

THESIS

ROLE OF BASIC AND HYDROPHOBIC RESIDUES IN THE POLIOVIRUS  
POLYMERASE ELONGATION COMPLEX AND THE STRUCTURE OF A  
COXSACKIEVIRUS POLYMERASE ELONGATION COMPLEX

Submitted by

Matt Kortus

Biochemistry and Molecular Biology

In partial fulfillment of the requirements

For the degree of Master of Science

Colorado State University

Fort Collins, Colorado

Fall 2011

Master's Committee:

Advisor: Olve Peersen

P. Shing Ho  
Erica Suchman

## ABSTRACT

### BIOCHEMICAL AND STRUCTURAL CHARACTERIZATION OF PICORNAVIRAL RNA-DEPENDENT RNA POLYMERASES

Picornaviruses encode for and require a viral RNA-dependent RNA polymerase (RdRP) for genome replication. This enzyme synthesizes negative-sense RNA from the infecting positive sense genome producing a replicative intermediate. The negative sense RNA then serves as a template for synthesis of additional positive-sense RNA. To efficiently replicate the genome, RdRPs must form a stable and processive elongation complex (EC) by binding RNA, incorporating the first templating nucleotide, and undergoing a necessary conformational. Upon completion of these steps that comprise initiation, the newly formed EC is capable of rapidly replicating the viral genome. The work presented in this thesis 1) investigates the role that several basic and hydrophobic residues serve in forming and maintaining the poliovirus (PV) EC and 2) presents the crystal structure of a coxsackievirus (CV) EC. To determine the role of that several arginines, lysines, and tyrosines play in the PV polymerase, we assessed whether mutations to these residues affect initiation, elongation, or stability of the EC. The data indicates the basic residues within the fingers domain of the PV polymerase have a major role in binding RNA. In addition, data shows two tyrosine residues in particular are critical for formation and maintenance of the EC. Overall, the data provides evidence the fingers domain interacts with the template RNA in a manner not captured by crystal

structures. Finally, we have solved the structure of a CVEC stalled after incorporation of four nucleotides. The CVEC structure closely matches the previously solved PVEC structure. In addition, one crystal form produced an elongation complex trapped in a translocation intermediate state.

## TABLE OF CONTENTS

### **1 Introduction**

1.1 Picornavirus Overview and History.....	1
1.2 Viral Life Cycle.....	3
1.2.1 Infection and Cell Entry.....	3
1.2.2 Viral Genome and Translation.....	7
1.2.3 Proteolytic Processing.....	11
1.2.4 RNA Replication.....	13
1.3 Mechanism and Kinetics of Picornaviral Polymerase.....	17
1.3.1 Polymerase Kinetics.....	17
1.3.2 Polymerase Fidelity.....	20
1.4 Structure of Picornaviral Polymerases.....	21
1.4.1 Apo PV 3D <sup>pol</sup> Structure.....	21
1.4.2 PVEC Structure.....	24

### **2 Role of Basic and Hydrophobic Residues in the Poliovirus Polymerase Elongation Complex**

2.1 Introduction.....	27
2.2 Material and Methods.....	29
2.2.1 Mutant Generation and Protein Expression.....	29
2.2.2 Protein Purification.....	31
2.2.3 RNA Oligonucleotides.....	32
2.2.4 RNA Binding.....	33

2.2.5	Elongation Complex Formation.....	33
2.2.6	Elongation Rates.....	34
2.2.7	Elongation Complex Stability.....	35
2.3	Results.....	36
2.3.1	Charge Residues Play Role in Binding RNA.....	36
2.3.2	Tyrosine Mutations Slow EC Formations Rates.....	40
2.3.3	Elongation Rate not Affected by Mutations.....	42
2.3.4	Mutations Affect EC Stability.....	44
2.4	Discussion.....	44
<b>3</b>	<b>Structure of a Coxsackievirus B3 Elongation Complex</b>	
3.1	Introduction.....	50
3.2	Material and Methods.....	54
3.2.1	Protein Expression and Purification.....	54
3.2.2	RNA Synthesis and Preparation.....	57
3.2.3	EC Assembly, Purification, and Crystallization.....	59
3.2.4	Diffraction Data Collection and Reduction.....	60
3.3	Results.....	61
3.3.1	CVEC Kinetic Assays.....	61
3.3.2	CV 3D <sup>pol</sup> Elongation Complex Structure.....	61
3.3.3	CVEC Translocation Intermediate.....	65
3.3.4	CVEC Crystal Soaks.....	68
3.4	Discussion.....	68
	References.....	73

## LIST OF FIGURES

<b>1</b>	<b>Introduction</b>	
1.1	Postulated pathway for poliovirus infection.....	4
1.2	Poliovirus entry into host cells.....	6
1.3	Structural organization of the poliovirus RNA genome and proteolytic processing.....	8
1.4	Picornaviral Life-Cycle.....	9
1.5	Translation initiation mechanisms.....	10
1.6	Four-step model for asymmetric replication of poliovirus RNA.....	14
1.7	Kinetic mechanisms for initiation and elongation by 3D <sup>pol</sup> .....	18
1.8	Structural overview of poliovirus 3D <sup>pol</sup> .....	22
1.9	Poliovirus 3D <sup>pol</sup> -RNA elongation complex structure.....	25
<b>2</b>	<b>Role of Basic and Hydrophobic Residues in Poliovirus Polymerase</b>	
2.1	Structural overview of basic and hydrophobic residues.....	30
2.2	RNA binding results.....	37
2.3	Tyrosine residues play role in elongation.....	39
2.4	Elongation complex formation results.....	41
2.5	Elongation rate experiments.....	43
2.6	Elongation complex stability results.....	45
<b>3</b>	<b>Structure of a Coxsackievirus Polymerase Elongation Complex</b>	
3.1	Comparison of picornaviral polymerase structures.....	51
3.2	Poliovirus 3D <sup>pol</sup> -RNA elongation complex structure.....	53

3.3 RNA constructs for EC assembly.....	55
3.4 Production of RNA constructs for EC assembly.....	58
3.5 Comparison of PV3D <sup>pol</sup> and CV3D <sup>pol</sup> Kinetics.....	62
3.6 Coxsackievirus Elongation Complex Crystals.....	63
3.7 Comparison of the CV3D <sup>pol</sup> apo structure and elongation complex.....	64
3.8 Comparison of the CVEC and PVEC structures.....	66
3.9 RNA backsliding in the CV Translocation Intermediate.....	67
3.10 Thr114 as a translocation gate.....	69

## LIST OF TABLES

2.1 Summary of Kinetic Data for PV3D <sup>pol</sup> mutants.....	38
3.1 Comparison of PV3D <sup>pol</sup> and CV3D <sup>pol</sup> kinetics.....	62
3.2 Crystallographic data and refinement statistics.....	63

# Chapter 1

## Introduction

### 1.1 Picornavirus Overview and History

Members of the Picornaviridae family of viruses contain a single-stranded positive-sense RNA genome between 6500 and 9500 nucleotides in length packaged within a small (~ 30 nm) non-enveloped icosahedral capsid. The family can be broken into twelve different genera, and contains many animal and human pathogens. Examples of medically relevant viruses within this family include poliovirus, coxsackievirus, rhinovirus, foot and mouth disease virus, and Hepatitis A virus. Due to the health implications associated with picornaviruses, a significant amount of research has gone into understanding the mechanisms by which they cause disease. In fact, much of our knowledge regarding viral processes has arisen from studying picornaviruses. In 1898, foot and mouth disease virus became the first identified animal virus (Loeffler and Frosch, 1964), with the discovery of poliovirus following ten years later (Landsteiner and Popper, 1908). The plaque assay, used throughout virology to quantify viral infectivity (Dulbecco, 1952), and the first infectious DNA clone of an animal virus were developed using poliovirus (reviewed in Racaniello, 1981). In addition, the first three-dimensional structures of virions were determined through X-ray crystallography studies on poliovirus and rhinovirus (Hogel *et al.*, 1985; Rossmann *et al.*, 1985). In general, poliovirus has become the prototypical virus to study in this family, due mainly to the historical impact of the virus.

The first significant outbreak of poliomyelitis in the United States occurred in 1894 (World Health Organization, 2010). After the initial outbreak, polio prevalence grew continually both in the United States as well as the rest of the world. In 1916, poliovirus became an epidemic in New York, leading to many seasonal national epidemics that peaked during the 1950's. These epidemics stimulated research that culminated in the development of two polio vaccines: the inactivated polio vaccine (IPV) by Salk and Youngner (Salk *et al.*, 1954) and the oral polio vaccine (OPV) by Sabin (Sabin *et al.*, 1960). The IPV employs chemically killed virus to provide immunity, while the OPV presents an attenuated virus. The two vaccines proved very effective in combating the epidemic and significantly reduced the number of cases occurring each year. However, a lack of consistent immunization hindered progress towards eliminating the threat so that in 1988 poliovirus still caused over 350,000 worldwide cases of paralysis every year. This prompted the World Health Assembly to launch resolution 41.28: *The Global Polio Eradication Initiative* with the resolve to eradicate polio from the planet by the year 2000.

The initiative led to a dramatic reduction in the number of cases occurring each year so that by 1994 the United States was declared polio-free and, by 2002, half of the world's population lived in polio-free regions (World Health Organization, 2010). Success stemmed from the combination of the two vaccines and efficient diagnostic techniques for detecting poliovirus within populations. However, approximately 1500 cases of poliomyelitis still occur worldwide every year. In addition, wild poliovirus remains endemic in four countries: Afghanistan, India, Nigeria, and Pakistan (World Health Organization, 2010). The eradication effort faces primarily political struggles in these countries. However, another issue stems from the live attenuated OPV's ability to revert back to wild-type infectiousness resulting in vaccine-derived poliovirus (VDPV). As a result, the OPV, which provides immunity in underdeveloped

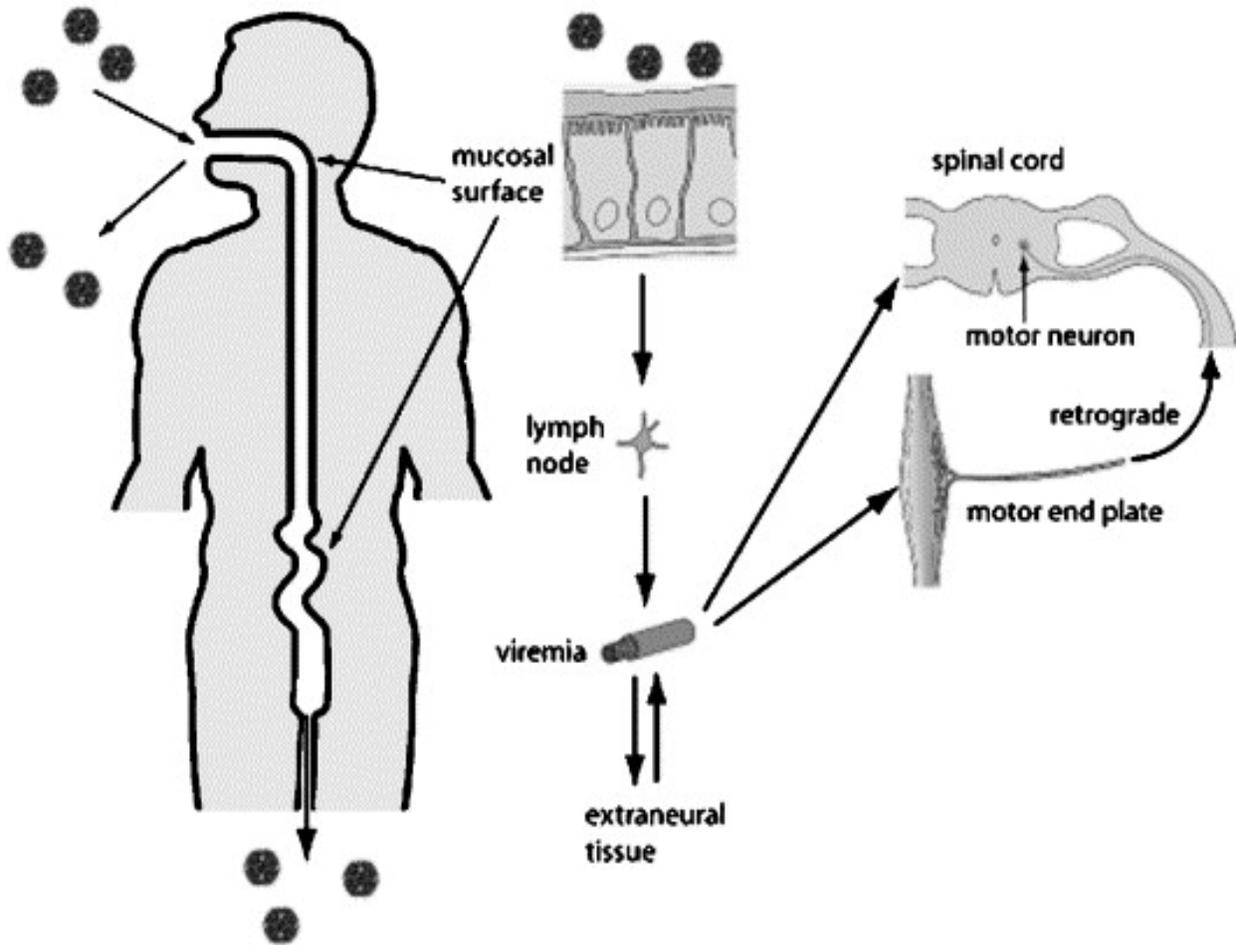
countries, is an additional source for poliovirus infection. Cases of VDPV have been reported in several areas, including China, Hispaniola, the Philippines, and Madagascar (World Health Organization, 2010). While the IPV circumvents this problem, the vaccine is far more expensive than the OPV and requires a trained health professional for administration, making it less suitable for underdeveloped countries. The weaknesses associated with each vaccine necessitate continued efforts to develop new antivirals and better vaccines.

## **1.2 Viral Life Cycle**

### *1.2.1 Infection and Cell Entry*

Poliovirus, a member of the enterovirus genus, spreads primarily through fecal-oral transmission. Once contracted, viral replication occurs in the upper gastrointestinal tract, specifically within the oropharyngeal and intestinal mucosa of the human host (Bodian and Hortsman, 1965; Sabin, 1956). After multiplication, viral particles pass into the lymph nodes from which they eventually enter the bloodstream, causing a transient viremia (Fig. 1.1) (Bodian and Hortsman, 1965). Most cases of poliovirus infection lead to nonspecific symptoms including fatigue, headache, fever, sore throat, and malaise. However, in rare instances (<1%), the virus enters the central nervous system (CNS), infecting and destroying the motor neurons within the brain stem or spinal cord and results in the characteristic paralytic poliomyelitis.

Viral entry into host cells is mediated by interactions with the poliovirus receptor (PVR), CD155. This integral membrane protein is a member of the immunoglobulin (Ig) super family of proteins and contains three Ig-like domains (Mendelsohn *et al.*, 1989). In host cells, CD155 interacts with Nectin-3, a member of a family of adhesion molecules that interacts with cadherin to form and stabilize cell-cell adherens junctions that mediate cell-cell adhesion (Mueller and

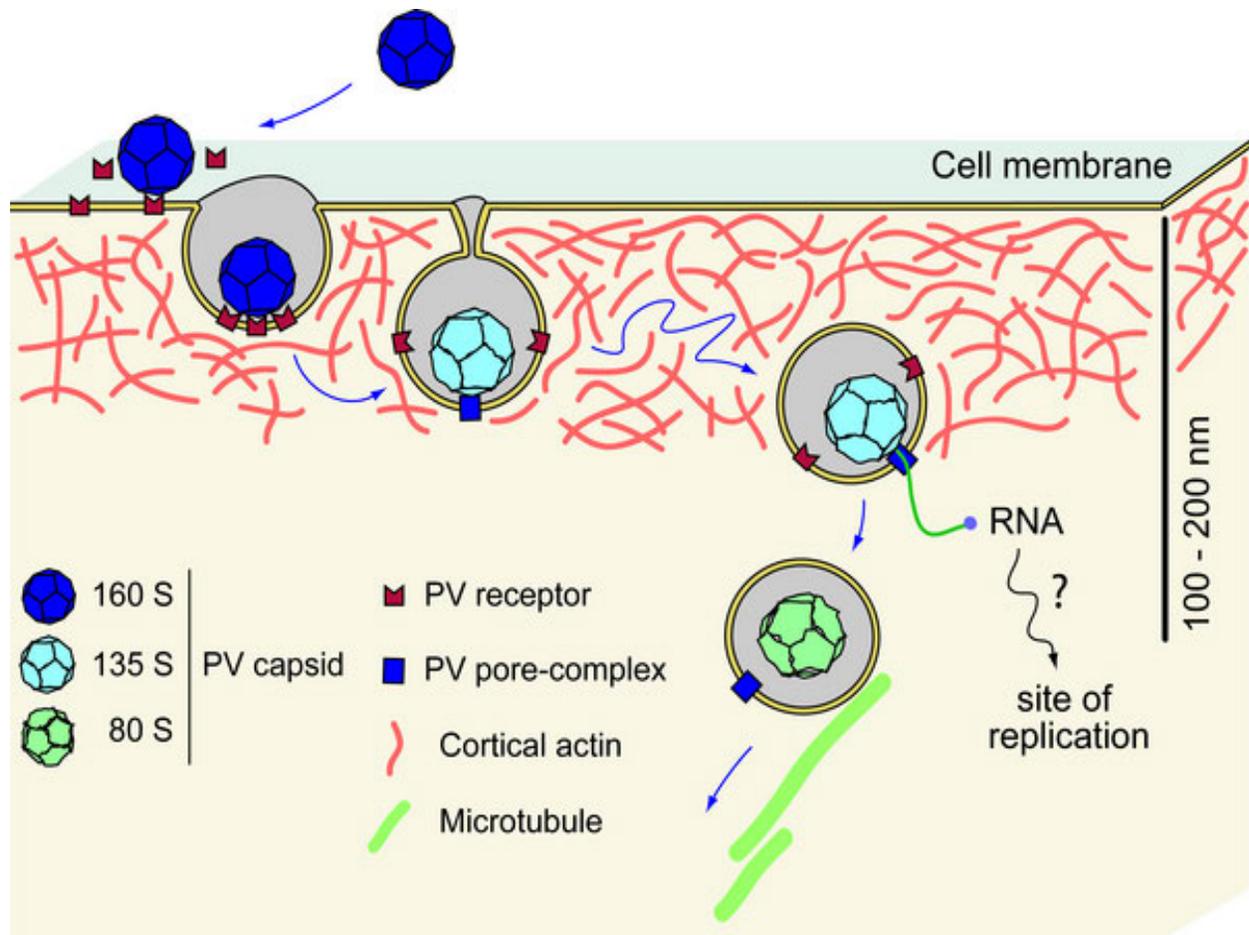


**FIG 1.1 Postulated pathway for poliovirus infection.** The virus begins replicating in the upper gastrointestinal tract and reaches the bloodstream through the lymph nodes, resulting in a primary viremia. In some instances, poliovirus continues on to the motor neurons, leading ultimately to paralysis. In most cases, the virus spreads through fecal-oral contamination, but may also be spread through the respiratory route (from Racaniello, 2005).

Wimmer, 2003). Upon binding to CD155, the poliovirus internalizes through an actin- and receptor tyrosine kinase dependent manner and releases the genome from the vesicle within the first 100-200 nm of the cytosol (Fig. 1.2) (Brandenburg *et al.*, 2007).

Closely related to poliovirus, coxsackievirus B3 (CVB3) is another enterovirus that presents a threat to human health. Scientists first isolated CVB3 as poliovirus, only later to find distinctions between the two. CVB3 spreads through fecal-oral contamination and follows a similar path of infection as poliovirus. In most cases, CVB3 infection produces only mild symptoms such as diarrhea, headache, and fever. However, CVB3 infection within heart cells causes myocarditis and can lead to cardiomyopathies (Beck *et al.*, 1990; Esfandiarei and McManus, 2008). CVB3 infects approximately 20,000 individuals each year in the US, of which between 10-20% will develop chronic myocarditis (reviewed in Kim and Nam, 2010). Despite all of this, no CVB3-specific vaccine exists at this time.

Upon infection, CVB3 binds to host cells through interactions with the coxsackievirus and adenovirus receptor (CAR), allowing for viral entry (Bergelson *et al.*, 1997). Since CAR resides in the tight junctions, rather than the apical surface of epithelial cells, CVB3 cannot immediately bind the receptor. Therefore, CVB3 must bind to the GPI anchored protein CD55, a decay-accelerating factor (DAF) (Shafren *et al.*, 1995). Binding activates AB1kinase, which in turn triggers Rac-dependent actin rearrangements, allowing for the virus to move to the tight junction. In addition, DAF facilitates activation of Fyn kinase, an essential enzyme for the formation of caveolar vesicles that are responsible for endocytosis of the CVB3 (Coyle and Bergelson, 2006). The mechanism of cell entry highlights one key difference between CVB3 and poliovirus; however, once the viral genome is inserted into the cell, both viruses employ nearly



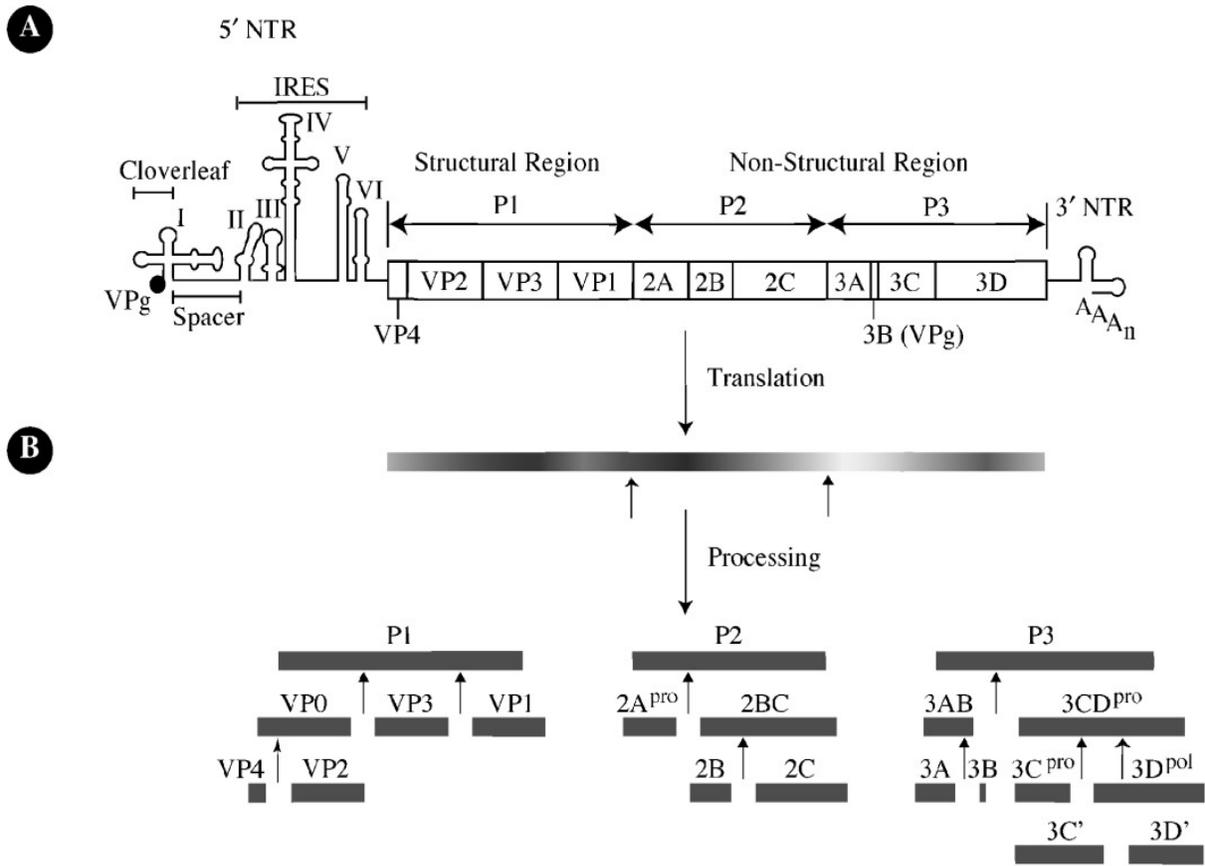
**FIG 1.2 Poliovirus entry into host cells.** Upon binding to CD155, the virus is internalized in actin- and receptor tyrosine kinase-dependent manner. The viral genome is released within the first 100-200nm from the plasma membrane, after which the capsid is transported along microtubules (from Brandenburg *et al.*, 2007).

identical means for replication and production of virions. Therefore, this chapter will focus primarily on poliovirus, the prototypical enterovirus, due to the substantial amount of information and research available on poliovirus replication.

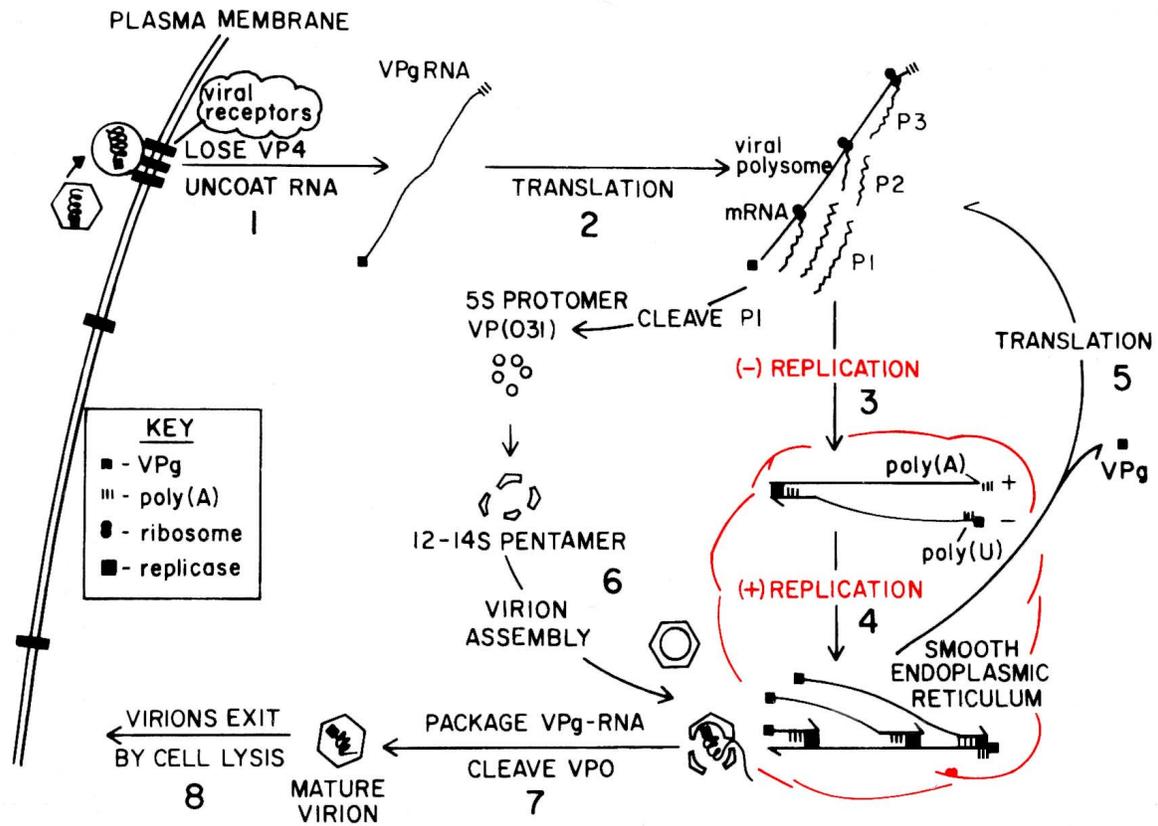
### *1.2.2 Viral Genome and Translation*

Picornaviral genomes consist of three basic components: the 5' untranslated region (5'-UTR), the coding region, and the 3' untranslated region (3'-UTR) (Fig. 1.3A). The 5'-UTR contains substantial secondary structures that play crucial roles in both translation and replication of the viral genome. Of the six stem loop structures present in the 5'-UTR, the first forms a cloverleaf comprised of three small stem-loops that interact with various proteins to direct replication (Rivera *et al.*, 1988; Andino *et al.*, 1990). The other five stem-loops constitute the internal ribosome entry site (IRES) allowing for cap-independent translation (Jang *et al.*, 1988; Pelletier and Sonenberg, 1988; Fitzgerald and Semler, 2009). The 3'-UTR (~70 nucleotides in poliovirus) is relatively small compared to the 5'-UTR (~740 nucleotides in poliovirus), but also contains secondary structures that mediate RNA-RNA interactions to form a tertiary pseudoknot structure that is important for RNA syntheses (Jacobson *et al.*, 1993). In addition, the 3'-UTR of both virion RNA and mRNA contain a poly(A) tail with an average length of 60 nucleotides in poliovirus (Yogo and Wimmer, 1972).

Upon insertion of the positive-sense viral genome, translation is carried out by host cell machinery that recognizes the IRES and, in turn, produces all the necessary viral proteins (Fig. 1.4) The absence of the 7-methyl guanosine cap typically found on cellular mRNAs, which initiates canonical cap-dependent translation through binding of translation initiation factor eIF-4F (Merrick, 1990), necessitates this alternative mechanism for initiation (Fig. 1.5). Picornaviral

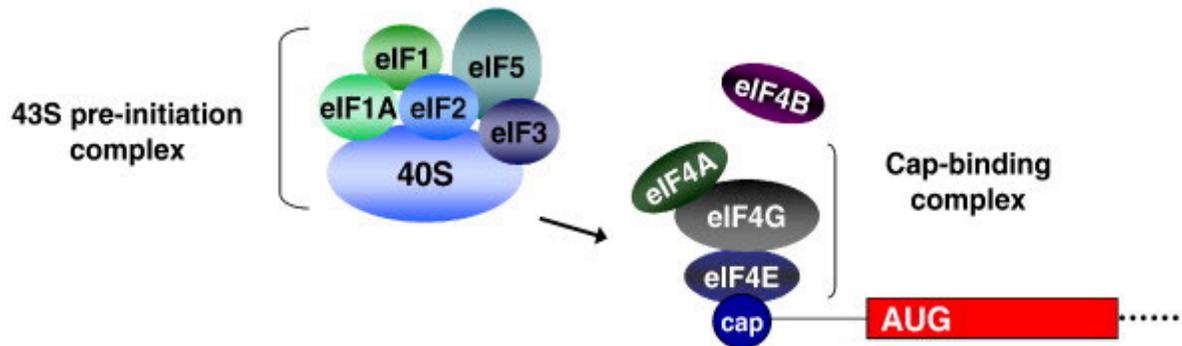


**FIG 1.3 Structural organization of the poliovirus RNA genome and proteolytic processing.** (A) The 5'NTR consists of substantial amounts of secondary structure, including the IRES element necessary for translation initiation. In addition, VPg is covalently linked to the 5' end of the genome to prime replication. The 3' NTR displays additional secondary structure as well as the poly(A) tail. Sandwiched in the middle is the single open reading frame comprised of eleven gene products. (B) The single open reading frame generates a 247kDa protein upon translation, which is subsequently cleaved by viral proteases into functional intermediates and mature proteins (from De Jesus, 2007).

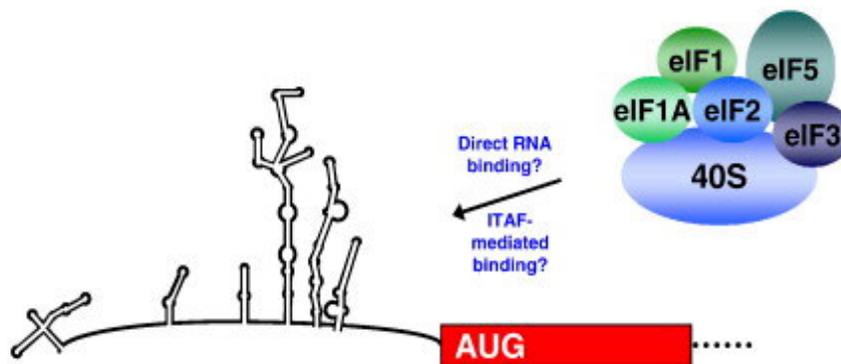


**FIG 1.4 Picornaviral Life-Cycle.** Diagram illustrating the typical life-cycle for picornaviruses. Upon infection, the genome is translated by the host cell protein synthesis machinery to generate functional intermediates and mature protein products. Non-structural proteins then carry out replication of the viral genome to produce large numbers of viral RNA that not only serve in additional rounds of translation and genome replication, but also in production of mature virions.

**A. Cap-dependent translation via cap recognition and ribosome scanning**



**B. Cap-independent translation via internal ribosome entry**



**FIG 1.5 Translation initiation mechanisms.** (A) Eukaryotes employ several initiation factors that recognize the 5' cap of mRNA in order to recruit ribosomes. (B) Poliovirus, which lacks the 5' cap, utilizes substantial secondary structure within the IRES to mediate ribosome recruitment. Currently, scientists believe this process is mediated by additional proteins thereby preventing direct interactions between the RNA and ribosome complex (Fitzgerald and Semler, 2009).

RNA molecules have VPg (virion protein, genome linked) covalently attached to their 5' ends, which must be cleaved by host cell proteases from the genome before translation can proceed (Nomoto *et al.*, 1977; Ambros and Baltimore, 1980; Ambros *et al.*, 1978). This cap-independent translation initiation is mediated by the binding of several canonical and non-canonical translation factors, such as poly(rC) binding protein 2 (PCBP2), to the stem-loop structures within the 5'-UTR (Sean and Semler, 2009). The translation factors serve to recruit host cell translation machinery to viral genome. To inhibit host cell translation, the viral proteinase 2A cleaves poly(A)-binding protein (PABP) (Joachims *et al.*, 1999) and eIF4G, which is a component of eIF-4F formerly known as p220 (Krausslich *et al.*, 1987). In addition, viral proteinase 3C inhibits host cell transcription of additional mRNA molecules by inactivating transcription factor TFIIC (Clark *et al.*, 1991) and cleaving TATA-binding protein (TBP) thereby, preventing the formation of TBP-TATA box complex (Yalamanchili *et al.*, 1996).

### 1.2.3 Proteolytic Processing

Translation of the single open reading frame (ORF) of the viral genome generates a 247 kDa polyprotein that is subsequently cleaved by viral proteinases (Fig. 1.3B) (Wimmer *et al.*, 1993). Cleavage produces eleven final viral gene products and several functional intermediates that are categorized into three broad groups: P1, P2, and P3. The P1 (VP1-VP4) region encodes for four structural proteins that later compose viral capsids while the P2 (2A-2C) and P3 (3A-3D) regions result in seven non-structural proteins necessary for protein processing and genome replication. The cleavage events occur both co- and post-translationally through activity of three viral proteinases: 2A<sup>pro</sup>, 3C<sup>pro</sup>, and 3CD<sup>pro</sup> (Fields 2007).

Two primary cleavage events occur co-translationally *in cis*: 2A<sup>pro</sup> cleaves between P1 and P2 (VP1 and 2A), releasing the capsid moieties from the non-structural proteins, and 3C<sup>pro</sup> cleaves between P2 and P3 at the 2C-3A junction. In poliovirus, 2A<sup>pro</sup> cleaves at tyrosine-glycine residues while 3C<sup>pro</sup> cleaves at glutamine-glycine pairs (Fields, 2005). After the initial co-translational proteolytic cleavage, 2A<sup>pro</sup> liberates itself from the polyprotein through self-cleavage to act on host cell proteins (as previously discussed), while 3C<sup>pro</sup> cleaves the 3AB-3CD junction to generate 3CD<sup>pro</sup>, which carries out additional *in trans* cleavage events.

The three viral proteinases display various levels of catalytic efficiency, depending on the environment found at the cleavage sites, such as steric accessibility and the surrounding amino acid sequence. The differences in efficiency yield disproportionate amounts of precursor and mature proteins throughout the course of the viral infection. The stability and function of many of the precursors allows tight control of the progression of the viral life cycle, despite the limited coding capacity of the small viral genome. As an example of this concept, the P3 region quickly undergoes two initial cleavage events to generate 3AB and 3CD. However, the subsequent cleavage of 3AB and 3CD is less efficient, resulting in a slow increase of the mature products 3A, 3B, 3C, and 3D. This slower pace provides 3CD<sup>pro</sup> sufficient time to carry out necessary cleavage events, such as the processing of the P1 coding sequence to generate the capsid proteins VP4, VP3, VP2, and VP1. The importance of this arises from 3CD<sup>pro</sup>'s increased efficiency for cleaving the P1 domain as compared to 3C<sup>pro</sup> (Jore *et al.*, 1988; Ypma-Wong *et al.*, 1988). This serves as direct evidence for the importance of differences in cleavage efficiency and how proteolysis can regulate protein function.

#### 1.2.4 RNA Replication

Proteolytic processing generates the proteins necessary for viral propagation, including one key enzyme responsible for replication of the viral genome: the RNA-dependent RNA polymerase (RdRP) known as 3D polymerase ( $3D^{pol}$ ) in picornaviruses. This 52 kDa protein carries out negative strand synthesis from the infecting genomic RNA and produces positive strands (from the negative sense template) that will serve for continued viral protein translation and become packaged into new virions. RNA replication occurs in membrane anchored replication centers. Viral infection results in the destruction of the endoplasmic reticulum and golgi apparatus and the subsequent formation of double-membraned vesicles within the host cell (Schlegel *et al.*, 1996). On the cytosolic face of these vesicles, viral proteins  $3D^{pol}$ ,  $3CD^{pro}$ , 3AB, and 2C gather to form replication centers and mediate RNA replication (Fig. 1.6) (Egger *et al.*, 2000). Membrane localization of the necessary viral proteins serves to protect the RNA genome from degradation by the host cell and to increase the local concentration of the replication machinery, thereby ensuring efficiency in the process (Fields, 2005).

To reiterate, negative sense RNA must first be synthesized from the infecting genomic RNA before high levels of positive stranded RNA molecules can be produced. However, upon insertion of the genome into the cytoplasm, host cell ribosomes bind and translate the RNA, a process that polymerase is unable to disrupt (Barton, 1999). Therefore, the virus switches from translation to replication through the buildup of viral protein  $3CD^{pro}$ , which serves not only as a protease, but also as an RNA binding protein (Andino *et al.*, 1990).  $3CD^{pro}$  binds to the 5'-cloverleaf structure and inhibits translation by preventing binding of the host cell translational machinery, namely eIF4G, and thereby allowing for replication to occur (Gamarnik and Andino,



1998). The current model suggests 3CD mediates circularization of the genome through an RNA-protein-protein-RNA interaction that draws the 5'-cloverleaf and 3'-stem loop together through interactions with cellular poly(C) binding protein 1 (PCBP1) and poly(A) binding protein (PABP) in a ribonucleoprotein complex (Fig. 1.6) (Herald and Andino, 2001). Genome circularization not only prevents further host cell translation, but also must occur for RNA replication to proceed (Herald and Andino, 2001; Lyons *et al.*, 2001). In addition, 2A<sup>pro</sup>, which is required for increased RNA stability, may stimulate negative-strand synthesis (Jurgens, 2005).

Similar to other polymerases, picornaviral RdRPs require a primer for initiation of replication (Flanegan and Baltimore, 1977). Previous studies have demonstrated the viral genome as well as newly synthesized plus and minus strand RNA are covalently linked to VPg, a 22 amino acid peptide corresponding to 3B (Nomoto *et al.*, 1977). The subsequent model for initiation of both positive and negative strand synthesis includes VPg as the primer for 3D<sup>pol</sup> (Wimmer, 1982). Interestingly, poliovirus-infected cells contain VPg as well as an uridylylated form: VPg-pUpU (Crawford and Baltimore, 1983). Uridylylation of VPg occurs at tyrosine 3 and is catalyzed by 3D<sup>pol</sup> (Paul *et al.*, 1998). The various forms of VPg allow for differentiation between positive and negative strand syntheses. The hydroxyl group of Tyr3 in VPg serves to prime negative strand synthesis while VPg-pUpU<sub>OH</sub> primes positive strand initiation. (Fig. 1.6) (Murray and Barton, 2003). The initial study showing 3D<sup>pol</sup> catalyzes the uridylylation reaction also demonstrated the necessity of a poly(A) template, suggesting that the 3' poly(A) tail of positive sense RNA serves as the template for uridylylation of at least minus strand synthesis (Paul *et al.*, 1998). However, a cis-replicating element (*cre*) within the 2C coding region of the ORF displays a conserved 5'-AAAC-3' sequence within a stem loop (Yang *et al.*, 2004) that templates uridylylation with greater efficiency than a poly(A) oligomer (Paul *et al.*, 2000; Rieder

*et al.*, 2000). Furthermore, mutating the *cre* abolishes the formation of VPg-pUpU despite the presence of a poly(A) tail. This led to a model that employs *cre* for templating uridylylation, rather than the poly(A) tail at the 3'NTR to allow for positive strand initiation (Fig. 1.6)(Murray and Barton, 2003). The use of the *cre* to initiate replication serves to ensure replication centers are composed of viral RNA instead of host mRNA (Paul, 2002).

Replication requires genome circularization as mediated by 3CD<sup>pro</sup>. This process may serve to recruit 3AB as another component of the ribonucleoprotein complex (Xiang *et al.*, 1995). Recruitment would place 3AB in close vicinity to 3CD<sup>pro</sup>, allowing for the cleavage that generates VPg for initiation of negative strand synthesis at the poly(A) tail. Genomic RNA isolated from replication centers remains in a duplex known as the replicative form, comprised of the complimentary negative and positive sense strands (Baltimore *et al.*, 1964; Barton and Flanagan, 1997). Initiation of positive strand RNA synthesis leads to melting of the duplex, leaving the negative strand to serve as a replicative intermediate to template additional rounds of replication. The buildup of VPg-pUpU primes the synthesis of large levels of positive strand resulting in approximately 50,000 copies of positive sense RNA per cell (Fields, 2005). This large copy number illustrates the stoichiometric asymmetry of replication in that positive stranded RNA molecules outnumber negative sense RNA molecules, which serves as the replication intermediate, by 50:1 (Novak and Kirkegaard, 1991). After synthesis, the positive stranded RNA can then serve in further rounds of translation to produce additional viral proteins for virion production, become replicated to increase levels of negative sense RNA, or packaged into virions.

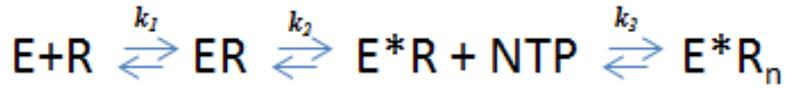
## 1.3 Mechanism and Kinetics of Picornaviral Polymerases

### 1.3.1 Polymerase Kinetics

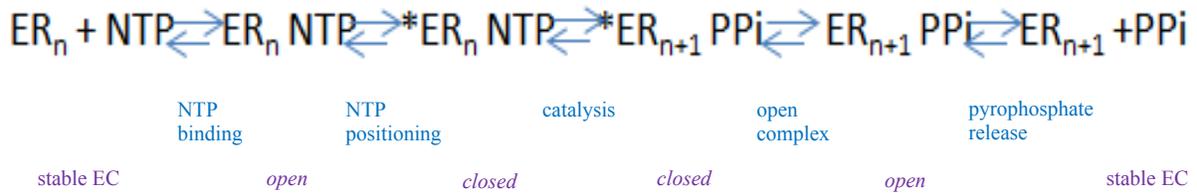
The RNA-dependent RNA Polymerase (RdRP) that catalyzes the synthesis reaction of both negative and positive strand RNA comes from the “D” protein within picornaviral polyprotein giving rise to the name 3D<sup>pol</sup>. As previously described, *in vivo* viral genome replication occurs in the context of replication centers and requires several additional viral enzymes. In order to efficiently study the mechanism of RNA replication by 3D<sup>pol</sup>, techniques have been developed to allow for replication using purified polymerase without the presence of any other viral proteins. Of special significance to this thesis has been the development of self-priming hairpin RNA constructs known as Primer Elongation Template Elements (PETEs) for use in determining the kinetics of 3D<sup>pol</sup> (Gong *et al.*, 2009; Mestas *et al.*, 2007). Using PETE prevents the need to include a separate primer strand in the reaction, thereby eliminating a second 3'-OH inherent with the annealing event and ensures duplex integrity even at lower RNA concentrations. Therefore, PETEs have only one priming site for 3D<sup>pol</sup> binding and initiation. In addition, PETE RNA constructs can be readily modified through the addition of molecular probes that allow for either fluorescent or infrared detection (Gong *et al.*, 2009; Mestas *et al.*, 2007). Using assays that utilize PETE and various other biochemical techniques, a better understanding has been developed regarding the three different phases of RNA replication: initiation, elongation, and termination.

The overall process of initiation involves two discrete steps: the binding event between 3D<sup>pol</sup> and the RNA substrate followed by incorporation of the first nucleotide (Fig. 1.7). Initiation results in the formation of an active and stable elongation complex (EC) capable of

## Initiation



## Elongation



**FIG 1.7 Kinetic mechanisms for initiation and elongation by 3D<sup>pol</sup>.** The kinetic mechanism for initiation (*top*) entails three different steps. First, the enzyme (E) binds the RNA (R) substrate ( $k_1$ ) followed by a conformational change in the enzyme ( $k_2$ ) and the slow incorporation of the first NTP ( $k_3$ ) to form a very stable elongation complex capable of elongating. The kinetic mechanism for elongation (*bottom*) occurs after initiation, and is a five step mechanism, each step is represented by blue arrows and the functions are described below also in blue. The polymerase states are described in purple.

rapidly elongating a template RNA with high processivity. The Cameron lab has performed extensive studies in order to characterize the initial binding event. By employing a symmetrical primer – template substrate (sym/sub) RNA complex, they found 3D<sup>pol</sup> binds RNA substrate slowly with an association rate of  $0.1 \mu\text{M}^{-1}\text{s}^{-1}$ . Binding is followed by an isomerization event that primes the complex for catalytic competence, which occurs with a rate of  $0.076 \text{ s}^{-1}$  (Arnold and Cameron, 2000). The 3D<sup>pol</sup>-RNA complex dissociates at a rate of  $0.1 \text{ s}^{-1}$ , with an equilibrium dissociation constant in the low  $\mu\text{M}$  range, although the latter is dependent upon experimental conditions such as pH, temperature, and RNA structure (Arnold and Cameron, 2000; Mestas *et al.*, 2007). After binding the RNA primer/template substrate, 3D<sup>pol</sup> incorporates the first nucleotide with an extremely slow rate of  $0.06 \text{ s}^{-1}$  in comparison to subsequent incorporation events that occur at a rate of  $0.0001 \text{ s}^{-1}$ . Although this first incorporation event is ~600-fold slower, the resulting EC becomes highly stable with a half-life of 8 h at 22°C (Arnold and Cameron, 2000). Upon formation of the EC, the polymerase is poised to rapidly replicate a viral RNA genome, suggesting some conformational changes must occur to prime the polymerase for genome replication. Despite the availability of a crystal structure of the poliovirus EC, evidence for these conformational changes remains undetermined (Gong and Peersen, 2010).

Similar to initiation, the overall mechanism of 3D<sup>pol</sup> elongation can be dissected into discrete steps that constitute a catalytic cycle for incorporation of a single nucleotide (Fig. 1.7). The five step cycle includes 1) nucleotide triphosphate binding, 2) a conformational change to reposition the NTP for catalysis, 3) catalysis and phosphodiester bond formation, 4) a second conformational change and 5) pyrophosphate release and translocation. Of these five steps, nucleotide repositioning (2) and catalysis (3) appear to be rate limiting if the complex is provided sufficient amounts of nucleotides. To determine the rate of nucleotide incorporation, the Peersen

lab has used a PETE RNA construct with a 26-nucleotide template for stopped-flow kinetics (Gong *et al.*, 2009). At 37°C and pH 7.5, PV 3D<sup>pol</sup> incorporates approximately 50 nt/sec when presented with 120 μM of each NTP. Factoring in increased intracellular NTP concentrations, 3D<sup>pol</sup> would replicate a poliovirus genome of 7500 nucleotides in less than 100 seconds at a rate of approximately 70-80 nt/sec (Gong *et al.*, 2009). Although initiation and elongation have been well characterized, there is no known mechanism for elongation termination aside from random dissociation and runoff of the polymerase from the end of the RNA.

### 1.3.2 Polymerase Fidelity

As is observed in many RNA viruses, 3D<sup>pol</sup> fidelity during genome replication is very low due to the lack of a proofreading mechanism. In contrast, DNA polymerases have 3'-exonuclease activity that enables them to correct a nucleotide misincorporation event. Errors occur at a frequency of approximately  $10^{-3}$  to  $10^{-5}$  per nucleotide incorporated, a value roughly a million fold higher than the error frequency approximated for cellular DNA replication (Drake, 1969; Drake 1993). The high error rate associated with RNA replication suggests that every genome contains at least one mutation (Drake and Holland, 1999). Therefore, one infected cell generates an assortment of closely related viruses, yielding a diverse population known as a quasispecies (Domingo *et al.*, 1997). Generation of a quasispecies allows viral evolution and adaptation to selective pressures, such as host cell defense mechanisms (Vignuzzi *et al.*, 2006). The error frequency responsible for the quasispecies lies at the threshold of “error catastrophe.” Slower error rates prevent viral evolution while faster rates result in a large proportion of non-viable viruses that prevents propagation (Crotty *et al.*, 2004). Scientists have exploited this feature of viral replication by using a mutagenic nucleotide analogue known as ribavirin to

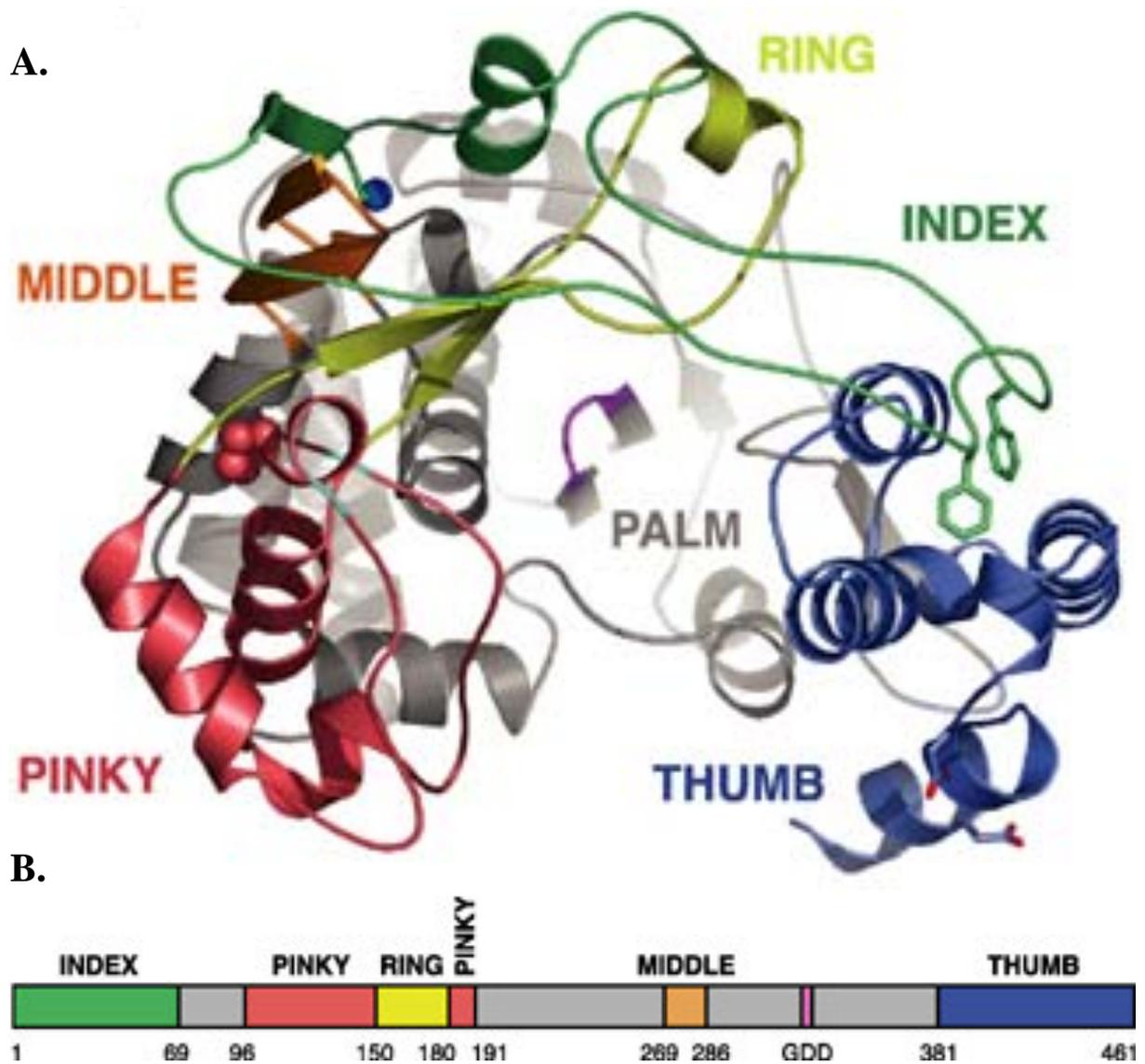
increase error rates past the threshold of error catastrophe (Pfeiffer and Kirkegaard, 2005). This small molecule is widely used in experimental studies, but also serves as drug for combating viral infections, especially hepatitis C.

## 1.4 Structure of Picornaviral Polymerases

### 1.4.1 Apo PV 3D<sup>pol</sup> Structure

Structurally, poliovirus 3D<sup>pol</sup> folds according to the classic analogy of a cupped right hand comprised of a palm, thumb, and finger domains, first used to characterize the DNA polymerase I Klenow fragment from *Escherichia coli* (Ollis *et al.*, 1985; Hansen *et al.*, 1997). The initial x-ray crystal structure of poliovirus polymerase provided insight into structure-function relationships; however, the structure contained a disordered fingers domain, thereby precluding key information regarding the overall mechanism. To obtain a complete structure, Interface I, which entails a head-to-tail oligomerization that may have biological importance in membrane-bound replication complexes (Hobson *et al.*, 2001; Lyle *et al.*, 2002), was disrupted through point mutations L446D and R455D (Thompson and Peersen, 2004). The resulting structure, resolved to 2 Å showed continuous electron density throughout the entire protein chain, including the fingers domain (Fig. 1.8A).

The core of the palm domain in poliovirus 3D<sup>pol</sup> contains four structural motifs that are highly conserved in other polymerase structures. The domain consists of four anti-parallel  $\beta$ -sheets packed above two  $\alpha$ -helices resembling an RNA recognition motif (RRM) found in various RNA-binding proteins, including ribosomal proteins and splicing proteins (Hansen *et al.*, 1997). The active site residing within the palm domain consists of a highly conserved YGDD motif centered on a type II  $\beta$ -turn of the RRM (Kamer and Argos, 1984; Koonin, 1991).



**FIG 1.8 Structural overview of poliovirus 3D<sup>pol</sup>.** (A) Top view of the complete 3D<sup>pol</sup> structure. The structure follows the analogy of primate right hand with the thumb in blue, the index finger in green, the middle finger in orange, the ring finger in yellow, and the pinky finger in pink, and finally the palm in gray. (B) Diagram of the amino acid sequence of 3D<sup>pol</sup> colored according to the structural elements shown in (A). The numbers correspond to the specific residues where domains begin and end (from Thompson and Peersen, 2004).

Replacing tyrosine 326 with either phenylalanine or methionine yields active polymerase in *in vitro* studies, although only Y326F produces viable viruses in transfection studies (Jablonski *et al.*, 1993). Mutation of the conserved glycine residue to either alanine or serine results in reduced but definitive activity, while mutagenesis to other residues abolishes enzyme activity (Jablonski *et al.*, 1991). Finally, mutation of one or both aspartic acids produces completely inactive enzymes, even when mutated to glutamic acid. However, a D329N mutation displays *in vitro* activity in the presence of  $Mn^{2+}$  or  $Fe^{2+}$  instead of  $Mg^{2+}$  as the divalent metal cation (Jablonski *et al.*, 1995). Overall, the conservation of the YGDD motif is required for a functional polymerase.

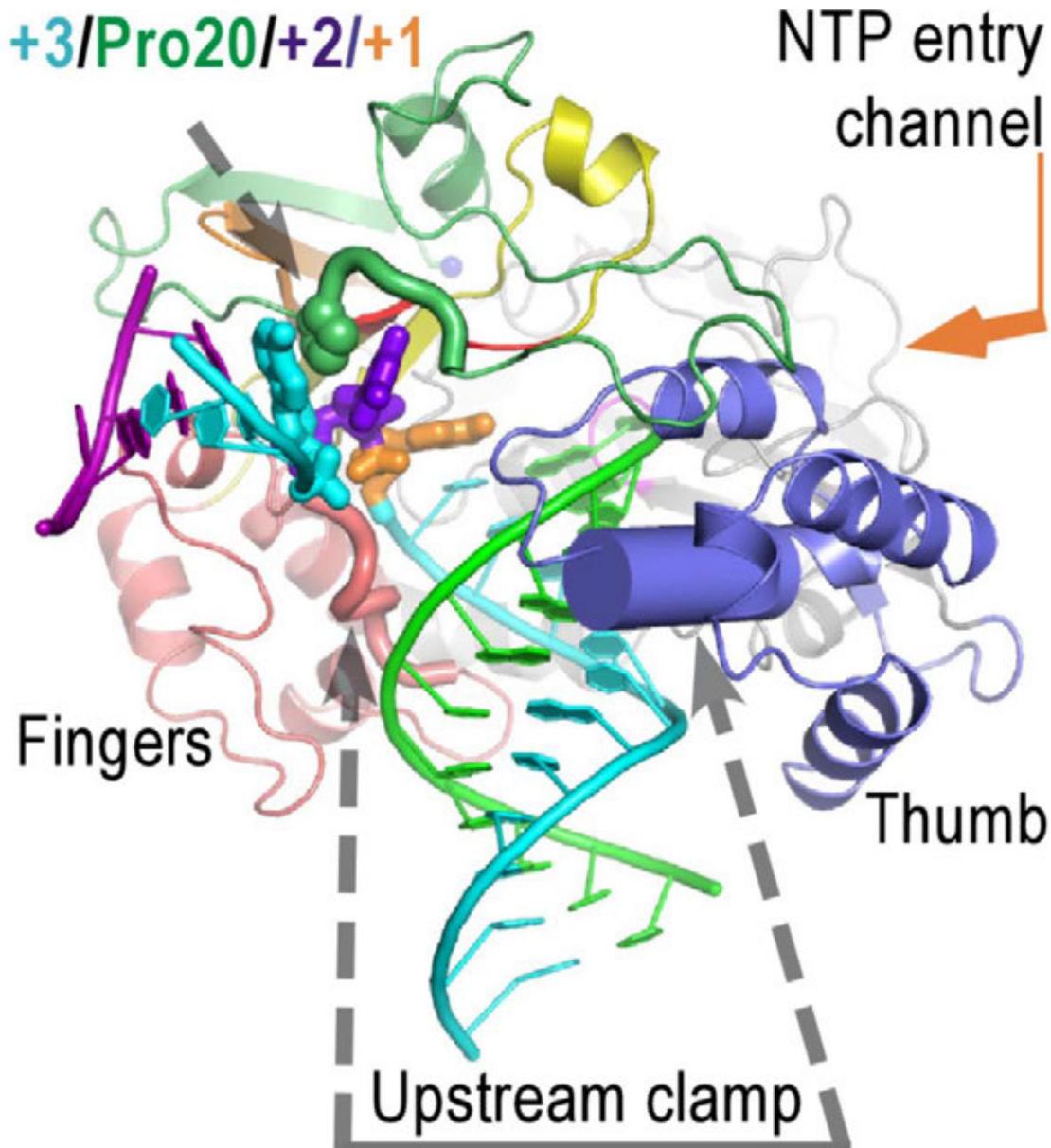
In addition to the two aspartates within the active site, Asp233 located on an adjacent  $\beta$  strand on the “back” of the palm domain is highly conserved and crucial in the two-metal-ion binding mechanism for catalysis (Steitz, 1998). This model involves two divalent metals (A & B) that each contributes in a different manner. Metal A serves as a Lewis acid that activates the 3'-OH group to an  $O^-$ , which allows for nucleophilic attack on the  $\alpha$ -phosphate of the incoming rNTP, resulting in phosphodiester bond formation. Metal B coordinates with the  $\beta$  and  $\gamma$  phosphates of the triphosphate moiety to hold it in a catalysis competent conformation. The conserved Asp residues serve to chelate the two divalent metals required for the phosphoryl transfer reaction (Steitz, 1998; Sosunov *et al.*, 2005).

The complete poliovirus polymerase structure reveals an enclosed active site formed by interactions between the thumb and fingers domain; specifically, phenylalanines 30 and 34 insert into a hydrophobic pocket at the top of the thumb (Thompson and Peersen, 2004). This interaction creates an NTP entry channel at the back of the polymerase. Furthermore, the previously uncharacterized fingers domain is divided by analogy to a primate hand with an index

finger (residues 1-68), middle finger (residues 269-285), ring finger (residues 150-179), and pinky finger (residues 96-146 and 180-190) (Fig. 1.8B) (Thompson and Peersen, 2004). The other unique and highly important feature of the polymerase is the buried N-terminal glycine residue in a pocket residing at the base of the fingers domain. This burial functions to properly position Asp238 relative to the active site, allowing for hydrogen bonding between the aspartate and the 2'OH of incoming NTPs; an important interaction for selecting rNTPs over dNTPs (Thompson and Peersen, 2004; Huang *et al.*, 1997). Thus, proteolysis of 3CD<sup>pro</sup> into mature 3D<sup>pol</sup> serves as an allosteric switch that activates 3D<sup>pol</sup> by repositioning Asp238 1.4Å closer to the active site relative to the surrounding structure. Mutational analysis of the N-terminal glycine shows abolished enzyme activity due to the addition or removal of a single amino acid, and 2- and 50-fold decrease for G1A and G1S mutations respectively (Thompson and Peersen, 2004).

#### 1.4.2 PVEC Structure

Recently, the Peersen lab obtained additional insight regarding the mechanism and conformational changes that accompany NTP recognition, active site closure, and catalysis through co-crystallizing poliovirus 3D<sup>pol</sup> with an RNA construct (Fig. 1.9) (Gong and Peersen, 2010). Using a newly designed experimental approach, they formed a highly stable EC stalled four nucleotides into the replication cycle with significant upstream and downstream duplex RNA. The structure shows template nucleotides +1, +2 and +3 are single stranded with strand separation of the downstream duplex RNA occurring at residues 18 and 19 of the index finger. The +2 nucleotide resides in a pocket formed by Pro 20 and the index finger that prevents it from base stacking with either the +3 or +1 nucleotides. The +1 nucleotide is positioned within the active site to allow for base stacking with the -1 upstream nucleotide and base pairing with the



**FIG 1.9. Poliovirus 3D<sup>pol</sup>-RNA elongation complex structure.** The poliovirus polymerase is colored with the palm in grey and the thumb in blue while the index, middle, ring and pinky fingers are in green, orange, yellow and pink, respectively. The structure shows the template RNA strand in cyan, downstream non-template strand in magenta and upstream non-template in bright green. The single stranded template nucleotides +1, +2 and +3 are colored orange, purple and cyan, respectively. The structure also highlights proline 20 in green and the upstream clamp that forms around the exiting upstream duplex with arrows pointing to regions of the pinky and thumb that interact with the RNA (from Gong and Peersen 2010).

incoming nucleotide. The product RNA exits as an upstream duplex through a clamp formed by the pinky and thumb domains of the polymerase.

Through soaking experiments with cognate CTP and CTP analogs, the elongation complex was trapped in different stages of the overall catalytic cycle, thereby providing insight into the conformational changes associated with catalysis. The catalytic cycle proposed after analysis of the elongation complex showed the absence of a major nucleotide-repositioning step that is commonly seen in other polymerases. Rather, NTPs bind to an *open* conformation of the polymerase that *closes* for the subsequent catalysis via a subtle repositioning at the palm domain. The driving force for this conformational change arises from a shift in the palm domain that repositions Asp 233 to recognize the ribose of the NTP. This differs from other polymerase structures, which utilize large movements in the fingers domain. The lack of the NTP-repositioning step reduces polymerase fidelity by removing a key step for selecting the appropriate nucleotide. This, in addition to the absence of a proofreading mechanism, may account for the rapid viral evolution that occurs in positive-sense RNA viruses. In addition, based on the structures from this study, the Peersen lab proposes a sixth step in the catalytic cycle of nucleotide incorporation constituting the translocation event.

Viral RdRPs play an essential role in viral propagation and deserve our attention as a potential target for antivirals. By elucidating the mechanism and structure of this enzyme, we gain a better understanding of replication processes in general, as well as valuable information regarding the viral life cycle. Specifically this information may assist in the design of antivirals and therapeutics with greater efficacies against both poliovirus and coxsackievirus.

## **Chapter 2**

### **Role of Charge and Hydrophobic Residues in the**

### **Poliovirus Polymerase Elongation Complex**

#### **2.1 Introduction**

Positive-strand RNA viruses encompass a wide range of serious human and animal pathogens that cause diseases including the common cold, liver disease, heart disease, meningitis, encephalitis, and paralytic poliomyelitis. As a group, these viruses encode for and require a viral RNA-dependent RNA polymerase (RdRP) for genome replication. This process involves replication of the positive sense genome into a negative sense intermediate that in turn templates the replication of additional positive sense strands. These positive-sense strands are subsequently used for additional rounds of translation or packaged as genomic RNA into the forming virions.

Viral RdRPs must form a stable and processive elongation complex (EC) to replicate the entire genome. Formation of this complex begins with RNA binding to form a polymerase-RNA complex, followed by a conformational change that poises the complex for incorporation of the first nucleoside triphosphate (NTP). After the addition of the first nucleotide, the complex undergoes another conformational change and converts to what is known as the elongation complex that is capable of rapidly replicating the viral genome. The rate limiting step of forming the elongation complex is followed by catalytic cycles for addition of each nucleotide. In this cycle, binding of the appropriate NTP to an open active site is followed by a conformational

change that closes the active site to allow for catalysis. After catalysis, the active site must reopen to allow for pyrophosphate release and subsequent translocation of the RNA. This results in an open post-translocation state capable of binding the next NTP.

Structurally, viral RdRPs contain an approximately 55 kDa core that folds in a conformation resembling a cupped right hand with palm, thumb, and fingers domains (Thompson and Peersen 2004). The finger domains reach across the palm domain to make interactions with the top of the thumb domain, encircling the active site and creating a loop that NTPs can be shuttled through for catalysis. This interaction between the fingers and thumb does not occur in other DNA- or RNA-dependent polymerase and serves to tether the two domains together. Overall, this interaction prevents the two domains from undergoing large-scale independent movements. Many DNA- and RNA-dependent polymerases utilize large conformational changes in the fingers and thumbs to open and close the active site during catalysis. Viral RdRPs must rely on rearrangements in the palm domain to facilitate this motion (Gong and Peersen, 2010). The driving force behind the conformational changes comes from NTP recognition. Incoming nucleotides bind the active site through base pairing interactions with the +1 template nucleotide. The subsequent ribose hydroxyl recognition causes the structural rearrangement in the palm that enables catalysis to occur. This method for active site closure does not involve movements in the fingers domain (Gong and Peersen, 2010).

Analysis of the poliovirus EC structure shows template RNA interacting with the fingers domain before reaching a conserved cleft between the pinky and index fingers that facilitates strand separation (Gong and Peersen, 2010). Although the structure shows definitive interactions between the polymerase and the first three base pairs of the downstream duplex, the interactions between the fingers domain of 3D<sup>pol</sup> and additional downstream duplex are poorly defined. A

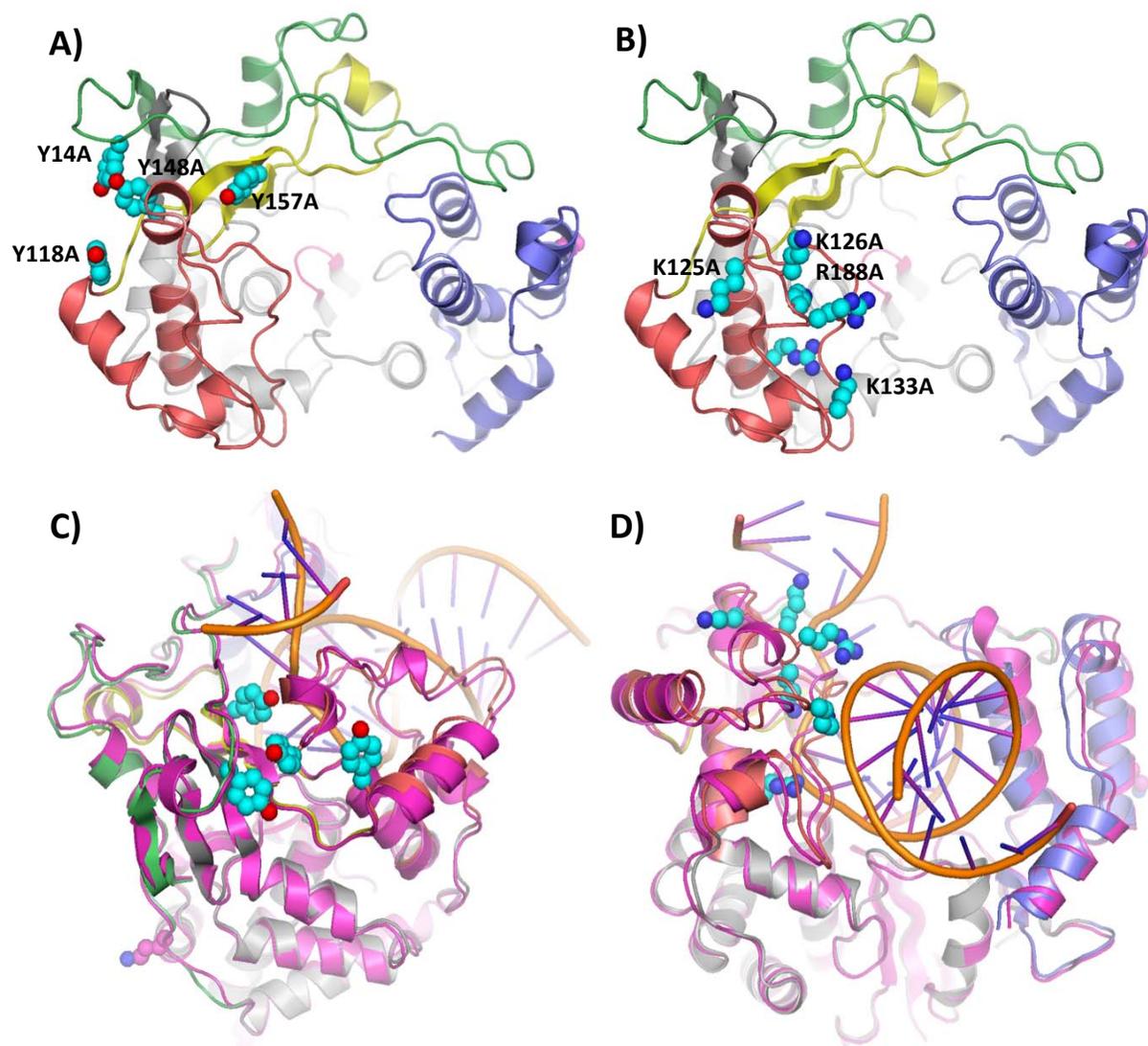
comparison of the fingers domain in the apo-3D<sup>pol</sup> structure and the elongation complex structure reveals no key differences in conformation between the two (Thompson and Peersen, 2004; Gong and Peersen, 2010). However, evidence from a mutational analysis of Trp5 suggests the fingers domain plays a key role in interacting with downstream template RNA (Hobdey and Peersen, 2010). This provides evidence for possible interactions between the downstream RNA and the fingers domain of the polymerase.

In this study, we investigated the role of several basic and hydrophobic residues within the fingers domain of poliovirus polymerase in an attempt to characterize their role in RNA replication (Fig. 2.1). We analyzed the effect that mutations to these residues have on RNA binding, elongation complex formation rates, elongation rates, and elongation complex stability. The data show that charged residues have a significant role in binding RNA, while the mutations to the hydrophobic residues show no significant effect on binding. In contrast, the hydrophobic residues, especially Y118 and Y148, play key roles in forming a stable elongation complex. Both types of mutations displayed no effect on elongation rates, while both types of mutations demonstrated a significant decrease in the stability of the polymerase elongation complex.

## 2.2 Material and Methods

### 2.2.1 Mutant Generation and Protein Expression.

3D<sup>pol</sup> constructs were expressed from the pET26-UbDH plasmid that was obtained from the Cameron lab at Penn State University. This plasmid contains the complete open reading frame of the poliovirus 3D<sup>pol</sup> gene sandwiched between an N-terminal ubiquitin (Ub) and a C-terminal GSSS-His<sub>6</sub> tag (Gohara *et al.*, 1999). Expression from this plasmid yields 3D<sup>pol</sup> fused with ubiquitin. After *in vivo* translation, Ubp1, an ubiquitin-specific carboxyl-terminal protease that is



**FIG 2.1 Structural overview of basic and hydrophobic residues.** PV 3D<sup>pol</sup> colored with the palm in grey and the thumb in blue while the index, ring and pinky fingers are in green, yellow and pink, respectively. A) Top view of the structure highlighting the four tyrosine residues mutated to alanines for this study. B) Top view of the structure highlighting the arginine and lysine residues mutated to alanines for this study. C) Tyrosine mutations in context of the PV elongation complex shown in magenta with RNA in orange. D) Lysine and arganines mutations in context of the PV elongation complex shown in magenta with RNA in orange (Thompson and Peersen, 2004; Gong and Peersen 2010).

co-expressed in *E. coli* pCGI cells, cleaves ubiquitin producing full-length 3D<sup>pol</sup> with an authentic N-terminal that has glycine as the first residue (Gohara *et al.*, 1999). In addition, the 3D<sup>pol</sup> gene contains a 3D<sup>pol</sup> L446D point mutation that effectively disrupts the Interface I interaction allowing the polymerase to exist as a monomer (Hobson *et al.*, 2001; Lyle *et al.*, 2002; Thompson and Peersen, 2004).

The appropriate mutations were cloned into the pET26-UbDH plasmid using primers synthesized by Integrated DNA Technologies and designed for Quick-Change site-directed mutagenesis (Invitrogen). Mutations were verified through sequencing, which is performed by the Proteomics and Metabolics Facility at Colorado State University. The mutant clones were subsequently transformed into *Escherichia coli* BL-21 DE3 pCGI cells for expression. Colonies from the transformation were used to inoculate 5 mL LB media and grown at 37° C in the presence of 50 µg/mL kanamycin and 34 µg/mL chloramphenicol until the optical density at 600 nm reached ~1.0. The culture were then used to inoculate 1L LB media and once again grown at 37°C and in the presence of 50 µg/mL kanamycin, but without chloramphenicol. Once the optical density at 600 nm reached ~0.5 the cells were cooled to room temperature and expression was induced through addition of IPTG to a final concentration of 0.4 mM. After an additional 14-18 hours of growth, the cells were pelleted through centrifugation at 4000 rpm in a model J-6 centrifuge (Beckman).

### 2.2.2 Protein Purification.

The resulting pellet from centrifugation were resuspended in lysis buffering consisting of 50 mM Tris (pH 8.0), 300 mM NaCl, 20% (vol/vol) glycerol, and 0.02% (wt/vol) NaN<sub>3</sub> and lysed at 18,000 lbs/cm<sup>2</sup> in a model M-110L microfluidizer (Microfluidics). NP-40 was added to

the lysate in a drop-wise manner to a final concentration of 0.1% (vol/vol) followed by DNA precipitation through the addition of PEI to a final concentration of 0.25% (vol/vol). The solution was centrifuged at 17,000 rpm for 45 minutes in a Sorvall RC2-B centrifuge. The supernatant was filtered through a 0.2 micron cellulose acetate filter, loaded onto a HisTrap HP nickel column (GE Healthcare), and eluted with 350 mM imidazole in buffer containing 50 mM Tris (pH 8.0), 300 mM NaCl, 20% (vol/vol) glycerol, and 0.02% (wt/vol) NaN<sub>3</sub>. Fractions containing the polymerase were pooled and diluted three-fold in low salt buffer consisting of 50 mM NaCl, 25 mM Tris (pH 8.5), 20% (vol/vol) glycerol, 0.02 (wt/vol) NaN<sub>3</sub> to allow for loading onto a HiTrap Q HP column (GE Healthcare). Polymerase was eluted via a linear gradient of 50 mM to 1 M NaCl, pooled and concentrated to a volume ~0.8 mL using 30 KDa MWCO centrifugal concentrator (Amicon) and a model 5810 R centrifuge (Eppendorf). The concentrated sample was loaded onto a Mono Q column (GE Healthcare) and eluted with a linear gradient using the same buffer conditions as for the Q-column. Fractions containing the polymerase were pooled, concentrated as before to a volume of ~0.8 mL, and loaded onto a Superdex 200 gel filtration column (GE Healthcare) equilibrated in 200 mM NaCl, 5 mM Tris (pH 7.5), 20% glycerol (vol/vol), and 0.02% (wt/vol) NaN<sub>3</sub>. Fractions containing the polymerase were pooled, provided 5 mM TCEP, concentrated to a final [3D<sup>pol</sup>] of 200 μM, flash frozen with liquid nitrogen, and stored at -80°C. Purification efficiency and sample purity were assessed by running samples from each step of the process on 12% acrylamide PAGE.

### 2.2.3 RNA Oligonucleotides

The RNA constructs used in all four assays were synthesized by Integrated DNA Technologies ([www.idtdna.com](http://www.idtdna.com)). The two different oligonucleotides used in the binding assay and elongation rate study contain a 5' fluorescein added by IDT. The constructs are suspended in TE buffer

(10mM Tris (pH 8.0), 1mM EDTA), heated to 95°C for three minutes, and snap cooled on ice to fold the RNA into the primer-template hairpin to allow for elongation. The RNA oligonucleotides used in the initiation and stability assays contain an amino modifier deoxythymidine residue within the loop of the RNA hairpin labeled with an IRdye 800RS NHS ester (LI-COR Biosciences) and folded into the primer-template form as described above.

#### *2.2.4 RNA Binding.*

RNA binding assays were performed using an “8-10” polymerase elongation template element (PETE) RNA consisting of an eight base pair (bp) hairpin followed by a ten-nucleotide (nt) template with a fluorescein labeled to the 5’end of the construct. 10 nM of RNA is mixed with 50 mM Hepes (pH 6.5), 75 mM NaCl, 1.5 mM MgAc<sub>2</sub>, 60 μM ZnCl<sub>2</sub>, 4 mM dithiothreitol (DTT), and 0.1% (vol/vol) NP-40. 3D<sup>pol</sup> was titrated into the RNA mix giving a range of protein concentrations from 1.0x10<sup>-8</sup> M to 3.5 x10<sup>-5</sup> M and incubated at 4°C for 30 minutes, followed by additional 30 minutes of incubation at room temperature. Fluorescence polarization (FP) was measured in a black 384-well polystyrene plate using a Perkin-Elmer Victor V multimode microplate reader. Using Kaleidagraph (Synergy Software), we constructed plots of FP versus [3D<sup>pol</sup>], and fit the data to single-site binding isotherm with a Hill coefficient to determine RNA dissociation constants.

#### *2.2.5 Elongation Complex Formation.*

To determine the rate for elongation complex formation we employed two LI-COR labeled RNA constructs that allow for formation of either a +1 or +2 product in the same reaction. The 10+1.12 PETE consists of a 10-bp hairpin primer, allows for incorporation of one nucleotide leading to a “locked” elongation complex, and leaves a 12-nt single-stranded template. The

9+2.24 PETE consists of a 9-bp hairpin primer, allows for incorporation of two nucleotides resulting in a “locked” elongation complex, and leaves a 24-nt downstream template. In addition, both RNA molecules lack a guanosine within the templating region until the sixth nucleotide from the 5' end. This allows us to pause elongation by omitting CTP, thereby preventing the polymerase from reaching the end of the template and avoiding end effects. The RNA constructs were mixed to a final concentration of 0.5  $\mu\text{M}$  each with 15  $\mu\text{M}$  3D<sup>pol</sup>, 40  $\mu\text{M}$  of each ATP and GTP in a reaction buffer consisting of 50 mM Hepes (pH 6.5), 75 mM NaCl, 1.5 mM MgCl<sub>2</sub>, and 5 mM TCEP. Reactions were quenched through equal volume mixing with quench buffer comprised of 50 mM Hepes, 75 mM NaCl, and 10 mM EDTA after various amounts of time. Samples were separated through gel electrophoresis on a 15% acrylamide, 7M Urea PAGE gel, analyzed with a LI-COR Odyssey infrared imager 9120, and quantified using Odyssey software. Plots of percent of total RNA elongated versus time were constructed using Kaleidagraph and fit to a single exponential curve providing an apparent  $K_m$ .

### *2.2.6 Elongation Rate.*

RNA elongation assays utilized the 10.26 PETE RNA that consists of a 10-bp hairpin and a 26-nt template with a fluorescein attached to the 5' end of the construct. 1  $\mu\text{M}$  RNA was premixed with 15  $\mu\text{M}$  3D<sup>pol</sup>, 40  $\mu\text{M}$  ATP and incubated at room temperature for 30 minutes to allow formation of the locked +1 elongation complex in the same buffer conditions used for the elongation complex formation assay. The resulting solution was diluted 10-fold in reaction buffer and stored on ice until ten minutes prior to use, at which point the reaction was placed in a water bath set to 22.5°C. Using an applied Photophysics SX-20 stopped-flow instrument, the preformed elongation complex was rapidly mixed with an equal volume of the remaining nucleotides to allow for full elongation of the 10.26 RNA. The final solution measured in the

instrument contains 50 nM RNA, 750 nM 3D<sup>pol</sup>, and a range of 5  $\mu$ M-120  $\mu$ M each NTP. Fluorescence excitation occurs at 492 nm from a monochromator source and emission was detected using a 515 nm high pass filter. From collecting the parallel and perpendicular polarized fluorescence, we calculated total fluorescence (TF) and fluorescence polarization (FP) as previously described (Gong *et al.*, 2009). Using Kaleidagraph, the lag time was calculated by fitting an initial decrease and the final increase in signal to two independent single exponential curves. The intersection between these curves provides the amount of time to elongate through 20-nt, which was used to calculate rate in nt/second. The rates were plotted against substrate concentration to construct a Michaelis-Menten curve enabling calculations of  $V_{\max}$  and  $K_m$ .

### 2.2.7 Elongation Complex Stability.

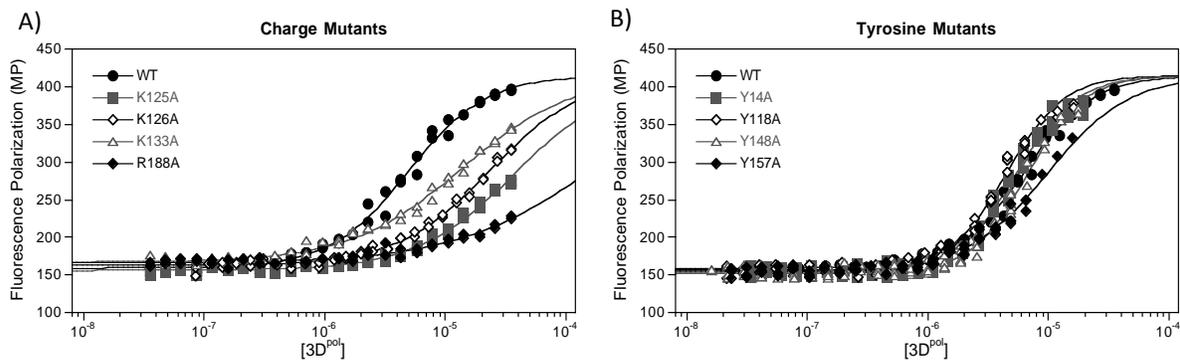
Stability studies followed the protocol for elongation complex formation with the 9+2.24 RNA, as outlined in section 2.2.4. Incubation times for complex assembly varied from 15 minutes to 90 minutes to provide ample time for complex formation based on results from formation rate studies in Table 2.1. To prevent degradation during the room temperature incubation, RiboLock RNase Inhibitor (Fermentas) was added to 3 units/uL. The reaction was then diluted ten-fold in the complex formation buffer containing 300 mM NaCl instead of 75 mM NaCl to prevent polymerase rebinding and 10+1.12 RNA at equivalent amounts as 9+2.24 to serve as a degradation control. After various amounts of time, samples were mixed with an equal volume solution that provides 40  $\mu$ M each of ATP, GTP, and UTP and incubated for an additional three minutes to allow for full elongation of the 9+2.24. Quench buffer (see 2.2.4) was added to stop the reaction and samples were separated using PAGE for analysis. Again, Odyssey software was employed to image and quantitate the gel, and stability was assessed using Kaleidagraph to calculate the percent of elongated RNA out of total RNA.

## 2.3 Results

### 2.3.1 Basic residues play role in RNA binding.

To determine the ability of the mutated polymerases to bind a synthetic RNA construct we employed a fluorescence polarization based binding assay. Titration of wild type protein into a solution containing a 5'-fluorescein labeled RNA construct shows an increase in fluorescence polarization corresponding to a decrease in the fluorescein's mobility in solution. From the data, we can construct a binding curve that yields an apparent disassociation constant ( $K_D$ ). The data show a wide range of reduced affinities, resulting from mutations to the basic residues (Fig. 2.2A). Mutation to alanine at residues K126 and K133 show only minor decreases in RNA affinity, while mutation at K125 shows a more severe decrease. The R188A mutation nearly abolished 3D<sup>pol</sup>'s ability to bind RNA. In contrast, none of the mutations to the hydrophobic residues displayed significant changes in RNA binding affinity, indicating these residues do not play a direct role in binding the templating RNA during replication (Fig. 2.2B).

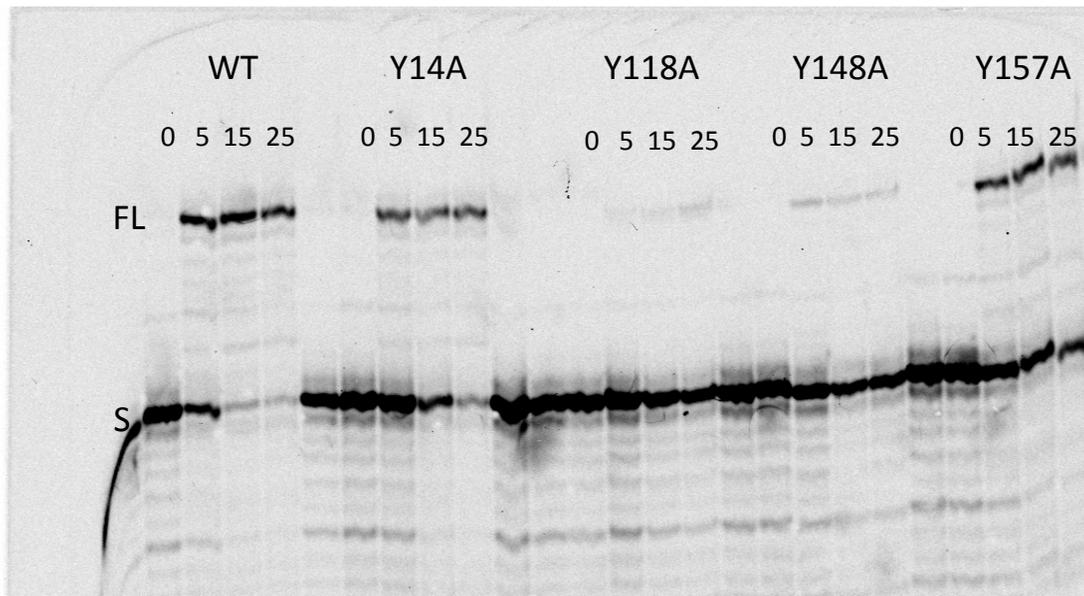
Since mutations to the selected tyrosine residues showed no effect on polymerase RNA binding affinity, we tested whether these residues displayed any effect on RNA replication. Using another synthetic PETE RNA construct, we tested these mutants' abilities to form fully elongated products after short incubation periods. The experiments showed that WT polymerase achieves full elongation after 15 minutes, while Y14A and Y157A show slowed kinetics, and Y118A and Y148A are almost devoid of elongation even after 25 minutes (Fig. 2.3). The data indicate that these hydrophobic residues have no direct effect on RNA binding, but do play some role in the overall initiation mechanism.



**FIG 2.2 RNA binding results.** Increase in fluorescence polarization upon addition of 3D<sup>pol</sup> to reaction fluorescein labeled 8.10 PETE RNA.  $K_d$  values calculated from single binding site model fits to the sigmoidal binding curves can be found in Table 2.1. The  $\approx 25$  mM solubility limit of 3D<sup>pol</sup> in these conditions prevented us from obtaining clear maximal binding plateaus at salt concentrations above 100 mM for many of the mutants. To obtain RNA dissociation constants when curve fitting such data we fixed the maximal FP signal to the stable plateau value observed with wild type protein, allowing the fitting algorithm to vary only the starting FP value and the midpoint of the binding curve (A) Comparison of wild-type 3D<sup>pol</sup> binding affinity with the mutations of basic residues. (B) Comparison of wild-type 3D<sup>pol</sup> binding affinity with the mutations of tyrosine residues.

**Table 2.1 Summary of biochemical data for PV 3D<sup>pol</sup> mutants.** NP indicates the assay was Not Performed on the given mutant. ND indicates the assay was performed, but produced No Data due to extreme effects associated with the mutation.

	<b>Binding Kd</b> <b>(<math>\mu</math>M)</b>	<b>Initiation (+1)</b> <b>(minutes)</b>	<b>Initiation (+2)</b> <b>(minutes)</b>	<b>V<sub>MAX</sub></b> <b>(nt/second)</b>	<b>K<sub>M</sub></b> <b>(<math>\mu</math>M)</b>	<b>Stability</b> <b>(minutes)</b>
WT	4.9 $\pm$ 0.1	4.8 $\pm$ 0.3	3.0 $\pm$ 0.2	3.1 $\pm$ 0.1	22 $\pm$ 3	150 $\pm$ 9
Y14A	5.6 $\pm$ 0.1	8.3 $\pm$ 0.4	8.7 $\pm$ 0.6	3.2 $\pm$ 0.1	17 $\pm$ 2	16 $\pm$ 2
Y118A	4.5 $\pm$ 0.1	99 $\pm$ 3	29.3 $\pm$ 0.8	3.34 $\pm$ 0.06	20 $\pm$ 1	< 5
Y148A	6.7 $\pm$ 0.1	230 $\pm$ 70	90 $\pm$ 6	3.24 $\pm$ 0.09	19 $\pm$ 2	< 5
Y157A	9.8 $\pm$ 0.3	10.2 $\pm$ 0.4	10.0 $\pm$ 0.6	2.9 $\pm$ 0.1	16 $\pm$ 2	26 $\pm$ 2
K125A	38.0 $\pm$ 1.0	38 $\pm$ 5	16 $\pm$ 1	NP	NP	59 $\pm$ 5
K126A	23.4 $\pm$ 0.5	20 $\pm$ 2	13 $\pm$ 1	NP	NP	58 $\pm$ 4
K133A	13.6 $\pm$ 0.4	17 $\pm$ 2	8 $\pm$ 1	NP	NP	65 $\pm$ 2
R188A	170 $\pm$ 20	ND	ND	NP	NP	ND
K66A	NP	7 $\pm$ 2	3.5 $\pm$ 0.6	3.1 $\pm$ 0.1	23 $\pm$ 4	113 $\pm$ 8
K314A	NP	7 $\pm$ 1	3.5 $\pm$ 0.6	3.1 $\pm$ 0.1	22 $\pm$ 3	110 $\pm$ 10
K431A	NP	8 $\pm$ 1	4.0 $\pm$ 0.6	3.2 $\pm$ 0.2	20 $\pm$ 4	120 $\pm$ 10

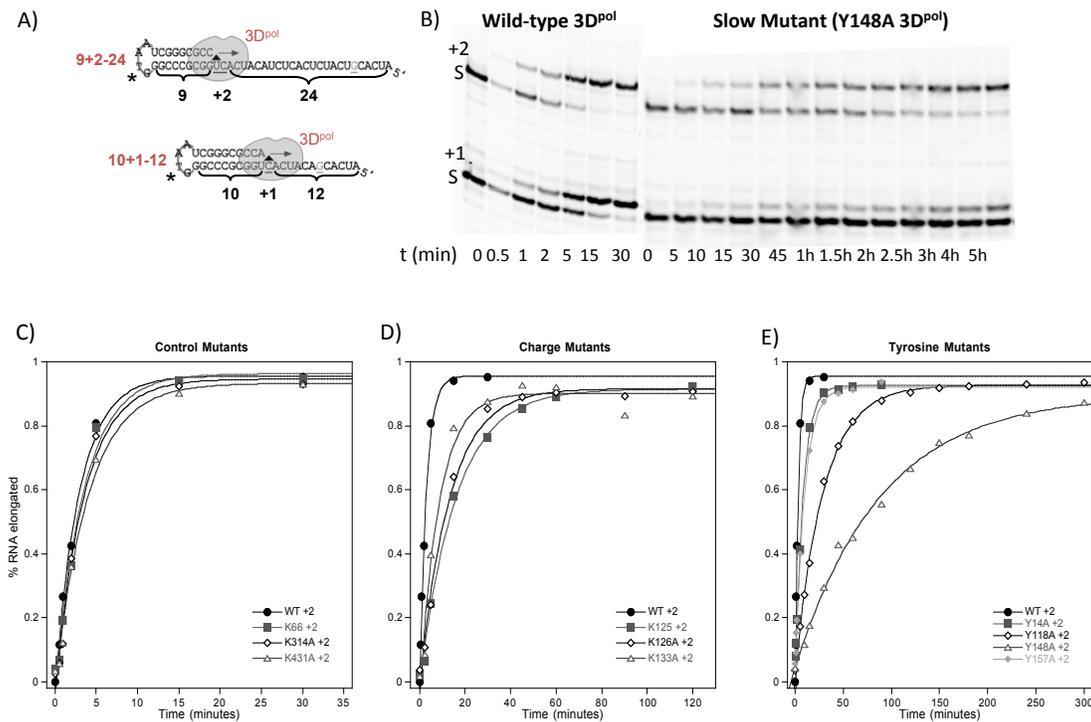


**FIG 2.3 Tyrosine residues play role in elongation.** PV 3D<sup>pol</sup> was incubated with a 10.26 PETE RNA consisting of a 10-bp hairpin primer and a 26-nt single-stranded template. Samples of the reaction were quenched after 5, 15, and 25 minutes and run over a 12% acrylamide 7 M urea gel. The gel shows the starting (S) material reaching a full-length (FL) product over time. Y14A and Y157A display slightly lower levels of activity, while Y118A and Y148A have greatly diminished activity.

### 2.3.2 Tyrosine Mutations slow EC formation rate.

To assess the roles that the tyrosine residues play in elongation and to determine if the charged residues displayed additional effects, we examined the rate at which these mutants form the stable elongation complex. Once again we employed self-priming RNA PETE constructs, and measured the rate of incorporation of the first few nucleotides, as previously described (Hobdey *et al.*; 2010). The two RNA constructs are both added to the same reaction along with ATP and GTP. The shorter of the two constructs allows for incorporation of only guanosine while the longer of the two allows incorporation of both adenine and guanosine (Fig. 2.4A). This allows us to measure the rate at which a +1 or +2 elongation complex is formed. Data show slight decreases in the formation rates for Y14A and Y157A, while Y118A and Y148A show ten and thirty fold decreases in formation rates respectively (Fig. 2.4E). This significant decrease partially explains the decreased amount of full elongation seen previously (Fig 2.3).

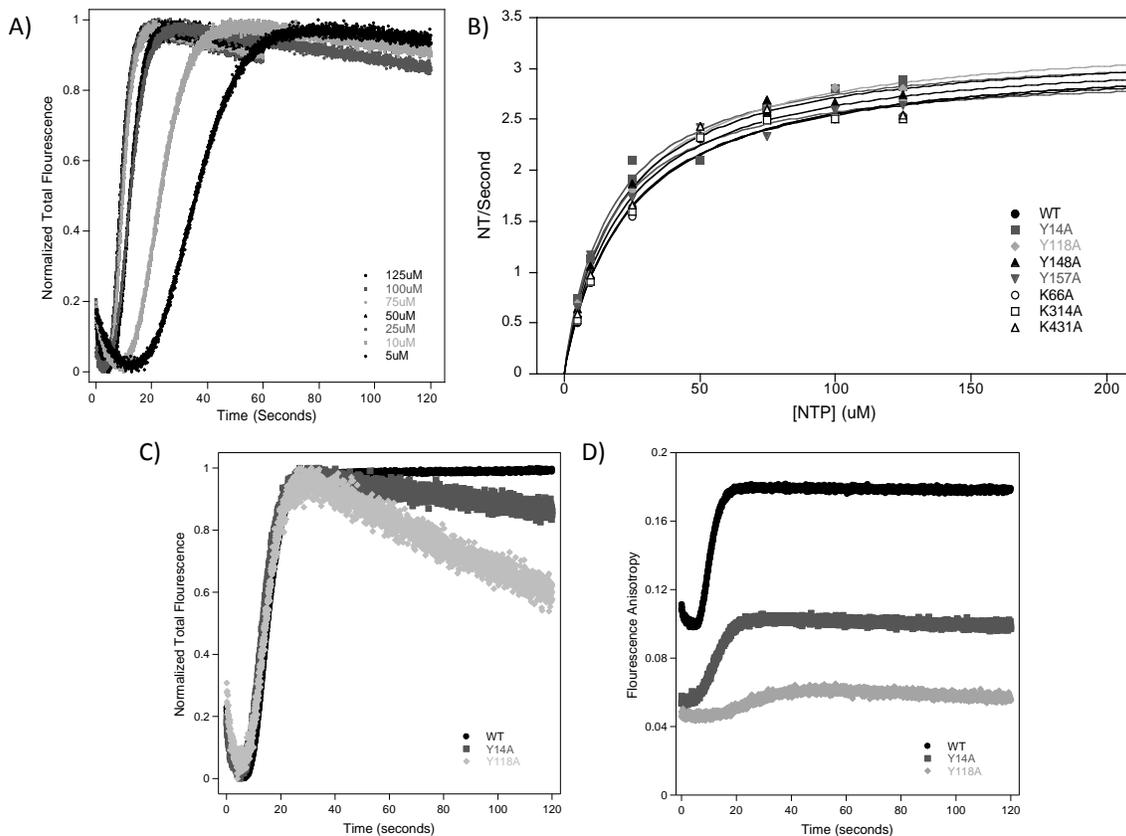
The mutations of basic residues also demonstrated a decrease in formation rate ranging from three to eight-fold (Fig. 2.4D). Since initiation as a whole entails RNA binding and incorporation of the first nucleotide, any effects on RNA binding will be apparent in this assay as well. Therefore, the diminished RNA affinities for the basic residues may provide an explanation for the reduced rates of EC formation. Indeed, the two values correlate very well, in that the mutants with higher disassociation constants also having slower initiation rates (Table 2.1). The control mutants K66A, K314A, and K431A displayed near WT activity for incorporation of the first one or two nucleotides (Fig. 2.4C).



**FIG 2.4 Elongation complex formation results.** (A) Cartoon of the polymerase elongation template element (PETE) RNA constructs used to determine the rate of elongation complex initiation. Nucleotide incorporation begins at the triangle and continues until reaching a templating GTP, six nucleotides from the end of the sequence. The asterisk marks the position the deoxythymidine residue is added to allow for imaging via an infrared fluorescence dye (B) Denaturing PAGE showing the formation of both the +1 and +2 products over time in minutes from the starting material (S) for both wild-type and an example of a slower mutant (Y148A). (C) Graph showing the formation of +2 complex over time. Comparison of the initiation rates for the +2 product between wild-type and the control mutants, (D) charge mutants, (E) and tyrosine mutants. The time constants associated with each curve-fit are given in Table 2.1.

### 2.3.3 Elongation Rate not affected by mutations.

To further dissect the effect of the tyrosine mutants, we performed experiments to determine the 3D<sup>pol</sup> elongation rate using the PETE assay in stopped-flow kinetics studies. The RNA construct used for the elongation rate experiments consists of a ten base pair hairpin for self-priming followed by a 26-nt templating region containing a 20-nt heterogeneous G-less cassette. The RNA is incubated with the first nucleotide for incorporation and then stalled at a +1 elongation complex. Stopped-flow experiments mix the complex with the remaining nucleotides for full elongation and measure the amount of time required by the 3D<sup>pol</sup> to reach the 5'-end of the RNA construct and restrict the mobility of the fluorescein. Data shows an initial lag phase corresponding to the incorporation of the first 20 nucleotides, comprising the G-less cassette, followed by a large increase in fluorescence polarization corresponding to the incorporation of the last five nucleotides (Gong *et al.*, 2009). This data enable us to determine a rate measured in nucleotides per second. Once carried out at multiple NTP concentrations (Fig. 2.5A), we can analyze the data according to Michaelis-Menton kinetics to determine  $V_{max}$  and apparent  $K_m$  values to compare WT 3D<sup>pol</sup> with the mutated polymerases. The data for wild-type polymerase matches previously published rate constants (Gong *et al.*, 2009). In addition, the tyrosine mutants and control mutants also displayed near wild-type values for both the  $V_{max}$  and apparent  $K_m$  (Fig. 2.5B). This indicates these residues have no direct effect on 3D<sup>pol</sup> elongation rates and confirms no direct effects on substrate affinity. Although the kinetics of the mutant polymerases matched those of wild-type 3D<sup>pol</sup>, several of the stopped-flow traces showed a decrease in total fluorescence after reaching the maximum level while wild-type and the control mutants displayed no such drop in total fluorescence (Fig. 2.5C). This phenomenon was evident in all four of the tyrosine mutants, but most profound in Y118A and Y148A.



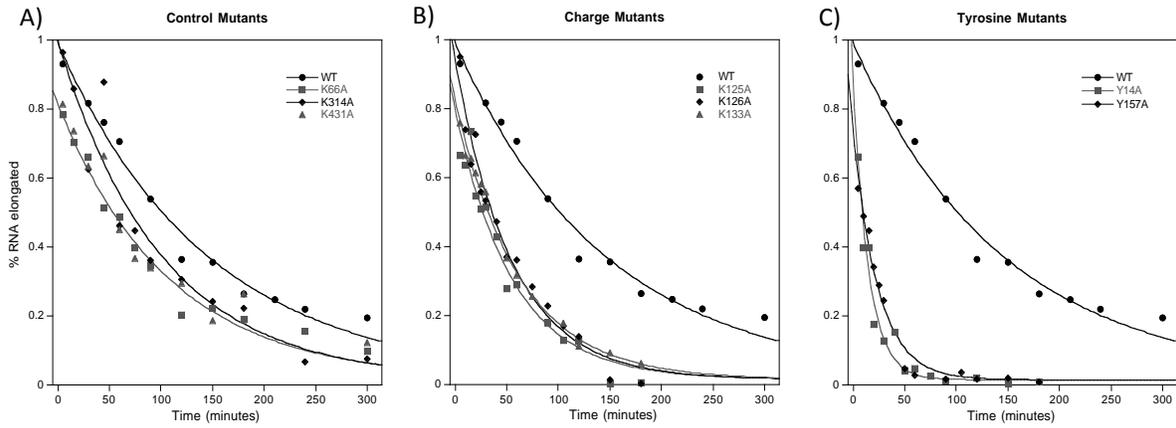
**FIG 2.5 Elongation rate experiments.** (A) Lag phase and the following increase in total fluorescence associated with the incorporation of the first twenty and last five nucleotides, respectively. The length of the lag phase decreases as the [NTP] increases. (B) Michaelis-Menton plot comparing the maximum elongation rate and apparent  $K_m$  of wild-type 3D<sup>pol</sup>, control mutants, and tyrosine mutants. The calculated values for these enzymatic parameters are listed in Table 2.1. (C) Normalized total fluorescence of wild-type, Y14A, and Y118A at 25uM NTP showing the differences in signal decay after reaching maximum total fluorescence. (D) Raw fluorescence anisotropy plots of wild-type, Y14A, and Y118A at 25uM NTP showing the differences in amplitude change and starting amplitude. The elongation curves starting with a lower initial fluorescence anisotropy also exhibit a smaller overall change in anisotropy, indicating that fewer complexes are capable of elongation.

### 2.3.4 Mutations affect EC stability

To determine whether these residues played a role in stabilizing the elongation complex, we designed an experiment in which the 9+2.24 RNA was stalled in the +2 complex. The elongation complex is then diluted 10-fold with a buffer containing 300 mM NaCl to prevent rebinding of the polymerase to any released RNA. The diluted complex is incubated at room temperature for various amounts of time and then provided with the remaining nucleotides to elongate to a +20 product. This allowed for measurement of the efficiency with which the pre-formed complex elongates to full-length product after a pause in the elongation cycle that correlates directly with the stability of the elongation complex. Data from these experiments for wild-type 3D<sup>pol</sup> matches previously published results (Hobdey *et al.*, 2010). In addition, the control mutants displayed comparable levels of stability to wild-type 3D<sup>pol</sup> (Fig. 2.6A); however, all of the remaining mutations led to significant decrease in the stability of the elongation complex. The charged mutants all displayed a roughly 3-fold decrease in stability and Y14A and Y157A showed an approximately 8-fold decrease (Fig. 2.6B&C). Mutations to residues 118 and 148 severely affected the stability of the complex so that no full-length product was seen after only 5 minutes of incubation.

## 2.4 Discussion

In analyzing the role of residue 5 in the 3D<sup>pol</sup> elongation complex, the Peersen lab proposed possible interactions between the downstream duplex RNA and Trp5 (Hobdey *et al.*, 2010). Although this interaction has not been captured through crystallography studies, the templating RNA could possibly wrap around the fingers domain of the polymerase before entering the active site. The 3D<sup>pol</sup> structure shows a putative channel that may provide the necessary



**FIG 2.6 Elongation complex stability results.** Percentage of RNA that forms full length product from a stalled elongation complex over time. (A) Comparison of the kinetic data for the stability of the +2 product over time between wild-type and the control mutants, (B) charge mutants, (C) and tyrosine mutants. The time constants associated with each curve-fit are given in Table 2.1.

interactions to facilitate wrapping of the RNA around the fingers domain. Along this channel are several basic residues that may form electrostatic interactions with the phosphate backbone of the incoming RNA. In addition, several tyrosine residues within this groove may stack with the bases of the single stranded RNA or facilitate other necessary interactions. Through a mutational study to select lysines, arginines, and tyrosines within this channel, we have assessed the importance of these residues in forming and stabilizing an elongation complex. This data is consistent with the hypothesis formulated from the Trp5 study that the downstream template RNA wraps around the polymerase and makes interactions with the fingers domain that serve in the preservation of the elongation complex.

In characterizing the role of the residues along this putative RNA entry channel, we first tested whether elimination of the basic residues or large planar groups of the tyrosine residues affected RNA binding. The results show the tyrosine residues within the fingers domain have no direct effect on RNA affinity. In contrast, the mutations of the charged residues located within the pinky finger of 3D<sup>pol</sup> diminished the affinity between protein and RNA. Since substitution of alanines for the tyrosine residues within the groove show no significant effects on binding, we conclude that even if the template RNA is interacting with these residues, that the interactions do not facilitate the initial binding event. On the other hand, since mutations to the basic residues within this channel did affect binding, we assume removal of the positively charged side chain disrupts electrostatic interactions between the pinky finger of 3D<sup>pol</sup> and the negatively charged phosphate backbone of the incoming RNA. This becomes especially evident from the elongation complex structure, which shows residue R188 clearly forming a salt bridge with the -2 phosphate of the upstream duplex. Other residues, such as K133, make clear interactions with the phosphate backbone of the downstream template. Although the elongation complex structure does not show

K125 and K126 making direct interactions with downstream RNA, they may serve a role in a more general electrostatic contribution to binding affinity.

While the tyrosine residues lining this possible template RNA binding channel do not play a direct role in binding RNA, the results show that these residues, especially Y118A and Y148A, play a key role in forming and maintaining a stable elongation complex. By analyzing the initiation rate associated with forming either a +1 or +2 elongation complex using PETE RNA constructs, we determined the effects of the mutations on initiation. The drastic decrease in elongation complex formation rate from Y118A and Y148A indicates these residues clearly have some function in initiation. Although the data does not provide an answer for the direct role of these two tyrosines, we propose these residues affect either an important conformational change or interact directly with the RNA in a manner not visible in the elongation complex structure. Both propositions may be associated with a possible *cis-trans* isomerization event associated with proline 119 (Andreotti, 2003). Such an event may serve to reposition Tyr118 for interaction with the templating RNA. In addition, these residues may simply have packaging properties that stabilize a local structural rearrangement from the isomerization event. Regarding Tyr157, the EC structure shows a stacking interaction with the ribose of the +2 nucleotide of the downstream RNA. Thus, we believe position 157 requires a planar residue that can accommodate this interaction. Finally, from the data and the associated structure, we cannot determine a specific role for Tyr14.

In testing initiation rates, we found the basic residues also displayed a decrease in elongation complex formation; however, this may be a result of these mutants' lowered binding affinity. Evidence for this comes from a proportional relationship between RNA affinity and formation rate. The mutants such as K126A and K133A have the highest affinity for RNA and are the

fastest charged mutants to form the stable elongation complex. In contrast, R188A is incapable of forming either the +1 or +2 product, but has an RNA dissociation constant of 170  $\mu\text{M}$ , suggesting the inability to form an elongation complex arises from an inability to bind RNA since the binding event occurs first in the overall process of initiation.

Data from the elongation rate experiments indicates the hydrophobic residues do not affect RNA elongation rates due to the wild-type values these mutants have for both  $V_{\text{max}}$  and  $K_m$ . Interestingly, the data from the stopped-flow kinetics does show some discrepancy between the tyrosine mutant and wild-type plots. Not only do we see a decrease in signal amplitude for some of the mutants, there is also a decrease in total fluorescence (TF) after the maximum TF is reached in the mutant traces, while the WT and control signals remain steady. These data suggest a possible decrease in stability of the elongation complex associated with the mutations. The smaller change in fluorescence anisotropy indicates fewer complexes in solution, and the decrease in TF, implies the complexes are disassociating after elongation. To test this, we expanded our initiation assay to enable us to evaluate EC stability and determine a half-life for the complex. From the data, Tyr118 and Tyr148 once again emerge as key residues for formation and maintenance of the stable EC. Mutations at either of these positions eliminates 3D<sup>pol</sup>'s ability to fully elongate from a stalled position after only five minutes. While Y118A and Y148A stand out, all of the mutations, both charge and tyrosine mutants, showed significant decreases in stability. Since tampering with the structure of the fingers domain has definite implications on stability, we believe this domain plays a key role in maintaining the EC in the correct conformation for processive elongation. Although we lack structural insight into the exact role the fingers domain plays in stabilizing the polymerase during replication, the lack of conformational changes in this region between the original 3D<sup>pol</sup> structure and the EC structure

indicates there may be an unknown conformational change that repackages this region. Such a structural rearrangement seems even more plausible when considering the initiation and stability data from this study.

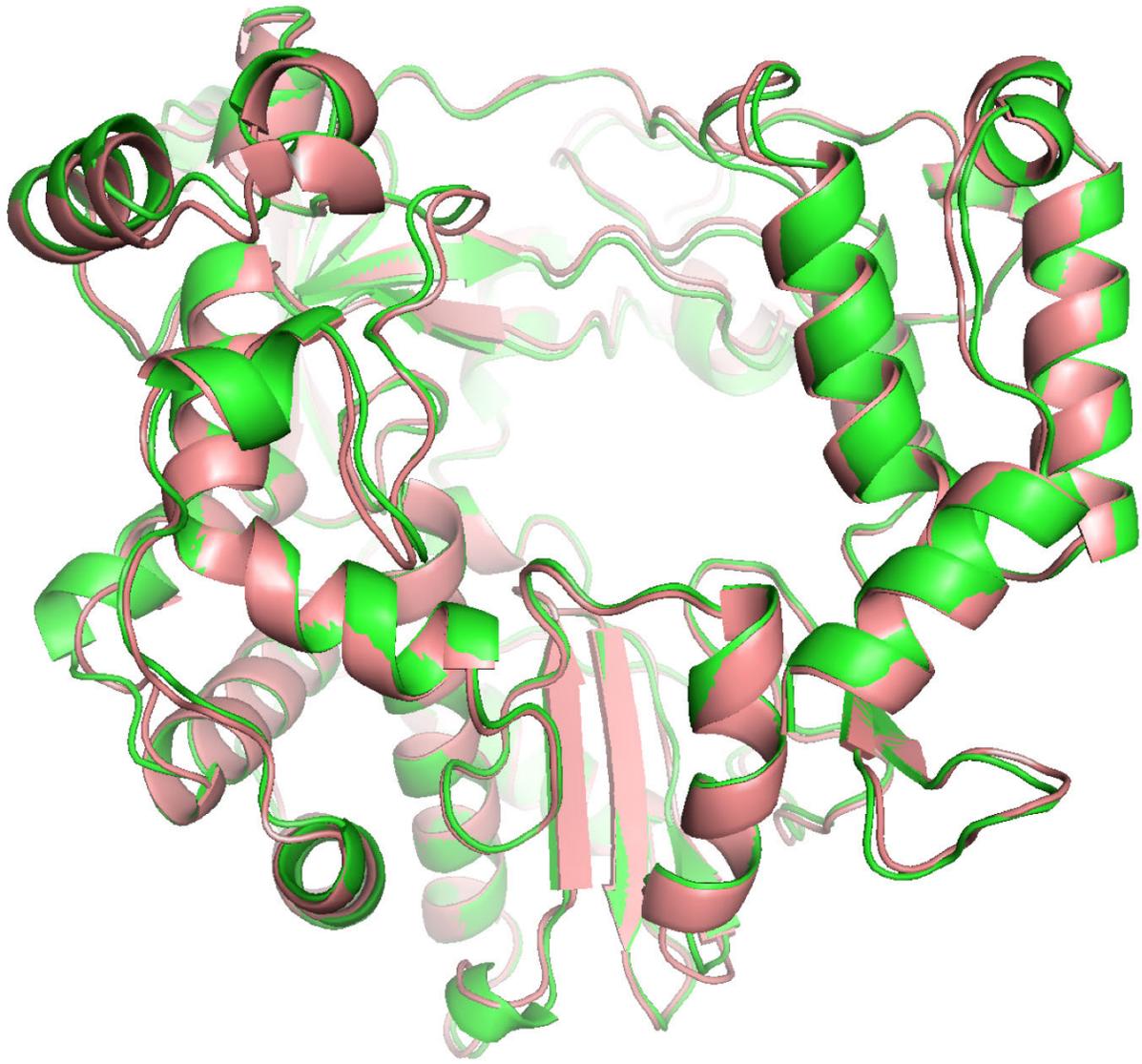
In conclusion, we have provided additional evidence supporting the hypothesis that the downstream template RNA interacts with the fingers domain before undergoing strand separation. Specifically, it appears several basic residues lining a channel that separates the pinky finger from the remaining fingers perform a role in binding the RNA. In addition, Tyr118 and Tyr148 have a critical role in forming and maintaining a stable elongation complex. Although, we currently have no structural evidence that provides insight regarding the exact role of these residues, we believe there is an additional conformation the polymerase can adopt that we have not captured. In this conformation, these residues may directly interact with the template RNA or have packaging properties.

## Chapter 3

### Structure of a Coxsackievirus Polymerase Elongation Complex

#### 3.1 Introduction

The structures of several viral polymerases have been solved, including those from poliovirus (PV) (Thompson and Peersen, 2004) and coxsackievirus B3 (CV) (Campagnola *et al.*, 2008; Gruez *et al.*, 2008). Phylogenetic analysis shows these two viruses are genetically closely related and, therefore, have highly homologous polymerases that differ in length by a single amino acid insertion in coxsackievirus. Indeed the two structures solved by the Peersen lab superimpose almost exactly and show no key differences in the overall structure (Fig. 3.1). Both polymerases fold according to the right hand analogy that includes palm, thumb, and fingers domains (Thompson and Peersen, 2004). The tip of the fingers domain reaches over the palm to make contacts with the tip of the thumb domain through a set of hydrophobic interactions. This interaction encloses the active site centered within the palm domain and creates a channel through which NTPs can be shuttled in for catalysis (Thompson and Peersen, 2004). In addition to the original PV 3D<sup>pol</sup> and CV 3D<sup>pol</sup> structures, the Peersen lab has recently solved the structure of the PV elongation complex (PVEC) (Gong and Peersen, 2010). The lab has developed an experimental system for assembling and purifying PV polymerase elongation complexes stalled after several rounds of nucleotide incorporation. This results in a highly stable complex paused in the replication cycle and poised for incorporation of the next nucleotide. After obtaining crystals of PVEC, soaking experiments with NTPs and various NTP analogs are performed to trap the polymerase in

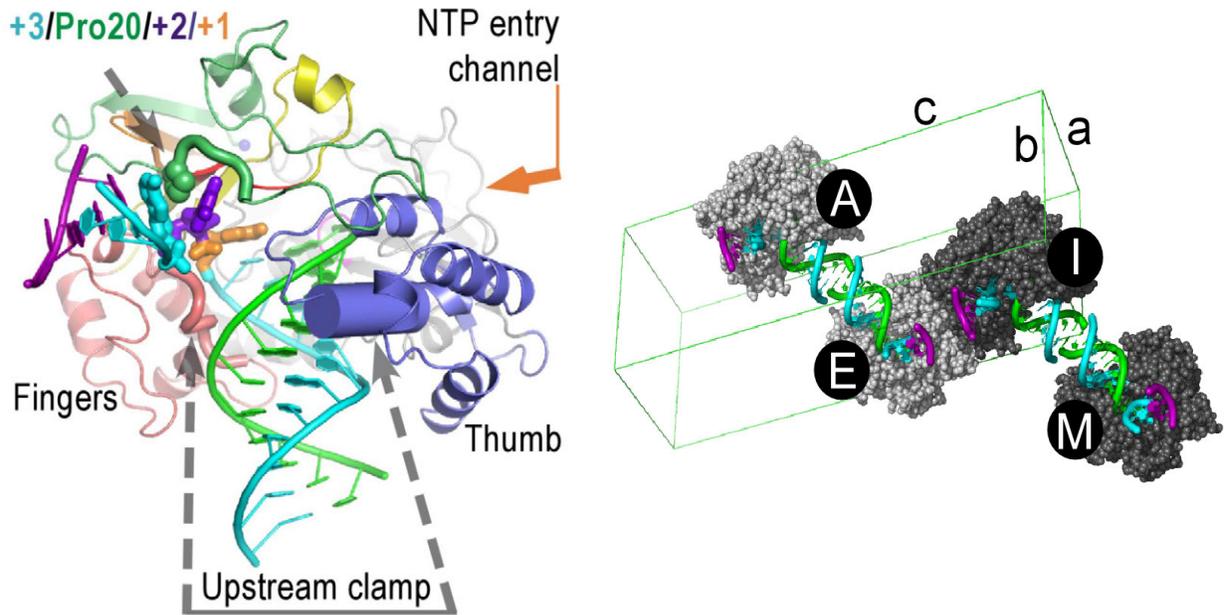


**FIG 3.1 Comparison of picornaviral polymerase structures.** Superimpositions of poliovirus (pink) and coxsackievirus (green) apo polymerase structures to show the high similarity between the viral polymerases.

different steps of the catalytic cycle. Using this approach, the lab has solved several structures that capture PV 3D<sup>pol</sup> in different stages of replication.

The PVEC structure shows template nucleotides +1, +2 and +3 are single stranded with strand separation of the downstream duplex RNA occurring at residues 18 and 19 of the index finger (Fig. 3.2) (Gong and Peersen, 2010). The +2 nucleotide resides in a pocket formed by Pro 20 and the index finger that prevents it from base stacking with either the +3 or +1 nucleotides. The +1 nucleotide is positioned within the active site to allow for base stacking with the -1 upstream nucleotide of the priming base pair and base pairing with the incoming nucleotide triphosphate. The product RNA exits as an upstream duplex that is locked in place by a clamp formed by the pinky and thumb domains of the polymerase. Analysis of the structures obtained from the soaking experiments showed an absence of a nucleotide-repositioning step, common in other polymerases. Rather, NTPs bind to an *open* conformation of the polymerase that *closes* for the subsequent catalysis via a subtle repositioning at the palm domain. This conformational change causes a shift in the palm domain that repositions Asp 233. This differs from other polymerase structures, which utilize large movements in the fingers domain (Gong and Peersen, 2010).

The experimental approach used to obtain crystals employs RNA constructs comprised of a 10mer primer annealed to a 35mer template strand with four nucleotides incorporated, resulting in a complex containing both upstream and downstream RNA duplexes (Fig. 3.2A) (Gong and Peersen, 2010). The upstream product strands coaxially stack with each other and form a set of crystal packing interactions (Fig. 3.2B). The RNA constructs all contain a pair of helical segments with an intervening templating sequence.



**FIG 3.2 Poliovirus 3D<sup>pol</sup>-RNA elongation complex structure.** (A) The poliovirus polymerase is colored with the palm in grey and the thumb in blue while the index, middle, ring and pinky fingers are in green, orange, yellow and pink, respectively. The structure shows the template RNA strand in cyan, downstream non-template strand in magenta and upstream non-template in bright green. The single stranded template nucleotides +1, +2 and +3 are colored orange, purple and cyan, respectively. The structure also highlights proline 20 in green and the upstream clamp that forms around the exiting upstream duplex with arrows pointing to regions of the pinky and thumb that interact with the RNA (from Gong and Peersen 2010). (B) The figure shows the staggered coaxial stacking interactions between the upstream template-product duplexes of two different elongation complexes (A+E and I+M) in context of the unit cell with dimensions a, b, and c. Again the template strand is in cyan, the product strand in bright green, and the downstream non-template in purple (from Gong and Peersen, 2010).

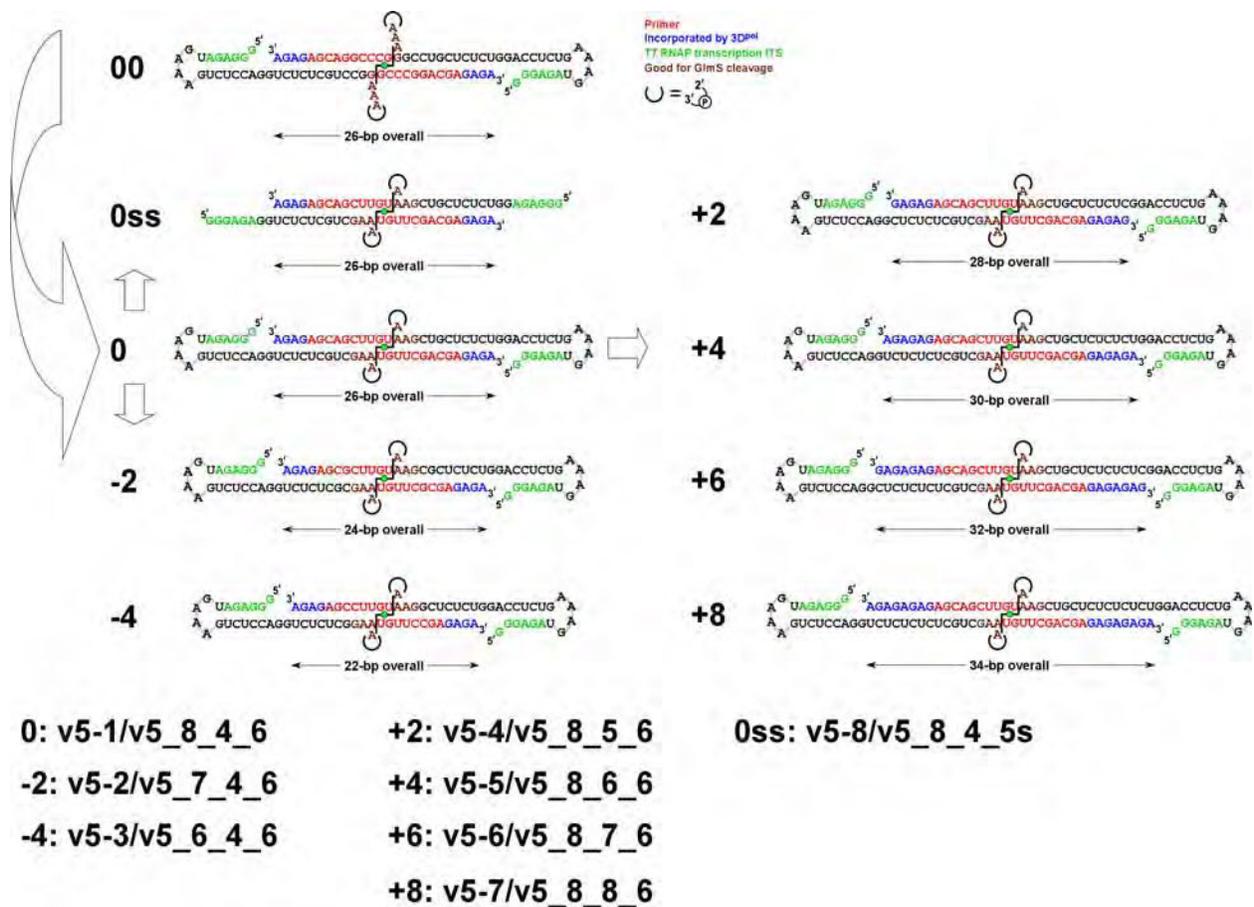
By varying the length of the coaxially stacked product and non-stacked template regions, the complexes may stack differently with one another and crystallize in different forms (Gong and Peersen, 2010).

In this work, we utilize the experimental approach developed by Gong and Peersen to crystallize a stalled coxsackievirus polymerase elongation complex (CVEC). Before attempting to crystallize the CVEC, we employed the kinetic assays from the PV mutagenic studies (Chapter 2) to establish that CV could efficiently form and maintain a stable elongation complex. To produce diffracting crystals, we synthesized variations of the RNA used to assemble the elongation complex that produced the PVEC structure by inserting or deleting sequence surrounding the core RNA construct (Fig 3.3). The ensuing crystallographic work produced four complete data sets with resolution at 2.75Å or better in two different crystal forms (C2 and P1). The CVEC structure from the crystal in the C2 space group closely resembles the PVEC structure, showing few differences. However, the crystal in the P1 space group produced a CV elongation complex trapped in intermediate state of translocation. This P1 CVEC structure displays key differences from the PVEC structure and the C2 CVEC structure, thereby providing insight to mechanism of nucleotide translocation in the replication cycle.

## **3.2 Material and Methods**

### *3.2.1 Protein Expression and Purification*

For this work, the CV 3D<sup>pol</sup> gene was inserted into the pET26-UbDH plasmid used for making mutations to basic and tyrosine residues in PV 3D<sup>pol</sup> and described in section 2.2.1. The resulting plasmid contains the complete open reading frame of the coxsackievirus 3D<sup>pol</sup> gene sandwiched between an N-terminal ubiquitin (Ub) and a



**FIG 3.3 RNA constructs for EC assembly.** The v4-1 RNA construct (00) that produced the original PVEC structure was modified to produce eight new constructs all containing the same core sequence with nucleotides added or removed from the template region to vary the overall length. Both the v4-1 and v5-1 constructs contain an 8-bp upstream duplex, allow for incorporation of 4-nt and form a 6-bp downstream duplex (8-4-6). To make variations of v5-1, nucleotides can be removed from the upstream duplex (v5-2 and v5-3) or added to the template region allowing for incorporation of additional nucleotides (v5-4 through v5-7). The v5-8 construct resembles the v5-1, but lacks downstream duplex (Figure from Peng Gong).

C-terminal GSSS-His<sub>6</sub> tag (Gohara *et al.*, 1999). Using *Escherichia coli* BL-21 DE3 pCGI cells, CV3D<sup>pol</sup> was expressed as described in section 2.2.1. After obtaining pellets from expression, we followed a purification protocol utilizing generally higher salt concentrations in later steps for CV 3D<sup>pol</sup> due to solubility concerns. Pellets were resuspended in lysis buffering consisting of 50 mM Tris (pH 8.0), 400 mM NaCl, 20% (vol/vol) glycerol, and 0.02% (wt/vol) NaN<sub>3</sub> and lysed at 18,000 lbs/cm<sup>2</sup> in a model M-110L microfluidizer (Microfluidics). NP-40 was added to the lysate in a drop-wise manner to a final concentration of 0.1% (vol/vol) followed by DNA precipitation through the addition of PEI to a final concentration of 0.25% (vol/vol). The solution was centrifuged at 17,000 rpm for 45 minutes in a Sorvall RC2-B centrifuge. The supernatant was filtered through a 0.2 micron cellulose acetate filter (Amicon), loaded onto a HisTrap HP nickel column (GE Healthcare), and eluted with 350 mM imidazole in buffer containing 50 mM Tris (pH 8.0), 400 mM NaCl, 20% (vol/vol) glycerol, and 0.02% (wt/vol) NaN<sub>3</sub>. Fractions containing the polymerase were pooled and diluted four-fold in low salt buffer consisting of 50 mM NaCl, 25 mM Tris (pH 8.5), 20% (vol/vol) glycerol, 0.02 NaN<sub>3</sub> to allow for loading onto a HiTrap Q HP column (GE Healthcare). Polymerase was eluted as 2 mL fractions via a linear gradient of 50 mM to 1 M NaCl into tubes containing 50 uL of 5 M NaCl and pooled. The resulting sample was diluted in low salt buffer to bring the [NaCl] below 125 mM and loaded onto a Mono Q column (GE Healthcare). Elution into 1 mL fractions was performed with a linear gradient using the same buffer conditions as for the Q-column into tubes containing 25 uL of 5 M NaCl bringing the salt concentration to approximately 400 mM. Fractions containing the polymerase were pooled, concentrated to a volume of ~0.8 mL, and loaded on to a

Superdex 200 gel filtration column (GE Healthcare) equilibrated in 300 mM NaCl, 5 mM Tris (pH 7.5), 20% glycerol (vol/vol), and 0.02% (wt/vol) NaN<sub>3</sub>. Fractions containing the polymerase were pooled, provided 5 mM TCEP, concentrated to a final [3D<sup>pol</sup>] of 200 μM, flash frozen with liquid nitrogen, and stored at -80°C. Purification efficiency and sample purity were assessed by running samples from each step of the process on 12% acrylamide PAGE.

### 3.2.2 RNA Synthesis and Preparation

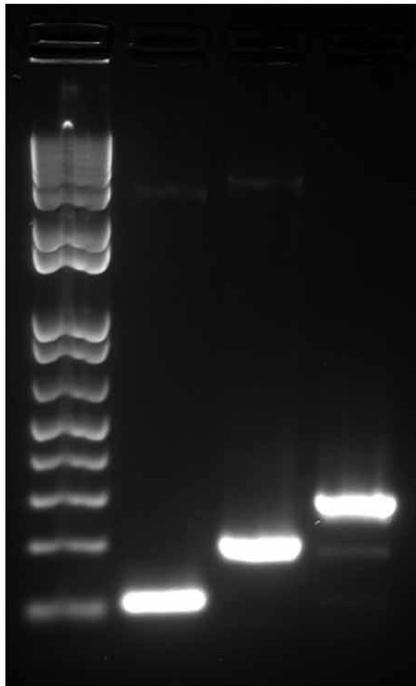
The RNA constructs were made through a series of PCR reactions to produce a DNA template used in a large-scale *in vitro* T7 transcription as previously described (Kieft and Batey, 2004; Batey and Kieft, 2007; Gong and Peersen, 2010). The primers used in the series of PCR reactions were synthesized by IDT, purified by 12% acrylamide 7M urea gel electrophoresis, excised from the gel, and isolated with an Elu-Trap electro-elution device (Whatman Inc.). Primers were designed to anneal to each through an 18-25 nucleotide overlap in the first PCR reaction, generating a product containing the sequence of interest (Fig 3.4A). In a simultaneous reaction using a parental plasmid pRAV23 provided by the Kieft lab at the University of Colorado, a PCR reaction produces the glmS ribozyme (Fig 3.4A). Products from the two PCR reactions were combined in a third PCR reaction producing a DNA template containing a T7 promoter followed by the RNA sequence, cleavage site, and glmS ribozyme (Fig 3.4A&B).

To generate full-length RNA, the PCR product was incubated for 3 hours in an 8 mL T7 RNA polymerase reaction containing 60 mM Hepes (pH 7.8), 2 mM spermidine, 0.01% (vol/vol) Triton X-100, 15 mM DTT, 22 mM MgCl<sub>2</sub>, 4 mM each rNTP, and 15 μg/mL T7 RNA polymerase. After removal of the pyrophosphate

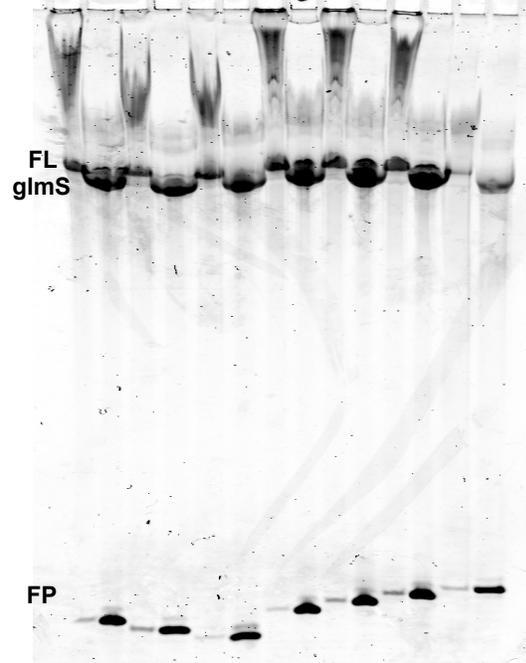
**A)** first PCR step:



**B)** MW 1a 1b 2



**C)** v5-1 v5-2 v5-3 v5-4 v5-5 v5-6 v5-7  
U C U C U C U C U C U C



**FIG 3.4 Production of RNA constructs for EC assembly.** A) Schematic of the three PCR reactions used to generate DNA template for T& transcription. Reaction 1a utilizes two primers synthesized by IDT that contain an 18-25 nucleotide overlap to produce the RNA sequence of interest. Reaction 1b utilizes pRAV23 as a template to produce the glmS ribozyme. Reaction 2 fuses the two first round products producing the template for transcription. B) 1% agarose gel with ethidium bromide staining showing the relative lengths of the three PCR products. C) 12% acrylamide 7 M urea gel with pre- and post-cleavage samples for each of the duplex RNA constructs. The gel shows uncut (U) samples containing full length RNA (FL) and cut (C) samples with the glmS ribozyme separated from the final product (FP).

precipitant through centrifugation, the transcription reaction was incubated at 65°C for three minutes and slowly cooled to room temperature to fold the glmS ribozyme into the catalytically active form. Addition of glucosamine-6-phosphate and NaCl to final concentrations of 5 mM and 100 mM, respectively, followed by a 30 minute incubation at 42°C constitutes the glmS ribozyme self-cleavage reaction, generating the RNA product (Fig 3.4C). To purify the RNA, the cleavage reaction was run on a 12% acrylamide 7 M urea gel, excised from the gel, and isolated with an Elu-Trap electro-elution device (Whatman Inc.). Purified RNA undergoes a self-annealing reaction driven by a 5 minute incubation at 95°C followed by a snap-cooling event on ice to fold the RNA in the appropriate form. The final RNA product was stored at -80°C in RNA annealing buffer (RAB) composed of 5 mM Tris (pH 7.5), 50 mM NaCl, and 5 mM MgCl<sub>2</sub>. Using this approach, nine different RNA constructs were assembled with sequence variation outlined in Figure 3.4.

### *3.2.3 EC Assembly, Purification, and Crystallization*

To prepare the RNA for elongation complex assembly, a 10-nt RNA primer was added to product RNA at 20% molar excess and annealed during a 5 minute incubation at 42°C in RAB. The assembled RNA was mixed with purified 3D<sup>pol</sup> to final concentrations of 15 μM and 20 μM, respectively, in a reaction buffer containing 50 mM Hepes (pH 6.5), 75 mM NaCl, 5 mM MgCl<sub>2</sub>, 2 mM TCEP, and 450 μM of each ATP and GTP. The 2 mL reaction was incubated at 22.5°C for 20 minutes and then purified over a MonoQ column (GE Healthcare) using a linear salt gradient from 75 mM to 1 M NaCl in buffer containing 50 mM Hepes (pH 6.5), 5 mM MgCl<sub>2</sub>, 2 mM TCEP, and 0.02% NaN<sub>3</sub>. The purified elongation complex was buffer exchanged into complex buffer (20 mM Hepes

(pH 6.5), 100 mM NaCl, 2 mM MgCl<sub>2</sub>, 2 mM TCEP) and concentrated to 8 mg/mL through centrifugation.

Crystallization conditions were screened using nine different RNA constructs generated by varying primer sequences in the PCR reactions described in section 3.2.2 and the Cryos Suite screening kit (Qiagen). Crystal trays were set up using a Douglas Instruments Oryx-8 robot to produce a 0.8 uL drops that were incubated at 16°C to allow for crystal growth by sitting drop vapor diffusion. Successful crystals were harvested and frozen in mother liquor with liquid nitrogen.

#### *3.2.4 Diffraction Data Collection and Reduction*

Crystals were sent to the Advanced Light Source in Berkeley, CA. Diffraction data was collected using the Molecular Biology Consortium synchrotron beam line 4.2.2, which was fully operated by remote control from Colorado. To ensure a complete dataset, 187° of data were collected in 0.5° increments. The diffraction data was integrated, merged, and scaled using the d\*Trek software suite. The output files (.ref) are converted to the appropriate form (.mtz) to obtain electron density maps using the CCP4 suite of programs. To obtain a model for molecular replacement, we superimposed the apo-CV3D (pdb file 3DDK) with the PVEC structure (pdb file 3OL6) and deleted both the PV polymerase and the portions of the upstream and downstream RNA that did not match the RNA construct that produced CVEC crystals. Superimposition was centered on residues 320-335 of the PV3D<sup>pol</sup> sequence (residues 321-336 of CV3D<sup>pol</sup>). Using Phenix, we performed molecular replacement to obtain an initial model and electron density map.

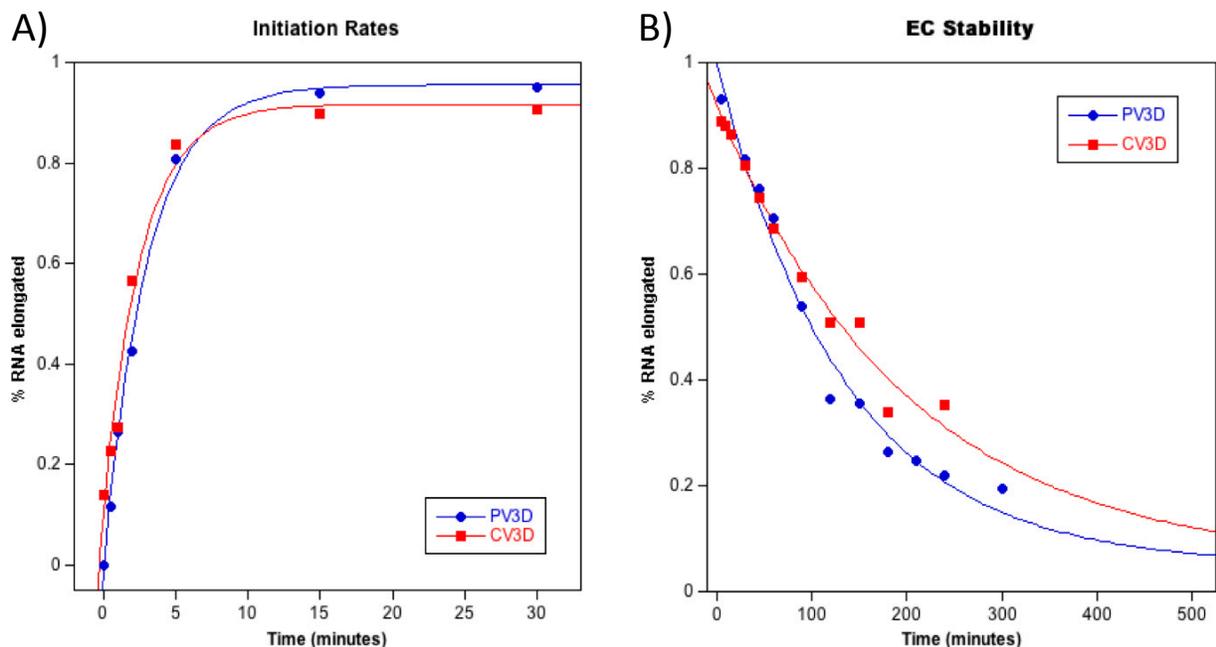
## 3.3 Results

### 3.3.1 CV3D Kinetic Assays

To ensure coxsackievirus would also efficiently form a stable elongation complex that will remain active for several hours, we employed the same assays used in the PV3D mutagenic studies (Section 2.2.5). Using both the 10+1.12 and 9+2.24 RNA constructs, we found the CV3D formed full-length product at approximately the same rate as the PV polymerase (Fig. 3.5A and Table 3.1). In addition, the CVEC demonstrates comparable levels of stability, evident from the continued elongation to full-length product of the 9+2.24 after extended amounts of time (Fig. 3.5B and Table 3.1). In fact, the CVEC appears to be slightly faster in forming the +2 complex and more stable upon the formation of the +2 complex than the corresponding PVEC.

### 3.3.2 CV 3D<sup>pol</sup> Elongation Complex Structure

By screening nine different RNA constructs with a cryos suite screening kit from Qiagen, we obtained and harvested approximately 50 crystals from multiple conditions (Fig. 3.6). Of the crystals, four (mk022, mk040, mk042, mk050) gave complete high-resolution data sets that yielded structures (Table 3.2). All four crystals came from elongation complex containing the v5-1 RNA construct (Fig. 3.3) but produced two different crystal forms (C2 and P1), from two different conditions (Table 3.2). The two structures in the C2 space group match the original PVEC structure in the catalytically open conformation, showing the polymerase trapped after incorporating four nucleotides. The fingers and thumb domains form a clamp around the upstream duplex as it exits the active site. To accommodate the product RNA, both domains must move outward as compared to the apo polymerase structure (Fig. 3.7). In addition, template binding forces



**FIG 3.5 Comparison of PV3D<sup>pol</sup> and CV3D<sup>pol</sup> kinetics.** A) Initiation rate. Percentage of RNA forming a +2 elongation complex over time using the 9+2<sub>-24</sub> (Data for +1 formation not shown). CV3D<sup>pol</sup> (red) initiates slightly faster than PV3D<sup>pol</sup> (blue). B) EC Stability. Graph shows the percentage of RNA that forms full length product from a stalled elongation complex over time. CV3D<sup>pol</sup> (red) remains stable for longer than PV3D<sup>pol</sup> (blue). Both graphs indicate both polymerase are highly capable of forming and maintaining a stable elongation complex.

**Table 3.1 Comparison of PV3D<sup>pol</sup> and CV3D<sup>pol</sup> Kinetics.** Values for all three columns correspond to time in minutes. The numerical values are obtained from the graphs above (Initiation rate for the +1 EC is not shown).

	Initiation (+1 EC)	Initiation (+2 EC)	EC Stability
PV3D	4.8 ± 0.3	3.0 ± 0.2	133 ± 8
CV3D	2.2 ± 0.4	2.6 ± 0.4	200 ± 10



**FIG 3.6 Coxsackievirus Elongation Complex Crystals.** Crystals produced from elongation complex containing the v5-1 RNA construct. Shown here are just a few of the approximately 50 crystals harvested throughout the screening of all nine RNA constructs.

**Table 3.2 Crystallographic data and refinement statistics.** Crystals are numbered in the order they are harvested. Data in parenthesis corresponds to highest resolution shell. No structural refinement data available for mk042 and mk050 because they have not yet been processed.

Crystal	mk022	mk040	mk042	mk050
Crystal Conditions	0.17 M NaOAc 0.085 M Tris pH 8.5 25.5% PEG 4000 15% glycerol	0.17 M NaOAc 0.085 M Tris pH 8.5 25.5% PEG 4000 15% glycerol	0.085 M NaOAc pH 4.6 1.7 M NaCl 15% glycerol	0.085 M NaOAc pH 5.1 1.6 M NaCl 15% Glycerol
Spacegroup	C2	C2	P1	P1
Unit Cell				
a,b,c (Å)	94.63, 77.08, 195.26	60.70, 60.68, 195.18	94.86, 76.88, 194.67	60.91, 60.87, 194.50
$\alpha,\beta,\gamma$ (°)	90.00, 90.02, 90.00	90.03, 90.03, 100.30	90.00, 89.73, 90.00	89.93, 89.76, 78.15
<b>Data Reduction</b>				
Resolution (Å)	48.82- <b>2.23</b> (2.31-2.23)	45.32- <b>2.73</b> (2.83-2.73)	46.03- <b>2.28</b> (2.36-2.28)	45.90- <b>2.44</b> (2.53-2.44)
Completeness (%)	99.9 (99.9)	97.9 (97.4)	99.8 (99.7)	97.8 (97.3)
R-merge	0.088 (0.488)	0.098 (0.302)	0.096 (0.382)	0.067 (0.332)
R-meas	0.102 (0.570)	0.138 (0.427)	0.112 (0.443)	0.094 (0.466)
Redundancy	3.81 (3.66)	1.90 (1.90)	3.84 (3.85)	2.00 (2.01)
I/Sigma	6.7 (1.5)	4.9 (1.5)	6.9 (1.6)	6.0 (1.5)
Mosaicity	0.32	0.92	0.39	0.55
<b>Structure Refinement</b>				
R-factor (%)	23.81		20.94	
R-free (5% of data)(%)	28.88		23.96	
Rmsd bond length (°)	0.008		0.011	
Rmsd bond angle (°)	1.056		1.154	

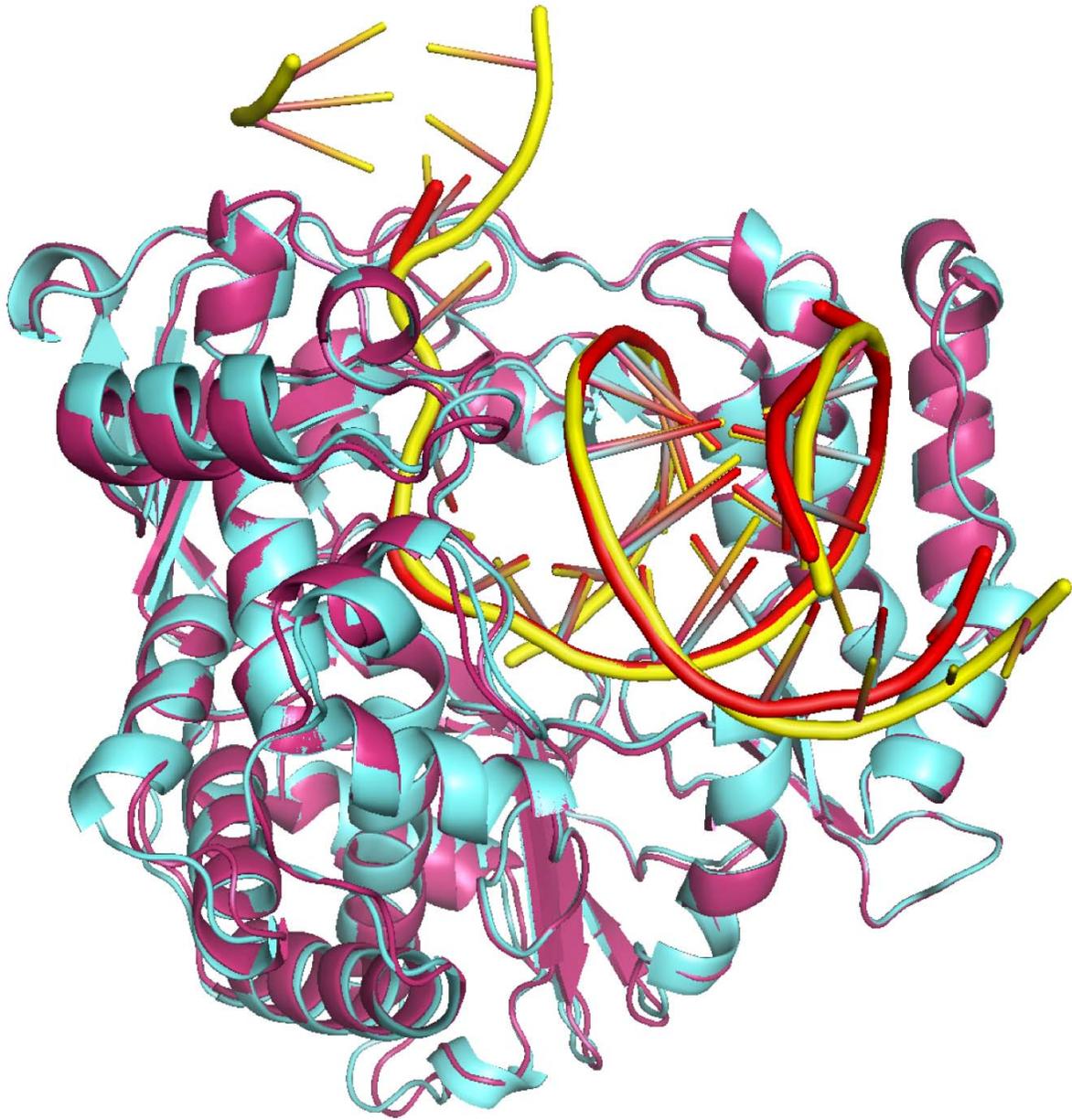


**FIG 3.7 Comparison of the CV 3D<sup>pol</sup> apo structure and elongation complex.** The apo CV3D<sup>pol</sup> structure (green) superimposed with the elongation complex structure (blue). Three main structural changes occur including outward movements of the fingers and thumb domain to accommodate RNA substrate. Also, Gly211 moves downward to prevent steric clash with the template strand.

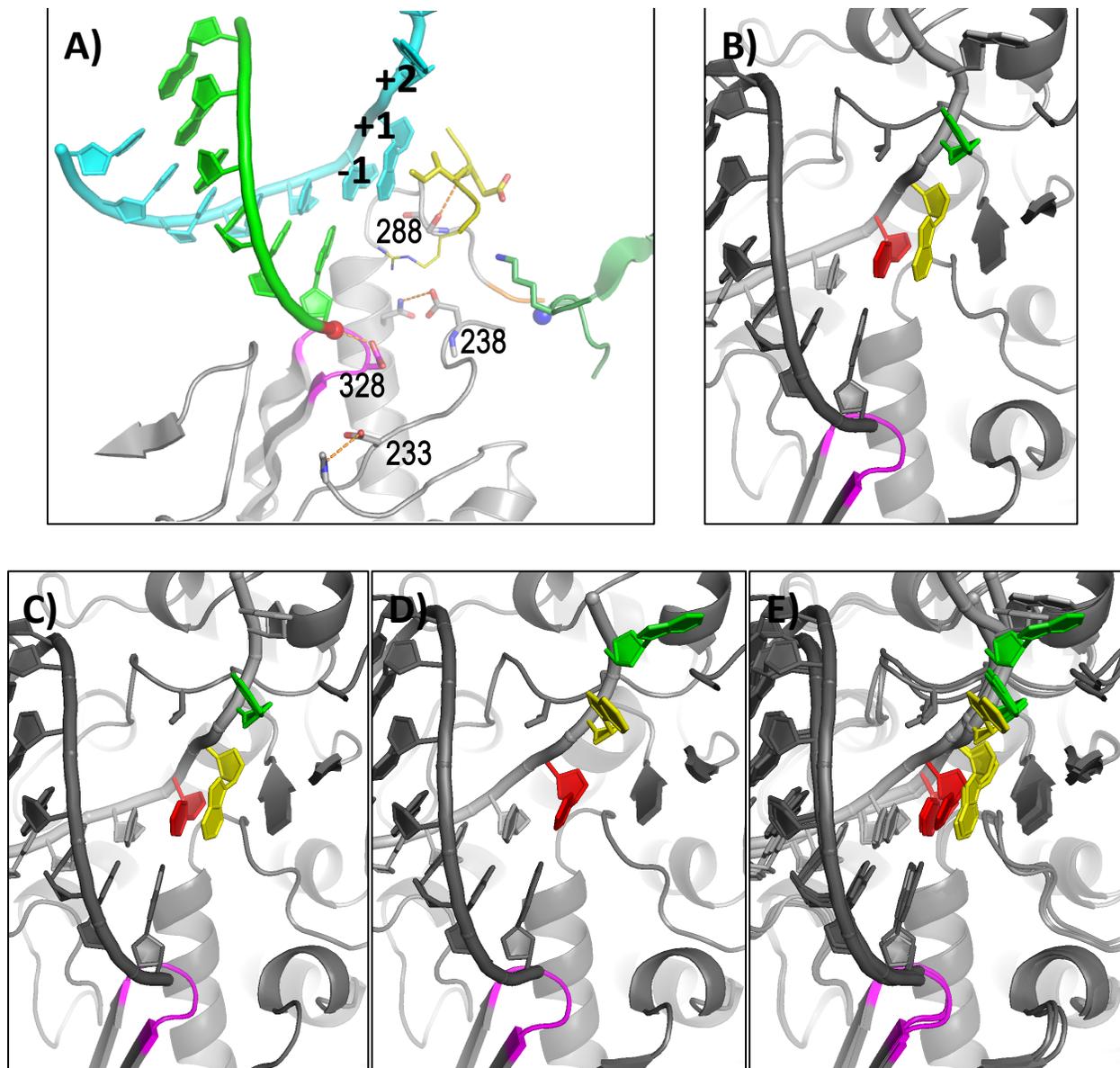
Gly211 down towards the palm to make room for the RNA and prevent any steric clashes. The structure also shows downstream RNA feeding into the active site; however, the weak electron density prevented modeling of the downstream duplex. Therefore, we see only three unpaired template nucleotides entering the active site at the same location shown in the PVEC structure. Strand separation occurs at a conserved cleft comprised of residues 18 and 19 of the fingers domain. The +3 and +2 bases have Pro20 wedged between them, preventing stacking between the bases and the +1 base stacks with the template product duplex, as observed in the PVEC (Fig 3.8).

### *3.3.3 CVEC Translocation Intermediate*

Along with the CVEC structure solved in the C2 space group, we obtained two crystals that captured the CVEC in a possible translocation intermediate. The structure, solved in the P1 space group, shows the template RNA shifted backwards from the active site relative to the PVEC structure and the C2 CVEC structure (Fig. 3.9). The template RNA in the PVEC structure (Fig. 3.9A&B) and the C2 CVEC structure (Fig. 3.9C) superimpose almost perfectly, with base stacking interactions occurring between the -1 and +1 bases and the +2 nucleotide shifted up into a small pocket. The P1 CVEC structure shows the -1 base still base pairing with the newly incorporated base of the product strand; however, the +1 base has shifted away from the active site (Fig. 3.9D). This backsliding of the template RNA prevents the stacking interaction between the -1 and +1 bases and places the +1 base in the pocket that is normally occupied by the +2 base (Fig. 3.9E). Furthermore, the RNA backsliding displaces the +2 base further from the active site so that it resides in essentially the +3 position of the PVEC and C2 CVEC structures (Fig. 3.9E).



**FIG 3.8 Comparison of the CVEC and PVEC structures.** The CVEC structure (cyan) superimposed with the PVEC (purple) with RNA in yellow and red, respectively. The structures show only minor differences in both the protein and RNA. The PVEC structure shows the downstream duplex (top) that does not appear in the CVEC due to weak density in that region.



**FIG 3.9 RNA backsliding in the CV Translocation Intermediate.** A) Active site of the poliovirus polymerase elongation complex (PVEC) with the template RNA strand in cyan and upstream non-template in bright green. The active site is colored magenta, and several key residues that facilitate catalysis are highlighted by numbering. Finally, the -1,+1, and +2 positions of the template RNA are numbered. (Gong and Peersen, 2010) B) Active site of the PVEC in same orientation as panel (A). The EC is colored grey with the exception of the -1 (red),+1 (yellow), and +2 (green) bases of the template RNA and the active site (magenta). C) Active site of the coxsackievirus polymerase elongation complex (CVEC) in the C2 form. D) Active site of the CVEC in the P1 form. E) Superimposition of panels B-D showing the alignment of the PVEC and CVEC in the C2 form and backsliding that occurs in the P1 form of the CVEC.

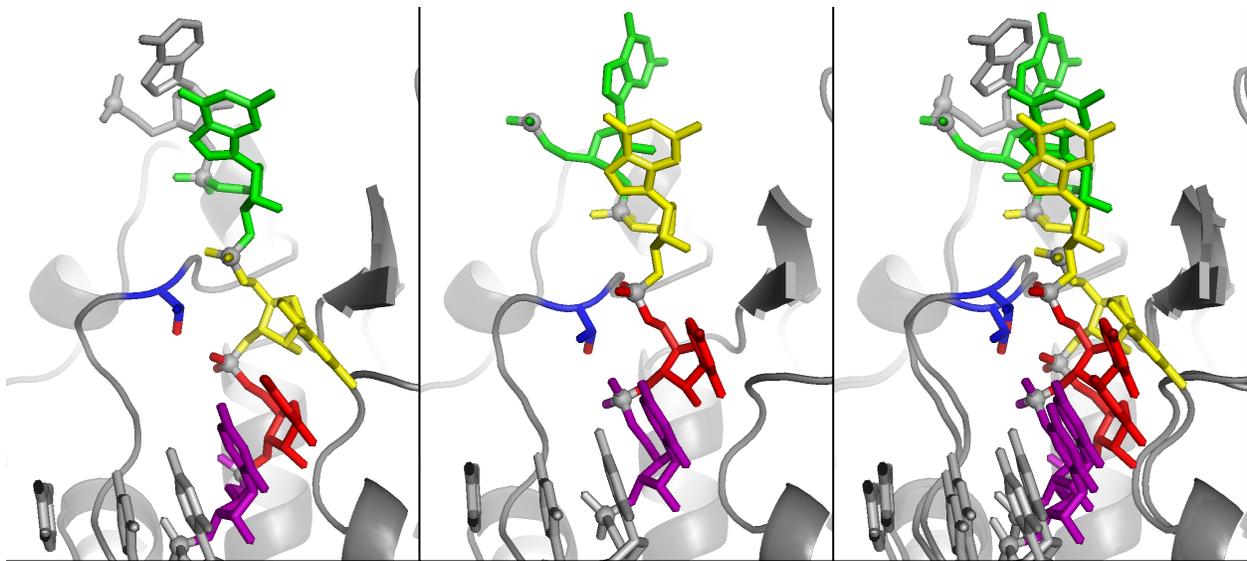
Template RNA backsliding in the P1 CVEC structure also causes shifts of the phosphate backbone relative to the surrounding protein structures. In both the PVEC structure and the C2 CVEC structure, Thr114 of the pinky finger resides between the -1 and +1 phosphates of the template RNA (Fig. 3.10A). The P1 CVEC structure shows Thr114 sitting between the -2 and -1 phosphates (Fig. 3.10B). This repositioning occurs due to the template RNA backsliding and causes the protein backbone surrounding Thr114 to shift downward to maintain the position of Thr114 in the middle of the surrounding phosphates (Fig. 3.10C). This occurs because the phosphate backbone does not shift by an entire nucleotide register during backsliding.

### *3.3.4 CVEC Crystal Soaks*

We also performed soaking experiments in an attempt to trap the elongation complex in different stages of the catalytic cycle, as was done in the PV studies. Soaking crystals produced by conditions 90 and 96 of the Qiagen screening kit with CTP or ddCTP caused the crystals to dissolve almost instantly. Other crystals remained intact upon the addition of CTP or various nucleoside analogs; however, after a short amount of time the crystals showed severe cracks that made them unworthy of harvesting. Finally, four crystals both remained intact and suffered only minimal damage due to soaking and were subsequently harvested for diffraction. However, these crystals gave only weak diffraction, and therefore were not processed any further.

## **3.4 Discussion**

The Peersen lab recently developed an experimental approach for obtaining crystals of viral polymerases stalled several nucleotides into replication. Using this method, they produced the structure of the poliovirus polymerase elongation complex



**FIG 3.11 Thr114 as a possible translocation gate.** The coxsackievirus polymerase elongation complex active site with the -2 (purple), -1 (red), +1 (yellow), and +2 (green) bases of the template RNA and Thr114 (blue) highlighted. A) Active site of the CVEC in the C2 form showing Thr114 positioned between the -1 and +1 phosphates (spheres) of the template RNA. B) Active site of the CVEC in the P1 form showing Thr114 positioned between the -2 and -1 phosphates (spheres) of the template RNA. C) Superimposition of panels A-B, illustrating the distortion of the protein backbone to maintain the position of Thr114 directly between two phosphates. Also, the differences in the positions of the RNA are readily seen.

(PVEC) capturing the polymerase with extensive upstream and downstream RNA (Gong and Peersen, 2010). The structure reveals key information regarding the interactions between the polymerase and the RNA. In addition, by soaking crystals with cognate CTP and various analogs, they were able to trap the complex in different stages of the overall catalytic cycle, thereby gaining insight into the conformational changes associated with replication. Of special significance, they determined the polymerase switches back and forth between open and closed conformations (Gong and Peersen, 2010). To complement this research, we utilized the same experimental approach in an attempt to produce the structure of the coxsackievirus polymerase elongation complex (CVEC). Results generated not only a CVEC structure comparable to the initial PVEC structure, but also the structure of a possible translocation intermediate. Comparison of these two structures with each other and the PVEC structures yields important information regarding picornaviral polymerase structure-function relationships.

Before carrying out crystallographic experiments, we assessed the kinetics of CV 3D<sup>pol</sup> using the initiation and stability assays employed to characterize the role of basic and hydrophobic residues in PV3D<sup>pol</sup> (Chapter 2). The results show the two viral polymerases exhibit relatively the same rates in both assays. The similarity in kinetics demonstrates the evolutionary conservation between these two viruses and their associated proteins. Since the data indicates CV 3D<sup>pol</sup> forms and maintains an elongation complex very efficiently it suggests the CVEC could potentially crystallize utilizing the same method designed for the PVEC structure.

In addition to the similarities found in the kinetic data, the structures of the PVEC and CVEC are highly homologous as well. Both structures display the same outward shift

of the fingers and thumb domains and the downward movement of Gly211 in order to accommodate RNA substrate (Gong and Peersen, 2010). Furthermore, the pathways of the template RNA into the active site and the upstream duplex through the clamp formed by the fingers and thumbs domain are nearly identical between the two structures. The high homology shared between these viral polymerase elongation complexes further demonstrates the evolutionary conservation between these two viruses. Furthermore, the success of the experimental approach used to obtain both structures suggests this method may provide structures of other viral RNA-dependent RNA polymerases.

Along with the kinetic data and the C2 CVEC structure, results from this project yielded an elongation complex trapped in a possible translocation intermediate. The P1 CVEC structure shows the template RNA shifted backwards from the active site relative to the C2 CVEC structure. This RNA backsliding causes a shift in both the location of the bases surrounding the active site and the phosphate backbone. Although the RNA in the P1 CVEC structure appears to have moved backwards relative to the existing PVEC structures and the C2 CVEC structure, we believe this may represent a complex trapped in a pre-translocation state. In other words, the P1 CVEC structure may show the position of the template RNA before translocation occurs, so that translocation causes the +1 base to swing down from the pocket above the active site into a position that allows for base stacking with the -1 base. This allows the incoming nucleotide to not only base pair with the +1 base of the template strand, but also base stack with the -1 base of product strand (Gong and Peersen, 2010). Thus, we believe this structure represents a biologically relevant conformation that normally exists in equilibrium with the previously captured

states. During translocation the RNA would move from the position observed in the P1 structure to the position seen in the C2 structure.

In addition, the observation that the phosphate backbone moves past Thr114 in the P1 CVEC structure to accommodate the RNA backsliding provides some evidence that this residue of the pinky finger may serve as a translocation gate. Again, we believe the P1 CVEC structure represents the state of the RNA before it moves into the position observed in the C2 CVEC structure. The phosphate group of the -1 base moves past Thr114 to allow for the +1 phosphate to move downward causing the +1 base to swing down into the active site for catalysis. Therefore, Thr114 appears to serve as the speed bump of translocation.

In conclusion, the comparison of the two CVEC structures with each other and the PVEC structures has confirmed many of the findings from the poliovirus studies while simultaneously providing insight into the mechanism of translocation. This information serves to increase our understanding of viral polymerases and may assist in the production of antiviral pharmaceuticals targeting the RNA-dependent RNA polymerase of these viruses.

## REFERENCES

- Ambros V, Pettersson RF, Baltimore D.** An enzymatic activity in uninfected cells that cleaves the linkage between poliovirion RNA and the 5' terminal protein. *Cell*. 1978 Dec;15(4):1439-46.
- Ambros V, Baltimore D.** Purification and properties of a HeLa cell enzyme able to remove the 5'-terminal protein from poliovirus RNA. *J Biol Chem*. 1980 Jul 25;255(14):6739-44.
- Andino R, Rieckhof GE, Baltimore D.** A functional ribonucleoprotein complex forms around the 5' end of poliovirus RNA. *Cell*. 1990 Oct 19;63(2):369-80.
- Arnold JJ, Cameron CE.** 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*. 2000 Feb 25;275(8):5329-36.
- Baltimore D.** In vitro synthesis of viral RNA by the poliovirus RNA polymerase. *Proc Natl Acad Sci U S A*. 1964 Mar;51:450-6.
- Barton DJ, Morasco BJ, Flanagan JB.** Translating ribosomes inhibit poliovirus negative-strand RNA synthesis. *J Virol*. 1999 Dec;73(12):10104-12.
- Beck MA, Chapman NM, McManus BM, Mullican JC, Tracy S.** Secondary enterovirus infection in the murine model of myocarditis. Pathologic and immunologic aspects. *Am J Pathol*. 1990 Mar;136(3):669-81.
- Bergelson JM, Cunningham JA, Droguett G, Kurt-Jones EA, Krithivas A, Hong JS, Horwitz MS, Crowell RL, Finberg RW.** Isolation of a common receptor for Coxsackie B viruses and adenoviruses 2 and 5. *Science*. 1997 Feb 28;275(5304):1320-3.
- Brandenburg B, Lee LY, Lakadamyali M, Rust MJ, Zhuang X, Hogle JM.** Imaging poliovirus entry in live cells. *PLoS Biol*. 2007 Jul;5(7):e183. Epub 2007 Jul 10.
- Campagnola G, Gong P, Peersen OB.** High-throughput screening identification of poliovirus RNA-dependent RNA polymerase inhibitors. *Antiviral Res*. 2011 Jun 21.

- Clark ME, Hämmerle T, Wimmer E, Dasgupta A.** Poliovirus proteinase 3C converts an active form of transcription factor IIC to an inactive form: a mechanism for inhibition of host cell polymerase III transcription by poliovirus. *EMBO J.* 1991 Oct;10(10):2941-7.
- Coyne CB, Bergelson JM.** Virus-induced Abl and Fyn kinase signals permit coxsackievirus entry through epithelial tight junctions. *Cell.* 2006 Jan 13;124(1):119-31.
- Crawford NM, Baltimore D.** Genome-linked protein VPg of poliovirus is present as free VPg and VPg-pUpU in poliovirus-infected cells. *Proc Natl Acad Sci U S A.* 1983 Dec; 80(24):7452-5.
- Crotty S, Saleh MC, Gitlin L, Beske O, Andino R.** The poliovirus replication machinery can escape inhibition by an antiviral drug that targets a host cell protein. *J Virol.* 2004 Apr;78(7):3378-86.
- De Jesus NH.** Epidemics to eradication: the modern history of poliomyelitis. *Virol J.* 2007 Jul 10;4:70.
- Drake JW.** Comparative rates of spontaneous mutation. *Nature.* 1969 Mar 22;221(5186):1132.
- Drake JW.** Rates of spontaneous mutation among RNA viruses. *Proc Natl Acad Sci U S A.* 1993 May 1;90(9):4171-5.
- Drake JW, Holland JJ.** Mutation rates among RNA viruses. *Proc Natl Acad Sci U S A.* 1999 Nov 23;96(24):13910-3.
- Domingo E, Menéndez-Arias L, Quiñones-Mateu ME, Holguín A, Gutiérrez-Rivas M, Martínez MA, Quer J, Novella IS, Holland JJ.** Viral quasispecies and the problem of vaccine-escape and drug-resistant mutants. *Prog Drug Res.* 1997;48:99-128.
- Dulbecco R.** Production of Plaques in Monolayer Tissue Cultures by Single Particles of an Animal Virus. *Proc Natl Acad Sci U S A.* 1952 Aug;38(8):747-52.
- Egger D, Teterina N, Ehrenfeld E, Bienz K.** Formation of the poliovirus replication complex requires coupled viral translation, vesicle production, and viral RNA synthesis. *J Virol.* 2000 Jul;74(14):6570-80.
- Esfandiarei M, McManus BM.** Molecular biology and pathogenesis of viral myocarditis. *Annu Rev Pathol.* 2008;3:127-55.

- Ferrer-Orta C, Arias A, Pérez-Luque R, Escarmís C, Domingo E, Verdaguer N.** Sequential structures provide insights into the fidelity of RNA replication. *Proc Natl Acad Sci U S A*. 2007 May 29;104(22):9463-8. Epub 2007 May 21.
- Fields.** *Virology* 5<sup>th</sup> Edition, Volume 1. Lippincott Williams and Wilkins. Philadelphia, 2007.
- Flanegan JB, Baltimore D.** Poliovirus-specific primer-dependent RNA polymerase able to copy poly(A). *Proc Natl Acad Sci U S A*. 1977 Sep;74(9):3677-80.
- Gamarnik AV, Andino R.** Switch from translation to RNA replication in a positive-stranded RNA virus. *Genes Dev*. 1998 Aug 1;12(15):2293-304.
- Gong P, Peersen OB.** Structural basis for active site closure by the poliovirus RNA-dependent RNA polymerase. *Proc Natl Acad Sci U S A*. 2010 Dec 28;107(52):22505-10. Epub 2010 Dec 10.
- Gong P, Campagnola G, Peersen OB.** A quantitative stopped-flow fluorescence assay for measuring polymerase elongation rates. *Anal Biochem*. 2009 Aug 1;391(1):45-55. Epub 2009 May 3.
- Hansen JL, Long AM, Schultz SC.** Structure of the RNA-dependent RNA polymerase of poliovirus. *Structure*. 1997 Aug 15;5(8):1109-22.
- Herold, J., Andino, R.** Poliovirus RNA replication requires genome circularization through a protein-protein bridge. *Mol. Cell* 2001. 7, 581-591.
- Hobson SD, Rosenblum ES, Richards OC, Richmond K, Kirkegaard K, Schultz SC.** Oligomeric structures of poliovirus polymerase are important for function. *EMBO J*. 2001 Mar 1;20(5):1153-63.
- Hogle JM, Chow M, Filman DJ.** Three-dimensional structure of poliovirus at 2.9 Å resolution. *Science*. 1985 Sep 27;229(4720):1358-65.
- Huang Y, Eckstein F, Padilla R, Sousa R.** Mechanism of ribose 2'-group discrimination by an RNA polymerase. *Biochemistry*. 1997 Jul 8;36(27):8231-42.
- Jablonski SA, Luo M, Morrow CD.** Enzymatic activity of poliovirus RNA polymerase mutants with single amino acid changes in the conserved YGDD amino acid motif. *J Virol*. 1991 Sep;65(9):4565-72.

- Jablonski SA, Morrow CD.** 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.* 1993 Jan;67(1):373-81.
- Jablonski SA, Morrow CD.** 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.* 1995 Mar;69(3):1532-9
- Jacobson SJ, Konings DA, Sarnow P.** 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.* 1993 Jun;67(6):2961-71.
- Jang SK, Kräusslich HG, Nicklin MJ, Duke GM, Palmenberg AC, Wimmer E.** A segment of the 5' nontranslated region of encephalomyocarditis virus RNA directs internal entry of ribosomes during in vitro translation. *J Virol.* 1988 Aug;62(8):2636-43.
- Joachims M, Van Breugel PC, Lloyd RE.** Cleavage of poly(A)-binding protein by enterovirus proteases concurrent with inhibition of translation in vitro. *J Virol.* 1999 Jan;73(1):718-27.
- Jore J, De Geus B, Jackson RJ, Pouwels PH, Enger-Valk BE.** Poliovirus protein 3CD is the active protease for processing of the precursor protein P1 in vitro. *J Gen Virol.* 1988 Jul;69 (Pt 7):1627-36.
- Jurgens CK, Barton DJ, Sharma N, Morasco BJ, Ogram SA, Flanagan JB.** 2Apro is a multifunctional protein that regulates the stability, translation and replication of poliovirus RNA. *Virology.* 2006 Feb 20;345(2):346-57. Epub 2005 Nov 17.
- Kamer G, Argos P.** Primary structural comparison of RNA-dependent polymerases from plant, animal and bacterial viruses. *Nucleic Acids Res.* 1984 Sep 25;12(18):7269-82.
- Kim DS, Nam JH.** Characterization of attenuated coxsackievirus B3 strains and prospects of their application as live-attenuated vaccines. *Expert Opin Biol Ther.* 2010 Feb;10(2):179-90.
- Kräusslich HG, Nicklin MJ, Toyoda H, Etchison D, Wimmer E.** Poliovirus proteinase 2A induces cleavage of eucaryotic initiation factor 4F polypeptide p220. *J Virol.* 1987 Sep;61(9):2711-8.

- Koonin EV.** The phylogeny of RNA-dependent RNA polymerases of positive-strand RNA viruses. *J Gen Virol.* 1991 Sep;72 ( Pt 9):2197-206.
- Landsteiner K, Popper E.** Mikroskopische preparate von eimen menschlichen und zwei affentuckermarker. *Weiner Klinische Wochenschrift* 1908;21:1930
- Loeffler F, Frosch P.** Report of the Commission for Research on foot-and-mouth-disease. IN: Hahon N,ed Selected Papers on Virology. Englewood Cliffs, NJ: Prentice Hall; 1964: 64-68
- Lyle JM, Bullitt E, Bienz K, Kirkegaard K.** Visualization and functional analysis of RNA-dependent RNA polymerase lattices. *Science.* 2002 Jun 21;296(5576):2218-22.
- Lyons T, Murray KE, Roberts AW, Barton DJ.** Poliovirus 5'-terminal cloverleaf RNA is required in cis for VPg uridylylation and the initiation of negative-strand RNA synthesis. *J Virol.* 2001 Nov;75(22):10696-708.
- Mendelsohn CL, Wimmer E, Racaniello VR.** Cellular receptor for poliovirus: molecular cloning, nucleotide sequence, and expression of a new member of the immunoglobulin superfamily. *Cell.* 1989 Mar 10;56(5):855-65.
- Merrick WC.** Overview: mechanism of translation initiation in eukaryotes. *Enzyme.* 1990;44(1-4):7-16.
- Mestas SP, Sholders AJ, Peersen OB.** A fluorescence polarization-based screening assay for nucleic acid polymerase elongation activity. *Anal Biochem.* 2007 Jun 15;365(2):194-200. Epub 2007 Apr 2.
- Mueller S, Wimmer E.** Recruitment of nectin-3 to cell-cell junctions through trans-heterophilic interaction with CD155, a vitronectin and poliovirus receptor that localizes to alpha(v)beta3 integrin-containing membrane microdomains. *J Biol Chem.* 2003 Aug 15;278(33):31251-60. Epub 2003 May 19.
- Murray KE, Barton DJ.** Poliovirus CRE-dependent VPg uridylylation is required for positive-strand RNA synthesis but not for negative-strand RNA synthesis. *J Virol.* 2003 Apr;77(8):4739-50.
- Nomoto A, Detjen B, Pozzatti R, Wimmer E.** The location of the polio genome protein in viral RNAs and its implication for RNA synthesis. *Nature.* 1977 Jul 21;268(5617):208-13.

- Novak JE, Kirkegaard K.** 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.* 1991 Jun;65(6):3384-7.
- Ollis DL, Brick P, Hamlin R, Xuong NG, Steitz TA.** Structure of large fragment of *Escherichia coli* DNA polymerase I complexed with dTMP. *Nature.* 1985 Feb 28-Mar 6;313(6005):762-6.
- Paul AV, van Boom JH, Filippov D, Wimmer E.** Protein-primed RNA synthesis by purified poliovirus RNA polymerase. *Nature.* 1998 May 21;393(6682):280-4.
- Paul AV, Rieder E, Kim DW, van Boom JH, Wimmer E.** Identification of an RNA hairpin in poliovirus RNA that serves as the primary template in the in vitro uridylylation of VPg. *J Virol.* 2000 Nov;74(22):10359-70.
- Pelletier J, Sonenberg N.** Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA. *Nature.* 1988 Jul 28;334(6180):320-5.
- Pfeiffer JK, Kirkegaard K.** Increased fidelity reduces poliovirus fitness and virulence under selective pressure in mice. *PLoS Pathog.* 2005 Oct;1(2):e11. Epub 2005 Oct 7.
- Racaniello VR.** One hundred years of poliovirus pathogenesis. *Virology.* 2006 Jan 5;344(1):9-16.
- Rieder E, Paul AV, Kim DW, van Boom JH, Wimmer E.** Genetic and biochemical studies of poliovirus cis-acting replication element cre in relation to VPg uridylylation. *J Virol.* 2000 Nov;74(22):10371-80.
- Rivera VM, Welsh JD, Maizel JV Jr.** Comparative sequence analysis of the 5' noncoding region of the enteroviruses and rhinoviruses. *Virology.* 1988 Jul;165(1):42-50.
- Rossmann MG, Arnold E, Erickson JW, Frankenberger EA, Griffith JP, Hecht HJ, Johnson JE, Kamer G, Luo M, Mosser AG, et al.** Structure of a human common cold virus and functional relationship to other picornaviruses. *Nature.* 1985 Sep 12-18;317(6033):145-53.
- Sabin, Ramos-Alvarez, Alvarez-Ameezquita, Pelon, Michaels.** Live, orally given poliovirus vaccine. Effects of rapid mass immunization on population under of massive enteric infection with other viruses. *JAMA.* 1960; 173:1521-1526

- Salk JE, Krech, Youngner, Bennet, Lewis, Bazeley.** Formaldehyde treatment and safety testing of experimental poliomyelitis. *Am. J. Public Health.* 1954; 44:56-70
- Schlegel A, Giddings TH Jr, Ladinsky MS, Kirkegaard K.** Cellular origin and ultrastructure of membranes induced during poliovirus infection. *J Virol.* 1996 Oct;70(10):6576-88.
- Sean P, Nguyen JH, Semler BL.** Altered interactions between stem-loop IV within the 5' noncoding region of coxsackievirus RNA and poly(rC) binding protein 2: effects on IRES-mediated translation and viral infectivity. *Virology.* 2009 Jun 20;389(1-2):45-58. Epub 2009 May 14.
- Shafren DR, Bates RC, Agrez MV, Herd RL, Burns GF, Barry RD.** Coxsackieviruses B1, B3, and B5 use decay accelerating factor as a receptor for cell attachment. *J Virol.* 1995 Jun;69(6):3873-7.
- Sosunov V, Zorov S, Sosunova E, Nikolaev A, Zakeyeva I, Bass I, Goldfarb A, Nikiforov V, Severinov K, Mustaev A.** The involvement of the aspartate triad of the active center in all catalytic activities of multisubunit RNA polymerase. *Nucleic Acids Res.* 2005 Jul 26;33(13):4202-11. Print 2005.
- Steitz TA.** A mechanism for all polymerases. *Nature.* 1998 Jan 15;391(6664):231-2.
- Takegami T, Kuhn RJ, Anderson CW, Wimmer E.** Membrane-dependent uridylylation of the genome-linked protein VPg of poliovirus. *Proc Natl Acad Sci U S A.* 1983 Dec;80(24):7447-51.
- Thompson AA, Peersen OB.** Structural basis for proteolysis-dependent activation of the poliovirus RNA-dependent RNA polymerase. *EMBO J.* 2004 Sep 1;23(17):3462-71. Epub 2004 Aug 12.
- Vignuzzi M, Gerbaud S, van der Werf S, Escriou N.** Naked RNA immunization with replicons derived from poliovirus and Semliki Forest virus genomes for the generation of a cytotoxic T cell response against the influenza A virus nucleoprotein. *J Gen Virol.* 2001 Jul;82(Pt 7):1737-47.
- Wimmer E.** Genome-linked proteins of viruses. *Cell.* 1982 Feb;28(2):199-201.
- Wimmer E, Nomoto A.** Molecular biology and cell-free synthesis of poliovirus. *Biologicals.* 1993 Dec;21(4):349-56.

- Xiang W, Harris KS, Alexander L, Wimmer E.** Interaction between the 5'-terminal cloverleaf and 3AB/3CDpro of poliovirus is essential for RNA replication. *J Virol.* 1995 Jun;69(6):3658-67.
- Yalamanchili P, Harris K, Wimmer E, Dasgupta A.** Inhibition of basal transcription by poliovirus: a virus- encoded protease (3Cpro) inhibits formation of TBP-TATA box complex in vitro. *J Virol.* 1996 May;70(5):2922-9.
- Yang Y, Rijnbrand R, Watowich S, Lemon SM.** Genetic evidence for an interaction between a picornaviral cis-acting RNA replication element and 3CD protein. *J Biol Chem.* 2004 Mar 26;279(13):12659-67. Epub 2004 Jan 7.
- Yogo Y, Wimmer E.** Polyadenylic acid at the 3'-terminus of poliovirus RNA. *Proc Natl Acad Sci U S A.* 1972 Jul;69(7):1877-82.
- Ypma-Wong MF, Dewalt PG, Johnson VH, Lamb JG, Semler BL.** Protein 3CD is the major poliovirus proteinase responsible for cleavage of the P1 capsid precursor. *Virology.* 1988 Sep;166(1):265-70.