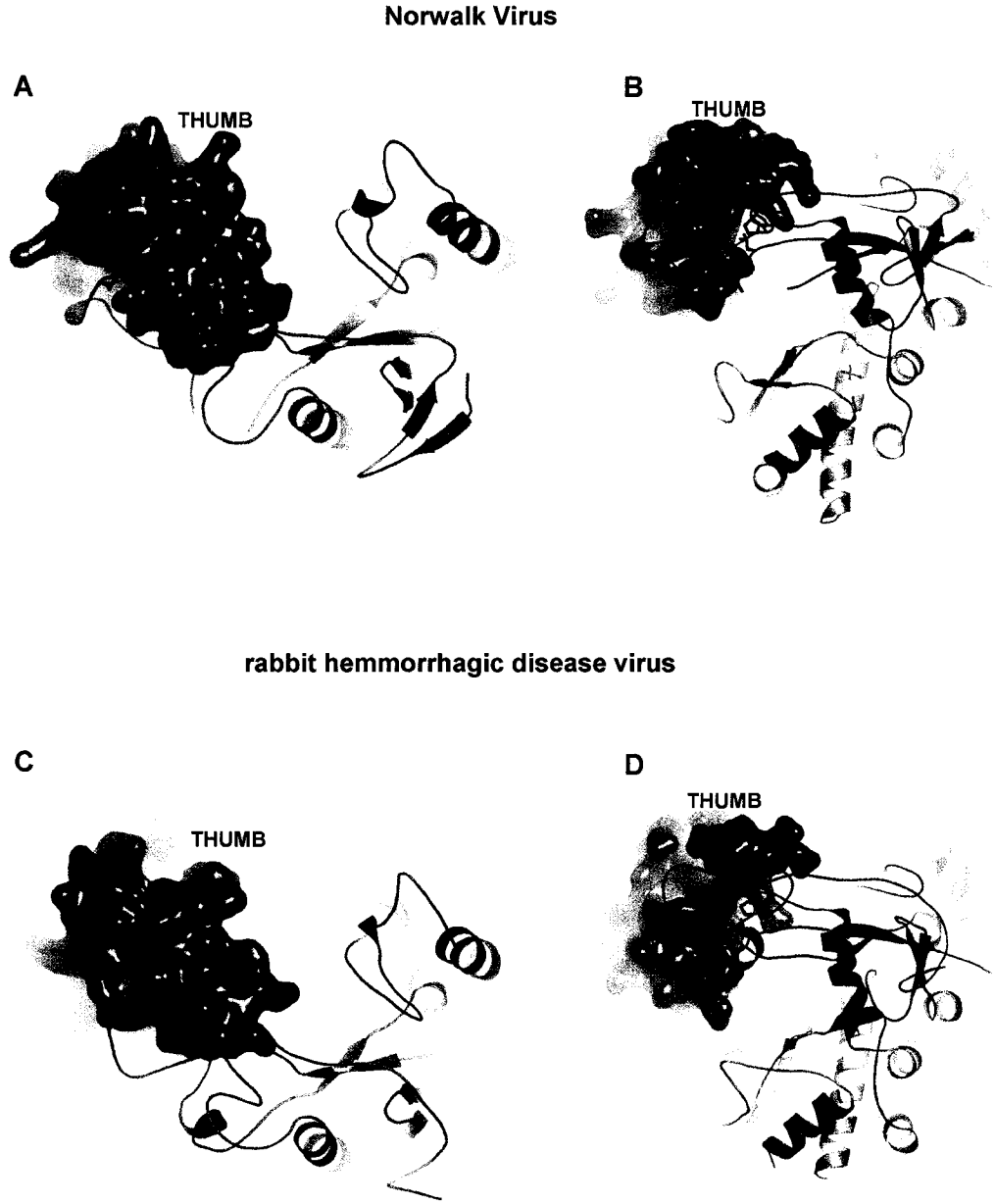


Figure 7.8 Polymerase Structures from the *Caliciviridae* Family showing Extensive Interactions between the Thumb and Index Finger.
(**A**) Top and (**B**) back view of the Norwalk virus polymerase (Ng *et al.*, 2004); and (**C**) top and (**D**) back view of the rabbit hemorrhagic disease virus polymerase (Ng *et al.*, 2002). These structures are depicted with surface representations of the thumb and cartoon representations of the palm and fingers domains. The thumbs are colored blue, palm is grey, fingers domain is red with the exception of the index finger which is green. These structures show extensive interactions between the top of the thumb and index finger.

unique and required feature of these polymerases clearly providing a target for the development of viral-specific inhibitory drugs.

CHAPTER VIII

CONCLUSIONS

The poliovirus $3D^{pol}$ structure provides the first structural insight into the molecular mechanism responsible for the proteolysis-dependent activation of a polymerase. It shows that the newly created $3D^{pol}$ N-terminus is buried in a shallow pocket at the junction of the fingers and palm domains (Fig 8.1). Mutations of the N-terminus or its binding pocket results in deleterious consequences for the proper catalytic activity of this polymerase. These structures show that the buried N-terminus positions the essential Asp238 residue in the active site for interactions with the 2' hydroxyl group of the incoming rNTP. Interestingly, these structures also suggest that mutation of this region may alter the positioning of the ring finger which functions in binding incoming nucleotides for catalysis. The ribonucleotide interactions with Asp238 and the ring finger are confirmed with the co-crystal structures in the presence of Mg^{2+} or Mn^{2+}.

The picornaviral polymerase structures from rhinoviruses and foot-and-mouth disease virus all show similar buried N-termini. Additionally, the recently solved coxsackie virus polymerase structure has a buried N-terminus that is sensitive to mutations like poliovirus $3D^{pol}$ (Peersen Lab., *unpublished results*).

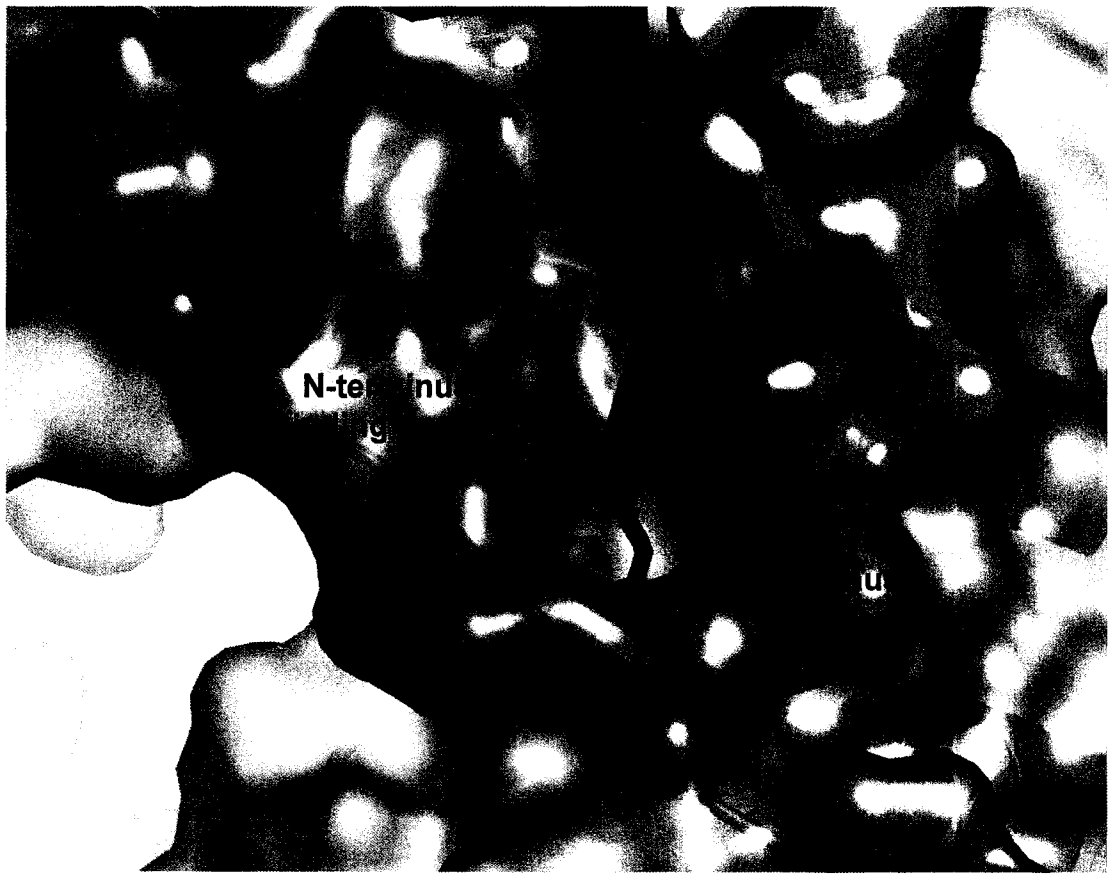


Figure 8.1 Surface representation of $3D^{pol}$ N-terminus binding Pocket. Surface representation of $3D^{pol}$ showing the shallow N-terminus binding pocket formed at the junction of the fingers (index in green, middle in orange) and palm (grey) domains.

Specific proteolytic processing of the RdRP N-terminus is likely a common strategy employed by this family of viruses to regulate proper activation of polymerases at the desired time and location. This enables the $3CD^{pro}$ protein to play an integral role in viral replication by binding to both ends of the viral RNA genome. This has the effect of localizing an inactive polymerase domain to the sites of negative- and positive-strand RNA synthesis, where it is poised to initiate RNA replication after cleavage to the mature $3D^{pol}$ protein.

The circular dichroism thermal denaturation studies highlight the critical importance of an enclosed poliovirus $3D^{pol}$ structure consisting of native fingers and thumb domain interactions. Mutations that disrupt the interdomain interactions between the thumb and index finger completely abolishes *in vitro* activity and results in perturbed thermal melting profiles that we attribute to a destabilization of the fingers domain. Unlike other polymerases, all viral RNA-dependent RNA polymerases to date show enclosed structures, which is emerging as a unique and required feature of these enzymes clearly providing a target for the development of viral specific inhibitory drugs.

The ANS binding and CD thermal denaturation studies surprisingly show that the T_m of $3D_{pol}$ is just above physiological temperatures. This is consistent with biochemical data from the closely related rhinovirus polymerase showing elongation activity is almost completely abolished at 37°C (Hung et al., 2002). Perhaps the reduction of global motion by virtue of the enclosed active site is compensated by the increase in potential thermal motion of the fingers domain such as the pinky (proposed in chapter VI). This would be consistent with the

high B-factors of the fingers and the correlation of increased flexibility of ligand binding sites to accommodate binding (Yang and Bahar, 2005). It is also possible that the thermolability of the enzyme provides a means of inactivating the polymerase. This may by necessary due to differential stoichometric requirements for various viral proteins. For example, far greater numbers of capsid proteins (60 per virus) are necessary for successful virus production compared to many of the non-structural proteins. Proteolytic cleavage of $3CD^{pro}$ bound at the 3' NCR releases active $3D^{pol}$, which is then able to replicate the genome before falling off the end of the template.

APPENDIX 1

Materials and Methods

Mutagenesis and $3D^{pol}$ 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 (r/t) 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 r/t 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% NaN3, 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 2 liters of the same solution overnight at 4°C. The protein was centrifuged, filtered through a 0.22 μM pore size membrane, and 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), 0.1 mM EDTA, 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). An extinction coefficient of 71950 $M^{-1}cm^{-1}$ was used for all full-length proteins which did not have mutations of Trp, Tyr, Phe, or Cys residues. The following values were used for other mutants.

Protein	Extinction Coefficient $M^{-1}cm^{-1}$
$3D^{pol}$-L446D-R455D	71950
$3D^{pol}$ Δ22	64980
$3D^{pol}$ Δ68	62420
$3D^{pol}$ Δ68-380	47200
$3D^{pol}$-F30X-F34X	71950
$3D^{pol}$-W403X	70375
$3CD^{pro}$	79690

Table A1-1 Extinction Coefficients for various $3D^{pol}$ Mutants Calculated from the Expasy web site.

3D^{pol} Crystallization

Crystals were grown by hanging drop vapor diffusion at 16 °C using 8-16 mg ml^{-1} protein. 3D^{pol} Δ68-L446A-R455D crystals grew in 8 days with a precipitant/well solution containing 1.5 M ammonium formate, 0.1 M sodium chloride, 50 mM HEPES (ph 7.0), 2 mM DTT and 0.02 % (w/v) sodium azide. 3D^{pol} L446D-R455D crystals (and all other mutants containing this mutation) 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 20-40 μL of precipitant/well solution and then transported into the 4°C cold room. Ribonucleotide-divalent metal ion crystal "soaks" were accomplished by transfer of the crystals into a precipitant/well solution containing 10 mM NTP and 10mM Mg^{2+} or Mn^{2+} for 1 - 10 hours. All crystals were then flash frozen with liquid nitrogen after slow (~25-30 min) equilibration with precipitant solutions containing 30 % (v/v) glycerol. It is important to have 10 mM NTP and 10 mM divalent metal ion in the cryo-preservant for successful co-crystal complexes.

As discussed in Chapter 2, crystallization of the 3D^{pol}-L446D-R455D mutant required cacodylic acid and DTT, consistent with a proposed mechanism (Fig A1.1A) (Tsao and Maki, 1991). In this reaction, pentavalent cacodylate is reduced by DTT yielding a highly reactive trivalent form that covalently binds with cysteine residues (Figure A1.1B). This modification appears to be required

for crystallization because the hydrophobic dimethyl arsenic adduct on Cys281 forms a major crystal lattice contact with a neighboring molecule (Figure A1.1C).

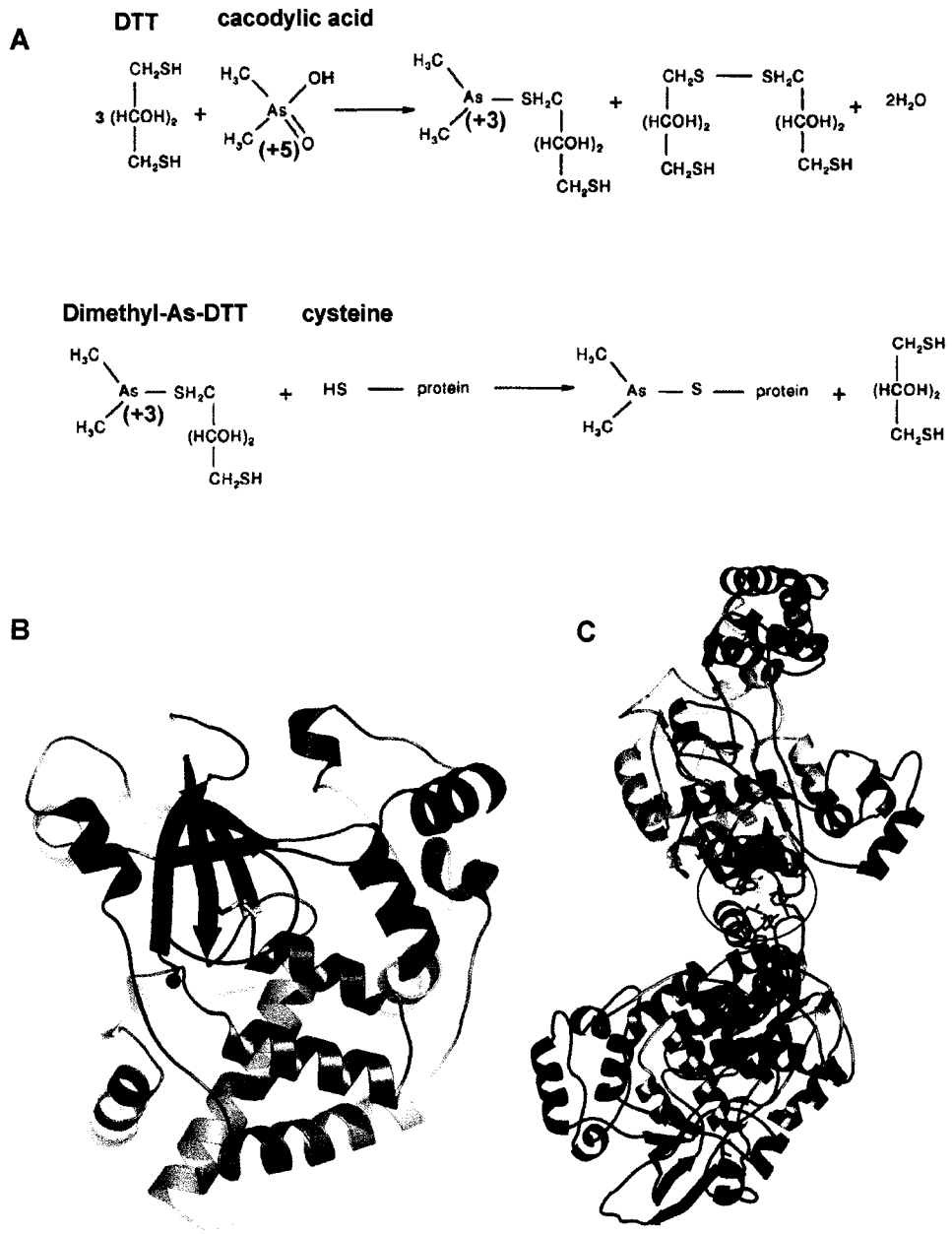


Figure A1.1 Covalent binding of Dimethyl Arsenic to Cysteines
(**A**) Two-step reaction pathway between cacodylate and cysteine. Adapted from Maignan et al., 1998. (**B**) Cysteine 281 of the fingers domain (red) bound with a dimethyl arsenic adduct. (**C**) Crystal lattice contact between the fingers domain centered around Cys281 of one molecule and the thumb domain of another.

Polymerase Activity Assays

Two different methods were developed for evaluating polymerase poly(A)/oligo(T) extension activity. Method 1 was used to evaluate the $3D^{pol}$-mutants in Thompson and Peersen et al., 2004 including $3D^{pol}$ WT, $3D^{pol}$-L446D-R455D, $3CD^{pro}$, $3D^{pro}$-Δ68, $3D^{pol}$-ΔG1, $3D^{pol}$ +Q, $3D^{pol}$ +SQ, $3D^{pol}$-G1A, $3D^{pol}$-G1S, $3D^{pol}$-P119A, $3D^{pol}$-P119G and $3D^{pol}$-D238A. The remaining mutants in this dissertation were evaluated using method 2.

Polymerase activity assay method 1

20 μL Poly(A)/oligo(dT) polymerase extensions reactions contained 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 $3D^{pol}$. Reactions were pre-incubated on ice for at least 5 minutes and then incubated for 30 minutes at 30 °C. 15μL of each reaction were then transferred to DE81 Whatman filters and washed three times with 5% (w/v) Na_2HPO_4, twice with ddH_2O, once with 95% ethanol, and once with 100% ethyl ether. The activity was then assessed by scintillation counting.

Polymerase activity assay method 2

20 μL Poly(A)/oligo(dT) polymerase extensions reactions 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 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 radiolabeled RNA bound to the membrane was quantitated using a phosphoimager. The relative activities were determined by graphing the amount of bound radiolabeled oligonucleotide retained by the Hybond N+ membrane vs. the various time points. Comparison of these slopes yielded the relative activities.

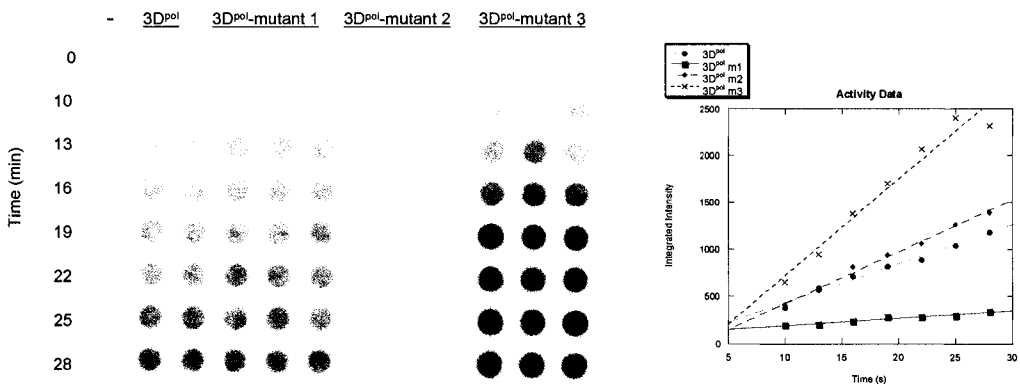


Figure A1.2 Example of Filter-binding Activity Assay
(**A**) Example of a Hybond N(+) filter bound with labeled nucleic acid after polymerase extension reactions are "pulled" through a 96-well dotblot apparatus. (**B**) Sample plot of integrated intensity from panel (A) versus time. The slopes relative to $3D^{pol}$ represent relative polymerase activity.

Fluorescence Polarization Assays

Measurements for fluorescence polarization were conducted in 384-well, black, flat-bottom, polystyrene plates (Corning). The 50 μL reactions 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 A1.3), 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 GF 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. Reactions were setup 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 $I_{\parallel}$ and $I_{\perp}$ is the fluorescence intensity 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. Data were fit to a Hill coefficient model with baseline offset and scaling factor adjustments as follows:

Figure A1.3 Double Fluorescein Labeled Duplex-RNA Construct used for Fluorescence Polarization Assays. Labeled RNA oligos were obtained from IDT and annealed by heating to 75°C in a 100 mL water bath, and allowed to cool to 4°C over 45 minutes. RNA was stored at -80°C

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

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 .

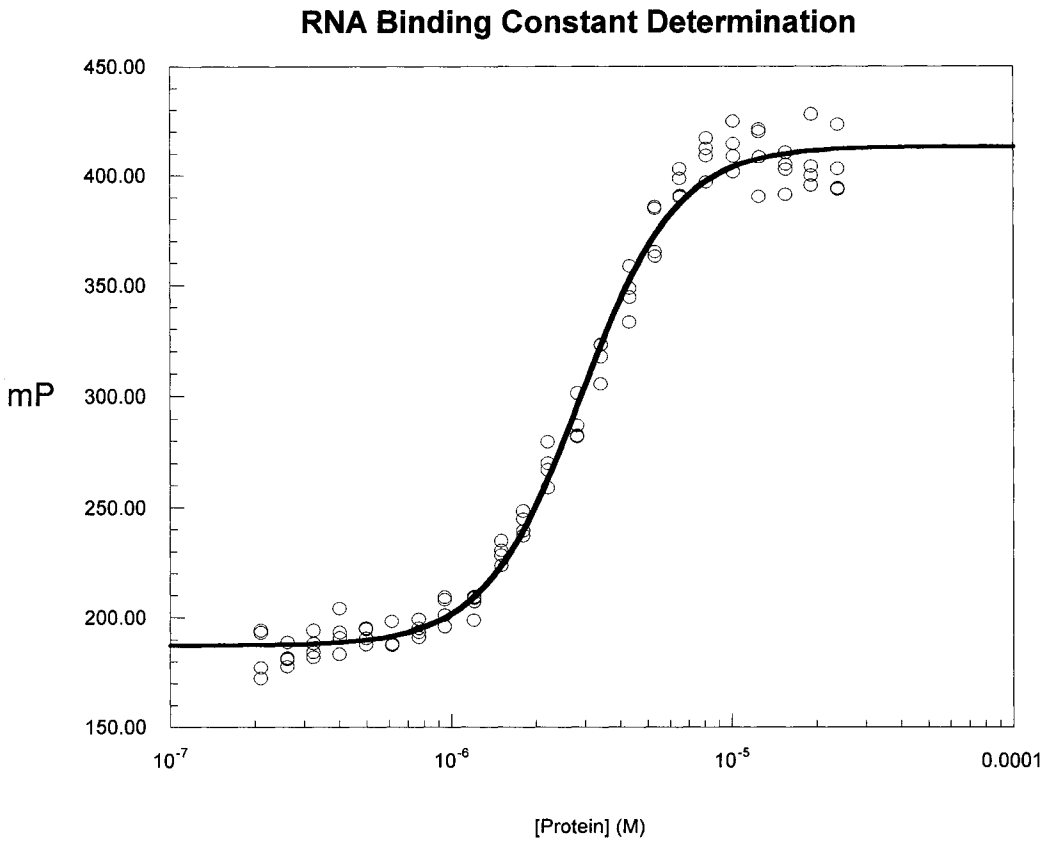


Figure A1.4 Example of Fluorescence Polarization Plot
Plot of milli-polarization versus log of protein concentration for determination of RNA binding constants to $3D^{pol}$. The points were fit to the equation above.

CD Spectroscopy

All CD spectra were obtained on an Aviv model 202 circular dichroism spectrometer. Thermal denaturation experiments at 222 nm were run from 0^{o} to 90^{o}C in one-degree steps, at a two-degrees/minute rate of increase with one-minute equilibration and data averaging at each temperature. Wavelength data are the average of three scans from 260 to 200 nm in 1-nm steps. A 1-cm pathlength cell containing 1.8 mL of ~2μM protein was used in all experiments. Protein concentrations were determined without further dilution of the sample using the extension coefficients listed above. T_m values were obtained from the maxima of d[θ] vs. dT plots.

References

Ackermann, M. and Padmanabhan, R. (2001) De novo synthesis of RNA by the dengue virus RNA-dependent RNA polymerase exhibits temperature dependence at the initiation but not elongation phase. J Biol Chem, 276, 39926-39937.

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 Fold Des, 7, 1417-1426.

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.

Andreotti, A.H. (2003) Native state proline isomerization: an intrinsic molecular switch. Biochemistry, 42, 9515-9524.

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. (1999) Poliovirus RNA-dependent RNA polymerase (3Dpol) is sufficient for template switching in vitro. J Biol Chem, 274, 2706-2716.

Arnold, J.J. and Cameron, C.E. (2000) Poliovirus RNA-dependent RNA polymerase (3D(pol)). Assembly of stable, elongation-competent complexes by using a symmetrical primer-template substrate (sym/sub). J Biol Chem, 275, 5329-5336.

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.

Astatke, M., Grindley, N.D. and Joyce, C.M. (1998a) How E. coli DNA polymerase I (Klenow fragment) distinguishes between deoxy- and dideoxynucleotides. J Mol Biol, 278, 147-165.

Astatke, M., Ng, K., Grindley, N.D. and Joyce, C.M. (1998b) A single side chain prevents Escherichia coli DNA polymerase I (Klenow fragment) from incorporating ribonucleotides. Proc Natl Acad Sci U S A, 95, 3402-3407.

Baltimore D, F.R. (1963) A new ribonucleic acid polymerase appearing after mengovirus infection of L-cells. J Biol Chem, 238, 3395-3400.

Barton, D.J., Morasco, B.J. and Flanegan, J.B. (1996) Assays for poliovirus polymerase, 3D(Pol), and authentic RNA replication in HeLa S10 extracts. Methods Enzymol, 275, 35-57.

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.

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

Benzaghou, I., Bougie, I. and Bisaillon, M. (2004) Effect of metal ion binding on the structural stability of the hepatitis C virus RNA polymerase. J Biol Chem, 279, 49755-49761.

Bienz, K., Egger, D., Rasser, Y. and Bossart, W. (1983) Intracellular distribution of poliovirus proteins and the induction of virus-specific cytoplasmic structures. Virology, 131, 39-48.

Biswal, B.K., Cherney, M.M., Wang, M., Chan, L., Yannopoulos, C.G., Bilimoria, D., Nicolas, O., Bedard, J. and James, M.N. (2005) Crystal structures of the RNA-dependent RNA polymerase genotype 2a of hepatitis C virus reveal two conformations and suggest mechanisms of inhibition by non-nucleoside inhibitors. J Biol Chem, 280, 18202-18210.

Bonneau, A.M. and Sonenberg, N. (1987) Proteolysis of the p220 component of the cap-binding protein complex is not sufficient for complete inhibition of host cell protein synthesis after poliovirus infection. J Virol, 61, 986-991.

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

Bosco, D.A., Eisenmesser, E.Z., Pochapsky, S., Sundquist, W.I. and Kern, D. (2002) Catalysis of cis/trans isomerization in native HIV-1 capsid by human cyclophilin A. Proc Natl Acad Sci U S A, 99, 5247-5252.

Bougie, I., Charpentier, S. and Bisaillon, M. (2003) Characterization of the metal ion binding properties of the hepatitis C virus RNA polymerase. J Biol Chem, 278, 3868-3875.

Braaten, D., Franke, E.K. and Luban, J. (1996a) Cyclophilin A is required for an early step in the life cycle of human immunodeficiency virus type 1 before the initiation of reverse transcription. J Virol, 70, 3551-3560.

Brandts, J.F., Halvorson, H.R. and Brennan, M. (1975) Consideration of the Possibility that the slow step in protein denaturation reactions is due to cis-trans isomerism of proline residues. Biochemistry, 14, 4953-4963.

Brautigam, C.A. and Steitz, T.A. (1998) Structural principles for the inhibition of the 3'-5' exonuclease activity of Escherichia coli DNA polymerase I by phosphorothioates. J Mol Biol, 277, 363-377.

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.

Bressanelli, S., Tomei, L., Rey, F.A. and De Francesco, R. (2002) Structural analysis of the hepatitis C virus RNA polymerase in complex with ribonucleotides. J Virol, 76, 3482-3492.

Bruenn, J.A. (2003) A structural and primary sequence comparison of the viral RNA-dependent RNA polymerases. Nucleic Acids Res, 31, 1821-1829.

Brunger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., Read, R.J., Rice, L.M., Simonson, T. and Warren, G.L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. Acta Crystallogr D Biol Crystallogr, 54, 905-921.

Bryan, T.M., Goodrich, K.J. and Cech, T.R. (2000) Telomerase RNA bound by protein motifs specific to telomerase reverse transcriptase. Mol Cell, 6, 493-499.

Burns, C.C., Lawson, M.A., Semler, B.L. and Ehrenfeld, E. (1989) Effects of mutations in poliovirus 3Dpol on RNA polymerase activity and on polyprotein cleavage. J Virol, 63, 4866-4874.

Butcher, S.J., Makeyev, E.V., Grimes, J.M., Stuart, D.I. and Bamford, D.H. (2000) Crystallization and preliminary X-ray crystallographic studies on the bacteriophage phi6 RNA-dependent RNA polymerase. Acta Crystallogr D Biol Crystallogr, 56 (Pt 11), 1473-1475.

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.

Cai, Z., Yi, M., Zhang, C. and Luo, G. (2005) Mutagenesis analysis of the rGTP-specific binding site of hepatitis C virus RNA-dependent RNA polymerase. J Virol, 79, 11607-11617.

Castro, C., Arnold, J.J. and Cameron, C.E. (2005) Incorporation fidelity of the viral RNA-dependent RNA polymerase: a kinetic, thermodynamic and structural perspective. Virus Res, 107, 141-149.

Celej, M.S., Montich, G.G. and Fidelio, G.D. (2003) Protein stability induced by ligand binding correlates with changes in protein flexibility. Protein Sci, 12, 1496-1506.

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

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

Cho, M.W., Richards, O.C., Dmitrieva, T.M., Agol, V. and Ehrenfeld, E. (1993) RNA duplex unwinding activity of poliovirus RNA-dependent RNA polymerase 3Dpol. J Virol, 67, 3010-3018.

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.

Clark, M.E., Hammerle, T., Wimmer, E. and Dasgupta, A. (1991) Poliovirus proteinase 3C converts an active form of transcription factor IIIC to an inactive form: a mechanism for inhibition of host cell polymerase III transcription by poliovirus. Embo J, 10, 2941-2947.

Cramer, P. (2002) Multisubunit RNA polymerases. Curr Opin Struct Biol, 12, 89-97.

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. and Andino, R. (2002) Implications of high RNA virus mutation rates: lethal mutagenesis and the antiviral drug ribavirin. Microbes Infect, 4, 1301-1307.

Datta, U. and Dasgupta, A. (1994) Expression and subcellular localization of poliovirus VPg-precursor protein 3AB in eukaryotic cells: evidence for glycosylation in vitro. J Virol, 68, 4468-4477.

Di Marco, S., Volpari, C., Tomei, L., Altamura, S., Harper, S., Narjes, F., Koch, U., Rowley, M., De Francesco, R., Migliaccio, G. and Carfi, A. (2005) Interdomain communication in hepatitis C virus polymerase abolished by small molecule inhibitors bound to a novel allosteric site. J Biol Chem, 280, 29765-29770.

Diamond, S.E. and Kirkegaard, K. (1994) Clustered charged-to-alanine mutagenesis of poliovirus RNA-dependent RNA polymerase yields multiple temperature-sensitive mutants defective in RNA synthesis. J Virol, 68, 863-876.

Dorner, A.J., Dorner, L.F., Larsen, G.R., Wimmer, E. and Anderson, C.W. (1982) Identification of the initiation site of poliovirus polyprotein synthesis. J Virol, 42, 1017-1028.

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

Drake, J.W. (1969) Comparative rates of spontaneous mutation. Nature, 221, 1132.

Drake, J.W. (1991) A constant rate of spontaneous mutation in DNA-based microbes. Proc Natl Acad Sci U S A, 88, 7160-7164.

Drake, J.W. and Holland, J.J. (1999) Mutation rates among RNA viruses. Proc Natl Acad Sci U S A, 96, 13910-13913.

Dulbecco, R. (1952) Production of plaques in monolayer tissue cultures by single particles of an animal virus. Proc Natl Acad Sci U S A, 38, 747-752.

Eckert, B., Martin, A., Balbach, J. and Schmid, F.X. (2005) Prolyl isomerization as a molecular timer in phage infection. Nat Struct Mol Biol, 12, 619-623.

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.

Ferrer-Orta, C., Arias, A., Escarmis, C. and Verdaguer, N. (2006a) A comparison of viral RNA-dependent RNA polymerases. Curr Opin Struct Biol, 16, 27-34.

Ferrer-Orta, C., Arias, A., Agudo, R., Perez-Luque, R., Escarmis, C., Domingo, E. and

Verdaguer, N. (2006b) The structure of a protein primer-polymerase complex in the initiation of genome replication. Embo J, 25, 880-888.

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.

Franke, E.K., Yuan, H.E. and Luban, J. (1994) Specific incorporation of cyclophilin A into HIV-1 virions. Nature, 372, 359-362.

Gamarnik, A.V. and Andino, R. (1998) Switch from translation to RNA replication in a positive-stranded RNA virus. Genes Dev, 12, 2293-2304.

Gamble, T.R., Vajdos, F.F., Yoo, S., Worthylake, D.K., Houseweart, M., Sundquist, W.I. and Hill, C.P. (1996) Crystal structure of human cyclophilin A bound to the amino-terminal domain of HIV-1 capsid. Cell, 87, 1285-1294.

Gao, G., Orlova, M., Georgiadis, M.M., Hendrickson, W.A. and Goff, S.P. (1997) Conferring RNA polymerase activity to a DNA polymerase: a single residue in reverse transcriptase controls substrate selection. Proc Natl Acad Sci U S A, 94, 407-411.

Gitti, R.K., Lee, B.M., Walker, J., Summers, M.F., Yoo, S. and Sundquist, W.I. (1996) Structure of the amino-terminal core domain of the HIV-1 capsid protein. Science, 273, 231-235.

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.

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.

Grathwohl, C., Wuthrich, K. (1981) NMR studies of the rates of proline cis/trans isomerisation in oligopeptides. Biopolymers, 20, 2623-2633.

Grochulski, P., Li, Y., Schrag, J.D., Bouthillier, F., Smith, P., Harrison, D., Rubin, B. and Cygler, M. (1993) Insights into interfacial activation from an open structure of Candida rugosa lipase. J Biol Chem, 268, 12843-12847.

Grochulski, P., Li, Y., Schrag, J.D. and Cygler, M. (1994) Two conformational states of Candida rugosa lipase. Protein Sci, 3, 82-91.

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., Reddigari, S.R., Nicklin, M.J., Hammerle, T. and Wimmer, E. (1992) Purification and characterization of poliovirus polypeptide 3CD, a proteinase and a precursor for RNA polymerase. J Virol, 66, 7481-7489.

Harris KS, H.C., Wimmer E. (1990) Proteolytic processin in the replication of picornaviruses. Semin Virol, 1, 323-333.

Harvey, S.a.P., M. (1986) Journal of the American Chemical Society, 108, 6128.

Herold, J. and Andino, R. (2001) Poliovirus RNA replication requires genome circularization through a protein-protein bridge. Mol Cell, 7, 581-591.

Hobson, S. (2000) Crystallographic and Biochemical Studies of Higher Order Poliovirus Polymerase Structures. Dept. of Chemistry and Biochemistry. University of Colorado, Boulder, CO USA, p. 174.

Hobson, S.D., Rosenblum, E.S., Richards, O.C., Richmond, K., Kirkegaard, K. and Schultz, S.C. (2001) Oligomeric structures of poliovirus polymerase are important for function. Embo J, 20, 1153-1163.

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

Hogle JM, C.M., Filman DJ. (1985) Three-dimensional structure of poliovirus at 2.9 A resolution. Science, 229, 1358-1365.

Hong, Z., Cameron, C.E., Walker, M.P., Castro, C., Yao, N., Lau, J.Y. and Zhong, W. (2001) A novel mechanism to ensure terminal initiation by hepatitis C virus NS5B polymerase. Virology, 285, 6-11.

Hope, D.A., Diamond, S.E. and Kirkegaard, K. (1997) Genetic dissection of interaction between poliovirus 3D polymerase and viral protein 3AB. J Virol, 71, 9490-9498.

Howard, B.R., Vajdos, F.F., Li, S., Sundquist, W.I. and Hill, C.P. (2003) Structural insights into the catalytic mechanism of cyclophilin A. Nat Struct Biol, 10, 475-481.

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

Hung, M., Gibbs, C.S. and Tsiang, M. (2002) Biochemical characterization of rhinovirus RNA-dependent RNA polymerase. Antiviral Res, 56, 99-114.

Ide, H., Okagami, M., Murayama, H., Kimura, Y. and Makino, K. (1993) Synthesis and characterization of oligonucleotides containing the alpha-anomer of deoxyadenosine to study its influence on DNA replication. Biochem Mol Biol Int, 31, 485-491.

Jablonski, S.A., Luo, M. and Morrow, C.D. (1991) Enzymatic activity of poliovirus RNA polymerase mutants with single amino acid changes in the conserved YGDD amino acid motif. J Virol, 65, 4565-4572.

Jablonski, S.A. and Morrow, C.D. (1993) Enzymatic activity of poliovirus RNA polymerases with mutations at the tyrosine residue of the conserved YGDD motif: isolation and characterization of polioviruses containing RNA polymerases with FGDD and MGDD sequences. J Virol, 67, 373-381.

Jablonski, S.A. and Morrow, C.D. (1995) Mutation of the aspartic acid residues of the GDD sequence motif of poliovirus RNA-dependent RNA polymerase results in enzymes with altered metal ion requirements for activity. J Virol, 69, 1532-1539.

Jackson, W.T., Giddings, T.H., Jr., Taylor, M.P., Mulinyawe, S., Rabinovitch, M., Kopito, R.R. and Kirkegaard, K. (2005) Subversion of cellular autophagosomal machinery by RNA viruses. PLoS Biol, 3, e156.

Jacobson, S.J., Konings, D.A. and Sarnow, P. (1993) Biochemical and genetic evidence for a pseudoknot structure at the 3' terminus of the poliovirus RNA genome and its role in viral RNA amplification. J Virol, 67, 2961-2971.

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.

Jang, S.K., Davies, M.V., Kaufman, R.J. and Wimmer, E. (1989) Initiation of protein synthesis by internal entry of ribosomes into the 5' nontranslated region of encephalomyocarditis virus RNA in vivo. J Virol, 63, 1651-1660.

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

Jones, T.A., Zou, J.Y., Cowan, S.W. and Kjeldgaard, M. (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. Acta Cryst, A47, 110-119.

Kang, Y.K. and Choi, H.Y. (2004) Cis-trans isomerization and puckering of proline residue. Biophys Chem, 111, 135-142.

Kao, C.C., Del Vecchio, A.M. and Zhong, W. (1999) De novo initiation of RNA synthesis by a recombinant flaviviridae RNA-dependent RNA polymerase. Virology, 253, 1-7.

Kao, C.C., Yang, X., Kline, A., Wang, Q.M., Barket, D. and Heinz, B.A. (2000) Template requirements for RNA synthesis by a recombinant hepatitis C virus RNA-dependent RNA polymerase. J Virol, 74, 11121-11128.

Kao, C.C., Singh, P. and Ecker, D.J. (2001) De novo initiation of viral RNA-dependent RNA synthesis. Virology, 287, 251-260.

Kitamura, N., Semler, B.L., Rothberg, P.G., Larsen, G.R., Adler, C.J., Dorner, A.J., Emini, E.A., Hanecak, R., Lee, J.J., van der Werf, S., Anderson, C.W. and Wimmer, E. (1981) Primary structure, gene organization and polypeptide expression of poliovirus RNA. Nature, 291, 547-553.

Kleywegt, G.J. and Jones, T.A. (1998) Databases in protein crystallography. Acta Crystallogr D Biol Crystallogr, 54, 1119-1131.

Koike, S., Horie, H., Ise, I., Okitsu, A., Yoshida, M., Iizuka, N., Takeuchi, K., Takegami, T. and Nomoto, A. (1990) The poliovirus receptor protein is produced both as membrane-bound and secreted forms. Embo J, 9, 3217-3224.

Koonin, E.V. (1991) The phylogeny of RNA-dependent RNA polymerases of positive-strand RNA viruses. J Gen Virol, 72 (Pt 9), 2197-2206.

Kornberg A, B.T. (1991) DNA Replication, 2nd ed. W. H. Freeman and Co, New York.

Krausslich, H.G., Nicklin, M.J., Toyoda, H., Etchison, D. and Wimmer, E. (1987) Poliovirus proteinase 2A induces cleavage of eucaryotic initiation factor 4F polypeptide p220. J Virol, 61, 2711-2718.

Labonte, P., Axelrod, V., Agarwal, A., Aulabaugh, A., Amin, A. and Mak, P. (2002) Modulation of hepatitis C virus RNA-dependent RNA polymerase activity by structure-based site-directed mutagenesis. J Biol Chem, 277, 38838-38846.

Lai, V.C., Kao, C.C., Ferrari, E., Park, J., Uss, A.S., Wright-Minogue, J., Hong, Z. and Lau, J.Y. (1999) Mutational analysis of bovine viral diarrhea virus RNA-dependent RNA polymerase. J Virol, 73, 10129-10136.

Lai, C.K., Mitchell, J.R. and Collins, K. (2001) RNA binding domain of telomerase reverse transcriptase. Mol Cell Biol, 21, 990-1000.

Lama, J., Paul, A.V., Harris, K.S. and Wimmer, E. (1994) Properties of purified recombinant poliovirus protein 3aB as substrate for viral proteinases and as co-factor for RNA polymerase 3Dpol. J Biol Chem, 269, 66-70.

Landsteiner K, P.E. (1908) Mikroscopische pruparate von einem menschlichen und zwei affentuckermarker. Wein Klin Wschr, 21, 1930.

Laskowski R A, M.M.W., Moss D S & Thornton J M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Cryst., 26, 283-291.

Laurila, M.R., Makeyev, E.V. and Bamford, D.H. (2002) Bacteriophage phi 6 RNA-dependent RNA polymerase: molecular details of initiating nucleic acid synthesis without primer. J Biol Chem, 277, 17117-17124.

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.

Lewis, P.N., Momany, F.A. and Scheraga, H.A. (1971) Folding of polypeptide chains in proteins: a proposed mechanism for folding. Proc Natl Acad Sci U S A, 68, 2293-2297.

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

Li, Y., Kong, Y., Korolev, S. and Waksman, G. (1998b) Crystal structures of the Klenow fragment of Thermus aquaticus DNA polymerase I complexed with deoxyribonucleoside triphosphates. Protein Sci, 7, 1116-1123.

Li, Y., Mitaxov, V. and Waksman, G. (1999) Structure-based design of Taq DNA polymerases with improved properties of dideoxynucleotide incorporation. Proc Natl Acad Sci U S A, 96, 9491-9496.

Liou, Y.C., Sun, A., Ryo, A., Zhou, X.Z., Yu, Z.X., Huang, H.K., Uchida, T., Bronson, R., Bing, G., Li, X., Hunter, T. and Lu, K.P. (2003) Role of the prolyl isomerase Pin1 in protecting against age-dependent neurodegeneration. Nature, 424, 556-561.

Loeffler F, F.P. (1964) Report of the Commission for Research on foot-and-mouth disease. Prentice Hall, Englewood Cliffs NJ, pp. 64-68.

Lohmann, V., Overton, H. and Bartenschlager, R. (1999) Selective stimulation of hepatitis C virus and pestivirus NS5B RNA polymerase activity by GTP. J Biol Chem, 274, 10807-10815.

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.

Lu, K.P., Hanes, S.D. and Hunter, T. (1996) A human peptidyl-prolyl isomerase essential for regulation of mitosis. Nature, 380, 544-547.

Lu, P.J., Wulf, G., Zhou, X.Z., Davies, P. and Lu, K.P. (1999) The prolyl isomerase Pin1 restores the function of Alzheimer-associated phosphorylated tau protein. Nature, 399, 784-788.

Lu, K.P., Liou, Y.C. and Vincent, I. (2003) Proline-directed phosphorylation and isomerization in mitotic regulation and in Alzheimer's Disease. Bioessays, 25, 174-181.

Luban, J., Bossolt, K.L., Franke, E.K., Kalpana, G.V. and Goff, S.P. (1993) Human immunodeficiency virus type 1 Gag protein binds to cyclophilins A and B. Cell, 73, 1067-1078.

Lue, N.F. (2005) A physical and functional constituent of telomerase anchor site. J Biol Chem, 280, 26586-26591.

Lyle, J.M., Bullitt, E., Bienz, K. and Kirkegaard, K. (2002a) Visualization and functional analysis of RNA-dependent RNA polymerase lattices. Science, 296, 2218-2222.

Lyle, J.M., Clewell, A., Richmond, K., Richards, O.C., Hope, D.A., Schultz, S.C. and Kirkegaard, K. (2002b) Similar structural basis for membrane localization and protein priming by an RNA-dependent RNA polymerase. J Biol Chem, 277, 16324-16331.

Lyons, T., Murray, K.E., Roberts, A.W. and Barton, D.J. (2001) Poliovirus 5'-terminal cloverleaf RNA is required in cis for VPg uridylylation and the initiation of negative-strand RNA synthesis. J Virol, 75, 10696-10708.

MacArthur, M.W. and Thornton, J.M. (1991) Influence of proline residues on protein conformation. J Mol Biol, 218, 397-412.

Maignan, S., Guilloteau, J.P., Zhou-Liu, Q., Clement-Mella, C. and Mikol, V. (1998) Crystal structures of the catalytic domain of HIV-1 integrase free and complexed with its metal cofactor: high level of similarity of the active site with other viral integrases. J Mol Biol, 282, 359-368.

Mallis, R.J., Brazin, K.N., Fulton, D.B. and Andreotti, A.H. (2002) Structural characterization of a proline-driven conformational switch within the Itk SH2 domain. Nat Struct Biol, 9, 900-905.

McKnight, K.L. and Lemon, S.M. (1996) Capsid coding sequence is required for efficient replication of human rhinovirus 14 RNA. J Virol, 70, 1941-1952.

McKnight, K.L. and Lemon, S.M. (1998) The rhinovirus type 14 genome contains an internally located RNA structure that is required for viral replication. Rna, 4, 1569-1584.

Mendelsohn, C.L., Wimmer, E. and Racaniello, V.R. (1989) Cellular receptor for poliovirus: molecular cloning, nucleotide sequence, and expression of a new member of the immunoglobulin superfamily. Cell, 56, 855-865.

Molla, A., Jang, S.K., Paul, A.V., Reuer, Q. and Wimmer, E. (1992) Cardioviral internal ribosomal entry site is functional in a genetically engineered dicistronic poliovirus. Nature, 356, 255-257.

Morris, D.P., Phatnani, H.P. and Greenleaf, A.L. (1999) Phospho-carboxyl-terminal domain binding and the role of a prolyl isomerase in pre-mRNA 3'-End formation. J Biol Chem, 274, 31583-31587.

Mosimann, S.C., Cherney, M.M., Sia, S., Plotch, S. and James, M.N. (1997) Refined X-ray crystallographic structure of the poliovirus 3C gene product. J Mol Biol, 273, 1032-1047.

Mueller, S., Wimmer, E. and Cello, J. (2005) Poliovirus and poliomyelitis: a tale of guts, brains, and an accidental event. Virus Res, 111, 175-193.

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.

Neidle, S. (2002) Nucleic Acid Structure and Recognition. Oxford University Press.

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.

Nomaguchi, M., Ackermann, M., Yon, C., You, S. and Padmanabhan, R. (2003) De novo synthesis of negative-strand RNA by Dengue virus RNA-dependent RNA polymerase in vitro: nucleotide, primer, and template parameters. J Virol, 77, 8831-8842.

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.

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.

Olson, W.a.S., JL. (1982) Journal of the American Chemical Society, 108, 2534.

Otwinowski, Z. and Minor, W. (1997) Processing of X-ray Diffraction Data Collected in Oscillation Mode. Methods Enzymol, 276, 307-326.

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.

Paul, A.V., Cao, X., Harris, K.S., Lama, J. and Wimmer, E. (1994) Studies with poliovirus polymerase 3Dpol. Stimulation of poly(U) synthesis in vitro by purified poliovirus protein 3AB. J Biol Chem, 269, 29173-29181.

Paul, A.V., van Boom, J.H., Filippov, D. and Wimmer, E. (1998) Protein-primed RNA synthesis by purified poliovirus RNA polymerase. Nature, 393, 280-284.

Paul, A.V., Mugavero, J., Yin, J., Hobson, S., Schultz, S., van Boom, J.H. and Wimmer, E. (2000a) Studies on the attenuation phenotype of polio vaccines: poliovirus RNA polymerase derived from Sabin type 1 sequence is temperature sensitive in the uridylylation of VPg. Virology, 272, 72-84.

Paul, A.V., Rieder, E., Kim, D.W., van Boom, J.H. and Wimmer, E. (2000b) 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.

Paul, A.V. (2002) Possible unifying mechanisms of picornavirus genome replication. In Semler BL, W.E. (ed.), Molecular Biology of Picornaviruses. ASM Press, pp. 227-246.

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.

Pflugrath, J.W. (1999) The finer things in X-ray diffraction data collection. Acta Crystallogr D Biol Crystallogr, 55 (Pt 10), 1718-1725.

Pilipenko, E.V., Maslova, S.V., Sinyakov, A.N. and Agol, V.I. (1992) Towards identification of cis-acting elements involved in the replication of enterovirus and rhinovirus RNAs: a proposal for the existence of tRNA-like terminal structures. Nucleic Acids Res, 20, 1739-1745.

Plotch, S.J., Palant, O. and Gluzman, Y. (1989) Purification and properties of poliovirus RNA polymerase expressed in Escherichia coli. J Virol, 63, 216-225.

Plotch, S.J. and Palant, O. (1995) Poliovirus protein 3AB forms a complex with and stimulates the activity of the viral RNA polymerase, 3Dpol. J Virol, 69, 7169-7179.

Racaniello VR, B.D. (1981a) Cloned poliovirus complementary DNA is infectious in mammalian cells. Science, 214, 916-919.

Racaniello VR, B.D. (1981b) Molecular cloning of poliovirus cDNA and determination of the complete nucleotide sequence of the viral genome. Proc Natl Acad Sci U S A, 78, 4887-4891.

Raman, C.S., Li, H., Martasek, P., Babu, B.R., Griffith, O.W., Masters, B.S. and Poulos, T.L. (2001) Implications for isoform-selective inhibitor design derived from the binding mode of bulky isothioureas to the heme domain of endothelial nitric-oxide synthase. J Biol Chem, 276, 26486-26491.

Ranjith-Kumar, C.T., Gutshall, L., Sarisky, R.T. and Kao, C.C. (2003) Multiple interactions within the hepatitis C virus RNA polymerase repress primer-dependent RNA synthesis. J Mol Biol, 330, 675-685.

Reimer, U., Scherer, G., Drewello, M., Kruber, S., Schutkowski, M. and Fischer, G. (1998) Side-chain effects on peptidyl-prolyl cis/trans isomerisation. J Mol Biol, 279, 449-460.

Richards, O.C., Yu, P., Neufeld, K.L. and Ehrenfeld, E. (1992) Nucleotide binding by the poliovirus RNA polymerase. J Biol Chem, 267, 17141-17146.

Richards, O.C., Baker, S. and Ehrenfeld, E. (1996) Mutation of lysine residues in the nucleotide binding segments of the poliovirus RNA-dependent RNA polymerase. J Virol, 70, 8564-8570.

Richards, O.C. and Ehrenfeld, E. (1997) One of two NTP binding sites in poliovirus RNA polymerase required for RNA replication. J Biol Chem, 272, 23261-23264.

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.

Rivera, V.M., Welsh, J.D. and Maizel, J.V., Jr. (1988) Comparative sequence analysis of the 5' noncoding region of the enteroviruses and rhinoviruses. Virology, 165, 42-50.

Rossmann MG, A.E., Erikson JW. (1985) Structure of a Human Common Cold Virus and Functional Relationship to other Picornaviruses. Nature, 317, 145-153.

Rothstein, M.A., Richards, O.C., Amin, C. and Ehrenfeld, E. (1988) Enzymatic activity of poliovirus RNA polymerase synthesized in Escherichia coli from viral cDNA. Virology, 164, 301-308.

Saphire, A.C., Bobardt, M.D. and Gallay, P.A. (2002) Cyclophilin a plays distinct roles in human immunodeficiency virus type 1 entry and postentry events, as revealed by spinoculation. J Virol, 76, 4671-4677.

Sarnow, P. (1989) Role of 3'-end sequences in infectivity of poliovirus transcripts made in vitro. J Virol, 63, 467-470.

Schlegel, A., Giddings, T.H., Jr., Ladinsky, M.S. and Kirkegaard, K. (1996) Cellular origin and ultrastructure of membranes induced during poliovirus infection. J Virol, 70, 6576-6588.

Schmid, F.X., Mayr, L.M., Mucke, M. and Schonbrunner, E.R. (1993) Prolyl isomerases: role in protein folding. Adv Protein Chem, 44, 25-66.

Schmid, F.X., Mayr, L.M., Mucke, M. and Schonbrunner, E.R. (1993) Prolyl isomerases: role in protein folding. Adv Protein Chem, 44, 25-66.

Semler, B.L., Anderson, C.W., Hanecak, R., Dorner, L.F. and Wimmer, E. (1982) A membrane-associated precursor to poliovirus VPg identified by immunoprecipitation with antibodies directed against a synthetic heptapeptide. Cell, 28, 405-412.

Semler BL, W.E. (2002) Molecular Biology of Picornaviruses. ASM Press, Washington DC.

Shen, M., Stukenberg, P.T., Kirschner, M.W. and Lu, K.P. (1998) The essential mitotic peptidyl-prolyl isomerase Pin1 binds and regulates mitosis-specific phosphoproteins. Genes Dev, 12, 706-720.

Sheng, X.R., Zhang, H.J., Pan, X.M., Li, X.F. and Zhou, J.M. (1997) Domain movement in rabbit muscle adenylate kinase might involve proline isomerization. FEBS Lett, 413, 429-432.

Shindyalov, I.N. and Bourne, P.E. (1998) Protein structure alignment by incremental combinatorial extension (CE) of the optimal path. Protein Eng, 11, 739-747.

Shirako, Y., Strauss, E.G. and Strauss, J.H. (2000) Suppressor mutations that allow sindbis virus RNA polymerase to function with nonaromatic amino acids at the N-terminus: evidence for interaction between nsP1 and nsP4 in minus-strand RNA synthesis. Virology, 276, 148-160.

Sholders, A. (2006) Structural Characterization of poliovirus RNA-dependent RNA Polymerase Fidelity and Translocation. Dept. of Biochemistry and Molecular Biology. Colorado State University, Fort Collins, CO.

Sholders, A., Thompson AA, Campagnola G, Peersen OB. (2006) Structure of the drug resistant G64S poliovirus polymerase in complex with ribavirin triphosphate. manuscript in preparation.

Silvestri, L.S., Parilla, J.M., Morasco, B.J., Ogram, S.A. and Flanegan, J.B. (2006) Relationship between poliovirus negative-strand RNA synthesis and the length of the 3' poly(A) tail. Virology, 345, 509-519.

Sonenberg, N. (1987) Regulation of translation by poliovirus. Adv Virus Res, 33, 175-204.

Sosunov, V., Sosunova, E., Mustaev, A., Bass, I., Nikiforov, V. and Goldfarb, A. (2003) Unified two-metal mechanism of RNA synthesis and degradation by RNA polymerase. Embo J, 22, 2234-2244.

Sosunov, V., Zorov, S., Sosunova, E., Nikolaev, A., Zakeyeva, I., Bass, I., Goldfarb, A., Nikiforov, V., Severinov, K. and Mustaev, A. (2005) The involvement of the aspartate triad of the active center in all catalytic activities of multisubunit RNA polymerase. Nucleic Acids Res, 33, 4202-4211.

Sousa, R. and Padilla, R. (1995) A mutant T7 RNA polymerase as a DNA polymerase. Embo J, 14, 4609-4621.

Spector, D.H. and Baltimore, D. (1974) Requirement of 3'-terminal poly(adenylic acid) for the infectivity of poliovirus RNA. Proc Natl Acad Sci U S A, 71, 2983-2987.

Steitz, T.A. (1998) A mechanism for all polymerases. Nature, 391, 231-232.

Stewart, D.E., Sarkar, A. and Wampler, J.E. (1990) Occurrence and role of cis peptide bonds in protein structures. J Mol Biol, 214, 253-260.

Summers, D.F. and Maizel, J.V., Jr. (1968) Evidence for large precursor proteins in poliovirus synthesis. Proc Natl Acad Sci U S A, 59, 966-971.

Tabor, S. and Richardson, C.C. (1989) Effect of manganese ions on the incorporation of dideoxynucleotides by bacteriophage T7 DNA polymerase and Escherichia coli DNA polymerase I. Proc Natl Acad Sci U S A, 86, 4076-4080.

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.

Temiakov, D., Mentesana, P.E., Ma, K., Mustaev, A., Borukhov, S. and McAllister, W.T. (2000) The specificity loop of T7 RNA polymerase interacts first with the promoter and then with the elongating transcript, suggesting a mechanism for promoter clearance. Proc Natl Acad Sci U S A, 97, 14109-14114.

Tete-Favier, F., Cobessi, D., Boschi-Muller, S., Azza, S., Branlant, G. and Aubry, A. (2000) Crystal structure of the Escherichia coli peptide methionine sulphoxide reductase at 1.9 A resolution. Structure Fold Des, 8, 1167-1178.

Teterina, N.L., Gorbalenya, A.E., Egger, D., Bienz, K. and Ehrenfeld, E. (1997) Poliovirus 2C protein determinants of membrane binding and rearrangements in mammalian cells. J Virol, 71, 8962-8972.

Thali, M., Bukovsky, A., Kondo, E., Rosenwirth, B., Walsh, C.T., Sodroski, J. and Gottlinger, H.G. (1994) Functional association of cyclophilin A with HIV-1 virions. Nature, 372, 363-365.

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.

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.

Tomei, L., Altamura, S., Bartholomew, L., Biroccio, A., Ceccacci, A., Pacini, L., Narjes, F., Gennari, N., Bisbocci, M., Incitti, I., Orsatti, L., Harper, S., Stansfield, I., Rowley, M., De Francesco, R. and Migliaccio, G. (2003) Mechanism of action and antiviral activity of benzimidazole-based allosteric inhibitors of the hepatitis C virus RNA-dependent RNA polymerase. J Virol, 77, 13225-13231.

Towner, J.S., Ho, T.V. and Semler, B.L. (1996) Determinants of membrane association for poliovirus protein 3AB. J Biol Chem, 271, 26810-26818.

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

Tsao, D.H. and Maki, A.H. (1991) Optically detected magnetic resonance study of the interaction of an arsenic(III) derivative of cacodylic acid with EcoRI methyl transferase. Biochemistry, 30, 4565-4572.

Tuthill, T.J., Bubeck, D., Rowlands, D.J. and Hogle, J.M. (2006) Characterization of early steps in the poliovirus infection process: receptor-decorated liposomes induce conversion of the virus to membrane-anchored entry-intermediate particles. J Virol, 80, 172-180.

Valegard, K., Liljas, L., Fridborg, K. and Unge, T. (1990) The three-dimensional structure of the bacterial virus MS2. Nature, 345, 36-41.

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.

Wedemeyer, W.J., Welker, E. and Scheraga, H.A. (2002) Proline cis-trans isomerization and protein folding. Biochemistry, 41, 14637-14644.

Weinrich, S.L., Pruzan, R., Ma, L., Ouellette, M., Tesmer, V.M., Holt, S.E., Bodnar, A.G., Lichtsteiner, S., Kim, N.W., Trager, J.B., Taylor, R.D., Carlos, R., Andrews, W.H., Wright, W.E., Shay, J.W., Harley, C.B. and Morin, G.B. (1997) Reconstitution of human telomerase with the template RNA component hTR and the catalytic protein subunit hTRT. Nat Genet, 17, 498-502.

Wimmer, E., Hellen, C.U. and Cao, X. (1993) Genetics of poliovirus. Annu Rev Genet, 27, 353-436.

Wulf, G., Finn, G., Suizu, F. and Lu, K.P. (2005) Phosphorylation-specific prolyl isomerization: is there an underlying theme? Nat Cell Biol, 7, 435-441.

Xiang, W., Cuconati, A., Hope, D., Kirkegaard, K. and Wimmer, E. (1998) Complete protein linkage map of poliovirus P3 proteins: interaction of polymerase 3Dpol with VPg and with genetic variants of 3AB. J Virol, 72, 6732-6741.

Yaffe, M.B., Schutkowski, M., Shen, M., Zhou, X.Z., Stukenberg, P.T., Rahfeld, J.U., Xu, J., Kuang, J., Kirschner, M.W., Fischer, G., Cantley, L.C. and Lu, K.P. (1997) Sequence-specific and phosphorylation-dependent proline isomerization: a potential mitotic regulatory mechanism. Science, 278, 1957-1960.

Yalamanchili, P., Harris, K., Wimmer, E. and Dasgupta, A. (1996) Inhibition of basal transcription by poliovirus: a virus- encoded protease (3Cpro) inhibits formation of TBP-TATA box complex in vitro. J Virol, 70, 2922-2929.

Yang, L.W. and Bahar, I. (2005) Coupling between catalytic site and collective dynamics: a requirement for mechanochemical activity of enzymes. Structure, 13, 893-904.

Yang, L.W. and Bahar, I. (2005) Coupling between catalytic site and collective dynamics: a requirement for mechanochemical activity of enzymes. Structure, 13, 893-904.

Yao, J., Dyson, H.J. and Wright, P.E. (1994) Three-dimensional structure of a type VI turn in a linear peptide in water solution. Evidence for stacking of aromatic rings as a major stabilizing factor. J Mol Biol, 243, 754-766.

Yi, G.H., Zhang, C.Y., Cao, S., Wu, H.X. and Wang, Y. (2003) De novo RNA synthesis by a recombinant classical swine fever virus RNA-dependent RNA polymerase. Eur J Biochem, 270, 4952-4961.

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.

Yogo, Y. and Wimmer, E. (1972) Polyadenylic acid at the 3'-terminus of poliovirus RNA. Proc Natl Acad Sci U S A, 69, 1877-1882.

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.

Yuan, X., Werner, J.M., Knott, V., Handford, P.A., Campbell, I.D. and Downing, K. (1998) Effects of proline cis-trans isomerization on TB domain secondary structure. Protein Sci, 7, 2127-2135.

Zheng, W., Brooks, B.R., Doniach, S. and Thirumalai, D. (2005) Network of dynamically important residues in the open/closed transition in polymerases is strongly conserved. Structure, 13, 565-577.