

DISSERTATION

CHARACTERIZATION OF POLIOVIRUS 2C^{ATPASE} BOUND TO BILAYER NANODISCS
AND INVOLVEMENT OF THE POLIOVIRUS 3D^{POL} THUMB α -HELIX
IN DETERMINING POLY(A) TAIL LENGTH

Submitted by

Courtney Lee Springer

Department of Biochemistry and Molecular Biology

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Fall 2013

Doctoral Committee:

Advisor: Olve B. Peersen

P. Shing Ho
Karolin Luger
Alan Kennan

ABSTRACT

CHARACTERIZATION OF POLIOVIRUS 2C^{ATPase} BOUND TO BILAYER NANODISCS AND INVOLVEMENT OF THE POLIOVIRUS 3D^{POL} THUMB α -HELIX IN DETERMINING POLY(A) TAIL LENGTH

Poliovirus (PV) is a small non-enveloped picornavirus with a ≈ 7.5 kb long single-stranded, positive-sense RNA genome. Upon infection, the RNA is translated to generate a ≈ 250 kDa polyprotein that is subsequently cleaved into about a dozen fully processed proteins and several functional intermediates. PV replication occurs in large membrane associated complexes involving the “non-structural” P2 and P3 region proteins and two of these proteins, 2C^{ATPase} and 3D^{pol}, are the subjects of this dissertation.

Part I of this work is focused on the 2C protein, an AAA+ family ATPase that plays a key role in host cell membrane rearrangements and virion assembly, but the membrane binding characteristics of 2C and its polyprotein precursors have made it difficult to elucidate their exact roles in virus replication. In this work I show that small lipid bilayers known as nanodiscs can be used to chaperone the *in vitro* expression of soluble poliovirus 2C and the precursor 2BC and 2BC3AB polyproteins in a membrane bound form. Biochemical analysis shows that the proteins are highly active over a wide range of salt concentrations, exhibit slight lipid headgroup dependence, and show significant stimulation by acetate. Notably, the ATPase activity of the core 2C domain is stimulated ≈ 60 -fold as compared to the larger 2BC3AB polyprotein, with most of this

stimulation occurring upon removal of 2B. This data leads to a model wherein the viral replication complex can be assembled with a minimally active form of 2C that then becomes fully activated upon proteolytic cleavage from the adjacent 2B viroporin domain.

In Part II of this dissertation, I focus on the role of the viral RNA polymerase, 3D^{pol}, in maintaining the \approx 20-150 nucleotides long 3' poly(A) tail of the viral genome. The length of the tail is important for viral replication and initiation of (-)-strand synthesis, but the means by which the RNA is polyadenylated and how poly(A) tail length is regulated is not well understood. We have identified several mutations in an α -helix of the 3D^{pol} thumb domain that directly impact poly(A) tail length. Here, I tested the impact of these mutations on reiterative transcription of poly(A), poly(U), and poly(C) templates as well as characterized their effect on 3D^{pol} initiation, stability, elongation rate, and fidelity. I found that mutations in the thumb have the greatest impact on elongation complex stability and that 3D^{pol} is able to reiteratively transcribe homopolymeric poly(U) and poly(A), but not poly(C) RNA templates. Interestingly, distinct poly(A) and poly(U) transcripts are generated from 10 nucleotide homopolymers that are 1, 7, or 8 nucleotides longer than the template. Based on these findings, we propose a poly(A) slippage model in which the elongation complex stalls at the end of the homopolymer stretch in the absence of additional nucleotides to promote a single nucleotide slippage. This is followed by a slow structural rearrangement in which 3D^{pol} slips back to the 3' end of the homopolymer sequence, where it is able to re-transcribe starting from the fifth poly(U) in the template.

ACKNOWLEDGEMENTS

I would like to thank my advisor, Olve Peersen, for everything he has taught me and for giving me the opportunity to develop a wide array of skills in my work with 3D^{pol}, membrane proteins, and molecular dynamics. I would also like to thank my committee members Dr. Karolin Luger, Dr. Shing Ho, and Dr. Alan Kennan. I truly appreciate all their guidance and support over the years, both scientific and otherwise.

I would like to express my gratitude to the many people in the biochemistry department, faculty, colleagues, and fellow students alike, who have enriched my time as a graduate student in countless ways. I have also been very fortunate to have the chance to work with so many incredible people in the lab and I thank them for all their help and advice as well as for making “shience” so much fun! I would especially like to acknowledge Grace, who has been an incredible mentor and friend throughout my time in the lab.

Mom, Dad, Carol, Kent, and Jeannie, thank you for your unwavering love and support and for always being my cheerleaders. Dusty, thank you for always being there for me, for always knowing how to stand me back up, and for being my person. Chris, thank you for being so sweet and understanding, for all the meals you’ve made for me while I’ve been working on this dissertation, and for always knowing how to make me laugh. To all my other wonderful friends and family, thank you for being here for me during this journey. I love you all and could not have done this without you.

TABLE OF CONTENTS

ABSTRACT	ii
ACKNOWLEDGEMENTS	iv
TABLE OF CONTENTS	v
LIST OF TABLES	ix
LIST OF FIGURES	x
CHAPTER 1 Introduction	1
1.1 Overview of Poliovirus: History, Disease and Eradication Efforts	1
1.1.1 History	1
1.1.2 Disease	2
1.1.3 Eradication Efforts	2
1.2 Poliovirus Replication	6
1.2.1 Poliovirus: The Prototypical Picornavirus	6
1.2.2 Overview of the Viral Replication Cycle	7
1.2.3 Proteolytic Processing	11
1.2.4 Replication Centers	13
1.2.5 RNA Replication	16
1.3 Role of Poliovirus 2C in Viral Replication	18
1.3.1 Known functions of 2C	18
1.3.2 2C Shares Homology to AAA+ ATPases	20
1.4 Structure and Mechanism of 3D ^{pol}	21
1.4.1 Crystal Structure of 3D ^{pol}	21
1.4.2 Initiation of RNA Replication	23
1.4.3 Crystal Structure of the Elongation Complex and Mechanism of Elongation	23
1.4.4 3D ^{pol} is a Low Fidelity Polymerase	25

1.5 Scope of this Dissertation.....	26
CHAPTER 2 Polyprotein Context Regulates the Activity of Poliovirus 2C^{ATPase} Bound to Bilayer Nanodiscs	27
2.1 Introduction	27
2.2 Materials and Methods	32
2.2.1 Cloning	32
2.2.2 Nanodisc Assembly	34
2.2.3 In vitro Expression and Purification	35
2.2.4 Protein Yield and Concentration Determination	36
2.2.5 ATPase Activity Assay.....	38
2.3 Results	40
2.3.1 Lipid Dependence of Nanodisc Based in vitro Expression.....	40
2.3.2 ATPase Assay	45
2.3.3 Salt, pH, and Temperature Dependence	46
2.3.4 Magnesium and Acetate Dependence	48
2.3.5 Michaelis-Menten Kinetics of 2C Constructs.....	49
2.4 Discussion.....	55
2.4.1 Nanodisc-Bound 2C is a Highly Active ATPase.....	55
2.4.2 Proteolytic Control of 2C ATPase Activity	58
CHAPTER 3 Assessment of 2C Helicase Activity and Additional Preliminary Characterization	61
3.1 Introduction	61
3.2 Materials and Methods	62
3.2.1 Design and Preparation of RNA for Helicase Assays	62
3.2.2 Electrophoretic Mobility Shift Assays (EMSA).....	64
3.2.3 RNase T1 Helicase Assays.....	65
3.3 Results of 2C Helicase Assays.....	66
3.3.1 Assessment of 2C Helicase Activity via EMSA	66

3.3.2 Assessment of 2C Helicase Activity via RNase T1 Assays	73
3.4 Additional Preliminary Characterization	79
3.4.1 Mixing Experiments	79
3.4.2 Preliminary AUC and CD Data	81
CHAPTER 4 Involvement of the Poliovirus 3D^{pol} Thumb α-Helix in Determining Poly(A) Tail Length	82
4.1 Introduction	82
4.2 Materials and Methods	86
4.2.1 Protein Expression and Purification	86
4.2.2 RNA Oligonucleotides and Labeling	87
4.2.3 Poly(N) Slippage Assays	87
4.2.4 Initiation and Stability Assays	88
4.2.5 Stopped-Flow Assays to Determine 3D ^{pol} Elongation Rates and Fidelity	92
4.3 Results	93
4.3.1 Identification and effect of 3D ^{pol} mutations on poly(A) length and fidelity in HeLa cells and in vitro	93
4.3.2 Mutations Made in the Thumb Domain of Purified PV 3D ^{pol}	98
4.3.3 Effect of Mutations on Poly(N) Transcripts	99
4.3.4 Mutations in the Thumb have Variable Effects on Initiation and Stability	121
4.3.5 Thumb Mutants have Minor Effects on Elongation and Fidelity	125
4.4 Discussion	132
4.4.1 Multiple Determinants of Altered Poly(A) Tail Lengths	133
4.4.2 A Model for Poly(N) Slippage by Poliovirus 3D ^{pol}	136
CHAPTER 5 Conclusions and Future Directions	142
5.1 Introduction	142
5.2 2C ^{ATPase} : Summary and Future Directions	143
5.3 3D ^{pol} : Summary and Future Directions	145
5.4 Conclusion	148

REFERENCES	149
APPENDIX I Molecular Dynamics Simulations of Poliovirus 3D^{pol}	164
AI.1 Introduction.....	164
AI.2 Results and Discussion.....	168
<i>AI.2.1 Validation of MD Simulations by Modeling 3D^{pol} Thermal Denaturation</i>	168
<i>AI.2.2 MD Simulations of G64 Mutants</i>	171

LIST OF TABLES

CHAPTER 2	27
Table 2.1. Michaelis-Menten parameters of batch purified protein.....	52
Table 2.2. Michaelis-Menten parameters of FPLC purified protein.	53
CHAPTER 4	82
Table 4.1. Impact of 3Dpol mutations on poly(A) length and fidelity <i>in vivo</i>	94
Table 4.2. Quantitation of poly(A) and poly(U) transcripts.	106
Table 4.3. Quantitation results of poly(A) smearing in high versus low salt.	117
Table 4.4. 3D ^{pol} elongation rates.	128
Table 4.5. Fidelity of poliovirus 3Dpol.....	131
Table 4.6. Summary of <i>in vitro</i> and <i>in vivo</i> characterization results.	134

LIST OF FIGURES

CHAPTER 1	1
Fig. 1.1. Model of poliovirus pathogenesis.	3
Fig. 1.2. Poliovirus eradication efforts.....	5
Fig. 1.3. Poliovirus replication cycle.	8
Fig. 1.4. Poliovirus entry.....	9
Fig. 1.5. The poliovirus genome and polyprotein processing.	12
Fig. 1.6. Poliovirus forms and replicates on membranous structures.....	14
Fig. 1.7. Model for assembly of replication centers.....	15
Fig. 1.8. Positive- and negative-RNA synthesis.....	17
Fig. 1.9. 2C and the AAA+ ATPase domain.	19
Fig. 1.10. Structure of PV 3D ^{pol} and kinetic mechanism of RNA replication.	22
Fig. 1.11. Crystal structure of the 3D ^{pol} elongation complex.....	24
CHAPTER 2	27
Fig. 2.1. Poliovirus polyprotein processing and nanodisc chaperoned expression.....	28
Fig. 2.2. Protein concentration determination.	37
Fig. 2.3. Coupled ATP/NADH ATPase assays.	39
Fig. 2.4. Phospholipid dependence of Ub2C and Ub2BC expression and activity.	41
Fig. 2.5. Purification of <i>in vitro</i> translated 2C proteins.....	43
Fig. 2.6. ATPase activity of 2C proteins bound to nanodisc bilayers.....	47
Fig. 2.7. Michaelis-Menten analysis of nanodisc-bound 2C ATPase activity.....	50
CHAPTER 3	61
Fig. 3.1. Design of RNA constructs for 2C helicase assays.....	63
Fig. 3.2 Ub2C EMSA helicase assay with RNA-A/B/C.....	67
Fig. 3.3. EMSA competition assay with RNA-C and no protein.....	69

Fig. 3.4. Ub2C EMSA helicase assay with RNA-C in MgOAc vs. MgCl ₂	71
Fig. 3.5. Overview of RNase T1 helicase assay.	74
Fig. 3.6. RNaseT1 helicase assay with 0.1 U/μL T1.....	76
Fig. 3.7. RNaseT1 helicase assay with 1 U/μL T1 and proteinase K.	78
Fig. 3.8. 2C oligomerization state	80
CHAPTER 4	82
Fig. 4.1. Structural orientation of residues in PV 3D ^{pol} involved in the polyadenylation of viral RNA.	85
Fig. 4.2. RNA constructs.	89
Fig. 4.3. Initiation and stability of the 3D ^{pol} elongation complex.	91
Fig. 4.4. Reiterative transcription of a poly(U) sequence by purified 3D ^{pol}	96
Fig. 4.5. Poly(N) slippage assay with alanine α-helix thumb 3D ^{pol} mutants and the T4 RNA.	100
Fig. 4.6. Testing 8+2_A10, U10, and C10 RNA with wild type 3D ^{pol}	103
Fig. 4.7. Testing 8+2_A10, U10, and C10 RNA with 3D ^{pol} mutants.....	105
Fig. 4.8. Histogram of +13 band in A10 and U10 RNA reactions.	109
Fig. 4.9. Poly(A) slippage of D412N 3D ^{pol} dependence on ATP concentration.....	112
Fig. 4.10. WT , D412N, and S416T poly(A) slippage in high vs. low salt.	116
Fig. 4.11. 3D ^{pol} initiation times.	122
Fig. 4.12. 3D ^{pol} elongation complex stability.	124
Fig. 4.13. Determination of 3D ^{pol} elongation rates.	127
Fig. 4.14. Determining the fidelity of poliovirus 3D ^{pol}	130
Fig. 4.15. Proposed poly(A) slippage model.....	139
APPENDIX I	164
Fig. AI.1. Crystal structures of 3D ^{pol} showing important structural features.....	167
Fig. AI.2. 3D ^{pol} MD simulation at 498K.	169
Fig. AI.3. Average RMSF of 3D ^{pol} domains at 298K, 323K, and 498K.	170
Fig. AI.4. Inter-residue distances among 3D ^{pol} Gly64 mutants.....	172

CHAPTER 1

INTRODUCTION

1.1 Overview of Poliovirus: History, Disease and Eradication Efforts

1.1.1 History

Historical records suggest that poliovirus (PV) may have been in existence since as early as the 1500s B.C.; however the first scientific description of PV was not published until 1789 by Michael Underwood (1, 2). Clinical investigations of PV began shortly after Underwood's book, marked by the publication of Jacob Heine's book in 1840 (1, 3). PV was first identified as virus in 1908 by Landsteiner and Popper (4, 5), and in 1949 Enders, Weller, and Robbins reported the first instance of cultivating the virus in human cell culture, signifying a major breakthrough in poliovirus research by eliminating the need for animals in pathogenesis studies (1, 6). This research laid the groundwork for the development of two PV vaccines in the 1950-1960s, the inactivated PV vaccine by Jonas Salk and the live oral vaccine by Albert Sabin.

Since then, PV has been extensively researched and has been at the forefront of many important discoveries (7), including the discovery of uncapped mammalian messenger RNA (8), the internal ribosome entry site (IRES) (9, 10), and the 3'-terminal poly(A) sequence (10). PV was also one of the first two animal viruses for which the crystal structure of the viral capsid was obtained (11). Indeed, the breadth of knowledge regarding PV biochemistry and genetics has also made PV a useful model system for studying other viruses.

1.1.2 Disease

Poliovirus is a member of the *Enterovirus* genus, thus infection occurs in the gastrointestinal tract and the virus is primarily spread via a fecal-oral transmission cycle (Fig. 1.1). Upon ingestion, the virus infects cells in the pharynx and intestinal mucosa by binding to the CD155 receptor (4). Chimpanzees and Old World monkeys possess homologues of this receptor and can be experimentally infected with PV, although these species are not known to be infected in the wild, rendering humans the only known natural hosts (12). Once in the cells, the virus replicates and releases new viral particles into the blood; however, the majority of persons infected with PV will show no clinical symptoms and only $\approx 4-8\%$ of those infected will manifest minor flu-like symptoms (i.e. headache, fever, sore throat, etc.) (7). In $\approx 1-2\%$ of infected individuals the virus is able to pass the blood-brain barrier and spread to the central nervous system (CNS) by a mechanism that is not fully understood (7). Once in the CNS, PV infection leads to the destruction of motor neurons causing temporary or permanent muscle paralysis- a condition known as poliomyelitis (7). Death rates for those that develop paralytic poliomyelitis range from 2-75%, depending on the age of the individual and type of nerves infected (13).

1.1.3 Eradication Efforts

Prior to the availability of polio vaccines in the 1950's-60's, people around the world were suffering from the paralyzing effects of PV, with an estimated 21,000 paralytic cases reported in the United States alone in 1952 (13). Since then, there have been international efforts to eliminate this disease, and several of the key events in the

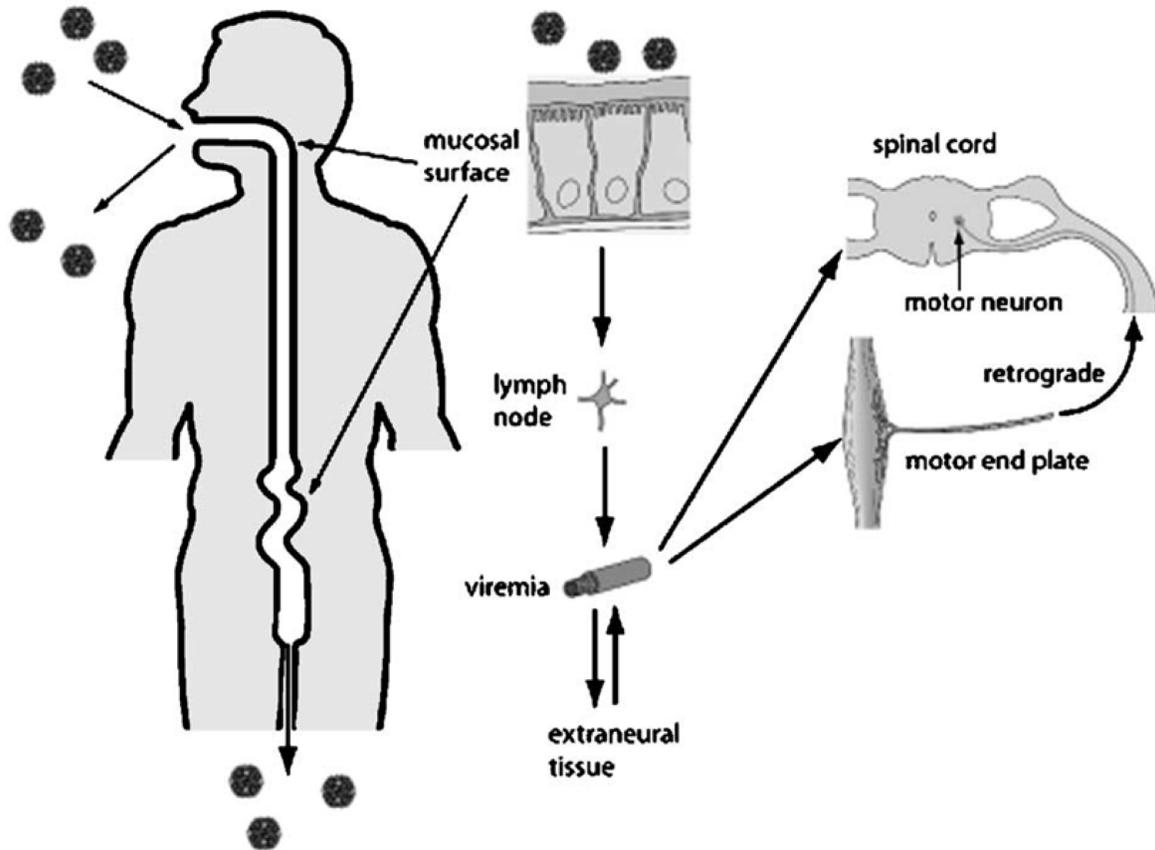


Fig. 1.1. Model of poliovirus pathogenesis.

Infection occurs upon ingestion of matter contaminated with virus containing fecal matter where it replicates in the cells at the mucosal surface in the gastrointestinal tract. The virus is introduced into the blood through the lymph nodes, where it continues to replicate (viremia), and new viral particles are shed in the fecal matter of the host. Occasionally, the virus is able to reach extraneural tissues, such as skeletal muscle and brown adipose tissue, resulting in sustained viremia and increasing the potential of entering the central nervous system. Figure from reference (4).

history of PV eradication are shown in Fig. 1.2A. Jonas Salk developed the first polio vaccine in 1955, the inactivated polio vaccine (IPV), which contained wild type virus that had been treated with formaldehyde in order to make it non-infectious (14). In 1963, the second PV vaccine, developed by Albert Sabin, was released. In contrast to the IPV, Sabin's vaccine contained "live" versions of the three main PV strains and could be administered orally. This oral polio vaccine (OPV) was generated by serially passaging the virus in monkeys as well as several different cell lines under sub-optimal conditions in order to select for attenuated virus that could still replicate in humans but were not able to invade the CNS (Fig. 1.2B) (14, 15).

International efforts to immunize against PV began in 1974 by the World Health Organization (WHO) with the expanded programme of immunization (EPI). Encouraged by the recent success of their campaign to eliminate smallpox, in 1988 the WHO launched a global campaign committed to the eradication of PV by the year 2000. The effectiveness, low cost, and the fact that the OPV did not require syringes made it the preferred method of vaccination for this campaign (14, 15). The efforts of the WHO made a significant impact, reducing the number of worldwide poliomyelitis cases from 350,000 in 1988 to just 2,971 in 2000 (16). Unfortunately, the goal of complete PV eradication still has not been achieved, ironically, because of the OPV itself; since OPV contains live virus that replicates in the gut of the vaccinated individual, it has the potential to develop mutations that alter its attenuated phenotype and this has resulted in the re-emergence of poliomyelitis (7, 14, 15, 17). Like wild type PV, OPV undergoes genetic reversions to a virulent, neuroinvasive strain on very rare occasions, causing vaccine-associated paralytic poliomyelitis (VAPP) in the vaccine recipient. The more

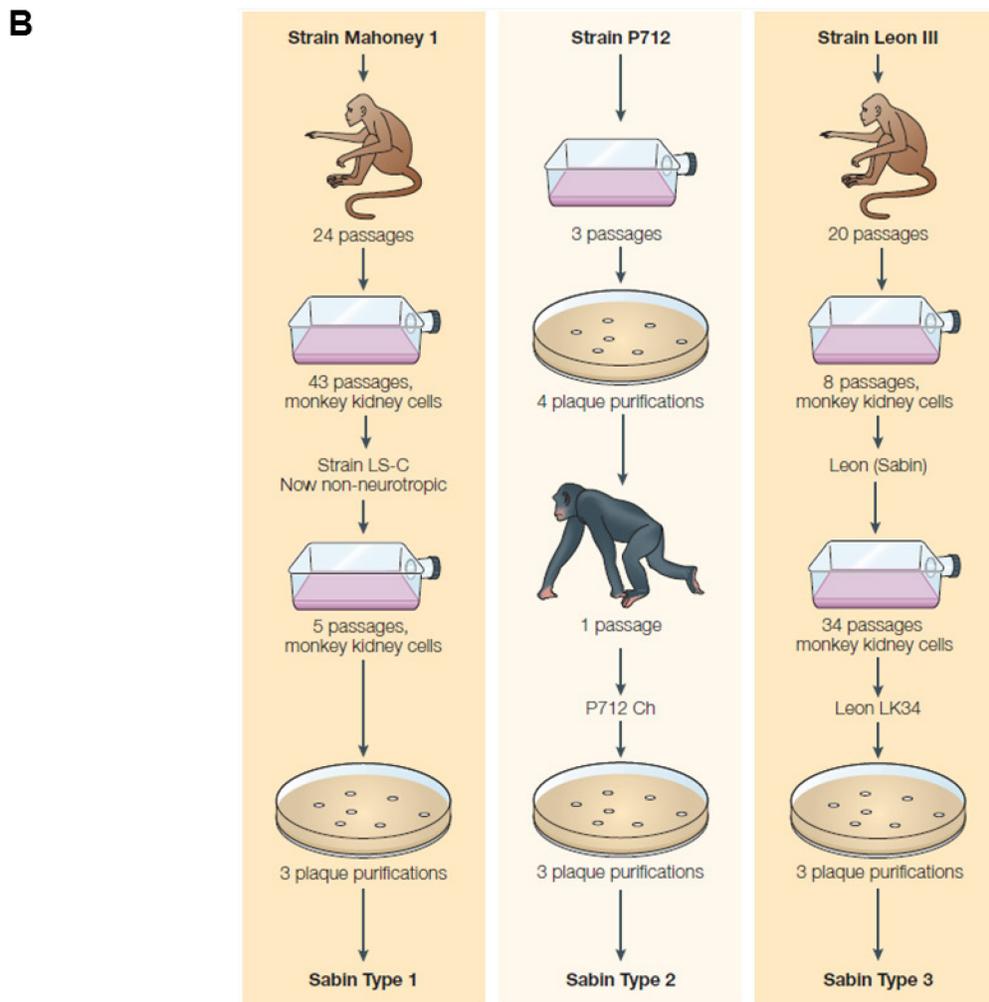
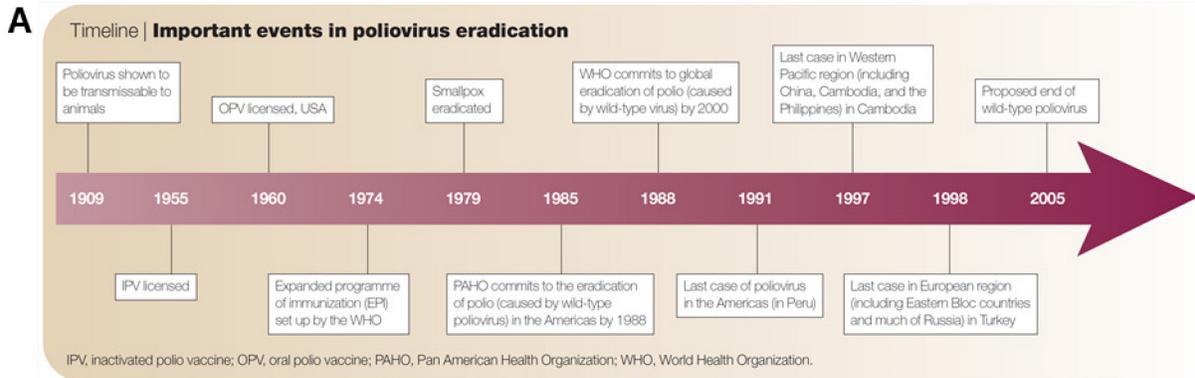


Fig. 1.2. Poliovirus eradication efforts.

A) Timeline of key events in the history of poliovirus eradication efforts. B) Diagram outlining the multiple rounds of selection that each of the three strains of PV were subjected to in order to isolate the attenuated versions of the virus that are used in Albert Sabin's oral polio vaccine. Figure from reference (15).

common cause of current poliomyelitis outbreaks are from circulating vaccine-derived polioviruses (cVDPVs), highly virulent strains believed to be the result of recombination between OPV strains and other enteroviruses. Populations that have been under vaccinated and have poor sewage management are particularly susceptible to cVDPV outbreaks. To overcome this hurdle in PV eradication, the WHO has made efforts to increase immunization and to improve global surveillance and outbreak response (www.polioeradication.org), reducing the total number of poliomyelitis cases to 291 in 2012 (16). OPV is still in use by the WHO due to its affordability and ability to impart gastrointestinal immunity, but the WHO is also aware that IPV is the long term solution to preventing VAPP and cVDPVs and is currently trying to find ways to make the IPV more affordable for low-income countries.

1.2 Poliovirus Replication

1.2.1 Poliovirus: The Prototypical Picornavirus

Poliovirus, which belongs to the genus Enterovirus, is a member of the family Picornaviridae- literally meaning “small RNA virus”, characterized by having small, single-stranded, positive-sense RNA genomes. Other notable members of the picornavirus family are hepatitis A virus, the rhinoviruses, foot-and-mouth disease virus, and coxsackie virus. These viruses cause health issues ranging from the common cold to liver disease to heart disease, and can potentially be fatal. Recent advances have been made towards developing compounds that inhibit various picornaviral proteins that are important for viral replication, such as the 2C and 3D^{pol} proteins that are the subjects of this dissertation. Indeed, several compounds have been designed that are

able to inhibit 2C and 3D^{pol} activity in multiple picornavirus species (reviewed in (18-20)). Despite the fact that PV is one of the most extensively studied picornaviruses, there are still many aspects of viral replication within the cell that we do not fully understand. Thus, continued research of the structure and mechanism of poliovirus proteins is not only important for understanding PV replication, but can provide insight into how replication occurs in other picornaviruses as well.

1.2.2 Overview of the Viral Replication Cycle

The replication cycle of poliovirus is representative of other picornaviruses and is outlined in Fig. 1.3. Infection begins upon binding of the virus to the CD155 cell receptor, also commonly referred to as the poliovirus receptor (PVR), at which point the virion is endocytosed into the cell via a tyrosine kinase- and actin-dependent mechanism (Fig. 1.3①) and the viral RNA is released (Fig. 1.3②). Depending on the type of cell PV has bound, there are currently two known pathways by which this process occurs (Fig. 1.4). Using HeLa cells (human cervical cancer cells), Brandenburg *et al.* (21) demonstrated that PV is internalized via an actin- and tyrosine kinase-dependent mechanism. After internalization, the viral RNA is released from the virion-containing endocytic vesicle within 20 min at a distance of ≈100-200 nanometers from the plasma membrane (Fig. 1.4A). A slightly different mechanism was seen by Coyne *et al.* (22) in human brain microvascular endothelial cells (HBMECs), which are representative of the blood-brain barrier that PV must infect in order to reach the CNS. Here, PV entry occurs by dynamin-dependent caveolar endocytosis, and PV RNA release occurs near the nucleus ≈2.5 hours after initial binding to CD155 (Fig. 1.4B).

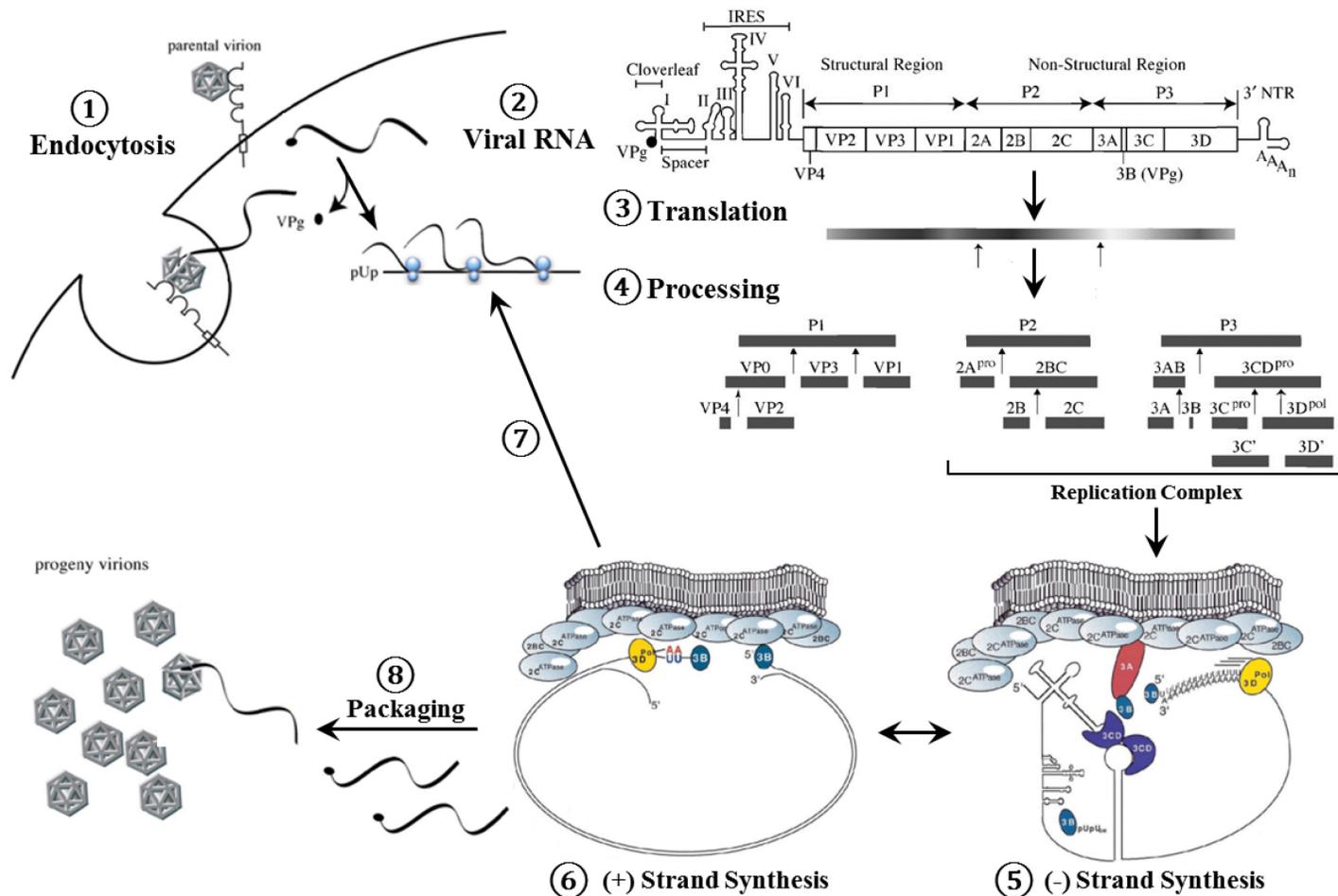


Fig. 1.3. Poliovirus replication cycle.

The PV replication cycle is initiated by binding to the CD155 cell surface receptor ①, followed by uncoating ② and translation ③ of the viral RNA. Proteolytic processing of the viral polyprotein yields structural and non-structural proteins ④. Positive-RNA acts as the template for complementary negative-strand synthesis ⑤ that is then used as a template to synthesize more positive-strands ⑥. The new positive-sense RNA can be translated to produce more viral proteins ⑦ or packaged into progeny virions ⑧. Derived from references (14, 23).

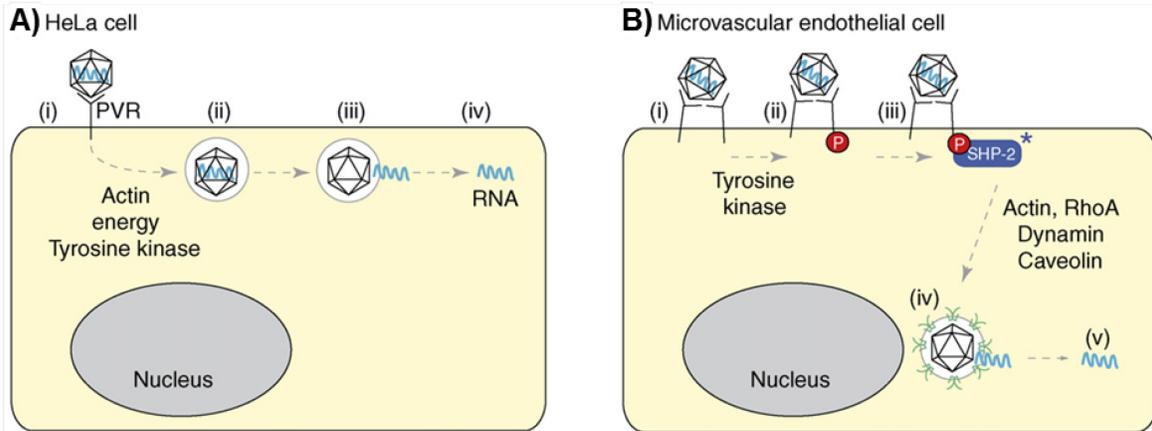


Fig. 1.4. Poliovirus entry.

A) In HeLa cells, binding to the PVR (CD155) instigates rapid (≈ 30 min) viral internalization and RNA release via a process that required a tyrosine kinase, ATP and an intact actin cytoskeleton (21). B) In microvascular endothelial cells, PV binding promotes dimerization and ligation of the PVR, which causes the receptor to be phosphorylated by a tyrosine kinase and activate SHP-2, a tyrosine phosphatase, leading to the activation of RhoA GTPase and causing actin rearrangement. PV enters the cell by a dynamin- and caveolin-dependent endocytosis and is transported near the nucleus where the viral RNA is release ≈ 2.5 hours after infection (22). Figure from reference (24).

Upon release of the viral RNA in the cell, the covalently linked protein primer VPg (or 3B) is cleaved from the 5'-end of the genome (25, 26), at which point the host cell ribosomes recognize and bind an internal ribosome entry site (IRES) in the viral RNA 5'-NTR (non-translated region) and translate the RNA into a single ≈ 250 kDa polyprotein (Fig. 1.3 ③). The polyprotein then undergoes a series of proteolytic cleavages, resulting in a total of eleven different proteins plus several important functional intermediates (Fig. 1.3 ④), discussed further in section 1.2.3). The eleven proteins in the viral polyprotein fall into one of three regions, P1, P2, or P3. The four proteins in the P1 region (VP1-VP4) are “structural proteins” that make up the virion capsid. The remaining seven proteins in the P2 and P3 regions (2A-2C and 3A-3D, respectively) are “non-structural” proteins, that are involved in viral replication and assembly of the viral structural proteins (27).

The non-structural viral proteins assemble on membranous structures to form the replication complex (discussed further in section 1.2.4) where $3D^{pol}$, the RNA dependent RNA Polymerase (RdRP), first synthesizes negative-strand RNA which it then uses as a template to synthesize positive-strand RNA (Fig. 1.3 ⑤, ⑥; discussed further in section 1.4). The resulting positive- sense RNA is either translated to generate additional viral proteins (Fig. 1.3 ⑦) or is packaged into progeny virions (Fig. 1.3 ⑧). The mechanism by which virions are released from the cell is not well understood, but there is evidence supporting both lytic and non-lytic release of virus particles (28-31).

1.2.3 Proteolytic Processing

Processing of the polyprotein plays an important role in the viral replication cycle, as different levels of processing impart different functions on the viral proteins, allowing PV proteins to carry out a wide array of tasks despite having a small genome (Fig. 1.5). Processing is carried out by the virally-encoded proteases 2A^{pro} and 3C^{pro}/3CD^{pro} to generate the fully processed proteins as well as several functional intermediates. 2A^{pro} carries out the first cleavage of the P1-P2 junction in *cis* as soon as it is translated and this is followed by the cleavage of the P2-P3 junction by 3C^{pro} (32). The proteases 3C^{pro} and 3CD^{pro}, a precursor protein whose protease activity is 3C-dependent, are responsible for further processing of the viral proteins to yield 2A and the 2BC precursor, which is further processed to yield 2B and 2C. 2BC and 2C sequester membranes from the host cell and induce the formation of tubular membrane structures (33-36). These structures serve as the foundation for replication centers (RCs) in which the P2 and P3 non-structural proteins act in concert to carry out viral RNA replication and assembly (34, 37, 38). The P3 region contains proteins that carry out viral RNA replication and synthesis. During processing, the P3 region is initially cleaved into the stable precursor proteins 3AB and 3CD^{pro}, and further cleavage of these two precursors yield 3A, 3B, 3C^{pro}, and 3D^{pol}. 3AB is believed to help form the RCs by recruiting 3CD^{pro} and 3D^{pol} to the membrane structures generated by 2BC and 2C (33). Free 3B is the 22 residue polypeptide VPg (viral protein genome-linked) that acts as a primer for RNA synthesis by the viral RNA polymerase 3D^{pol} (39). In addition to having protease activity, 3CD^{pro} plays an important role in replication by serving as a RNA binding protein, but it displays no intrinsic polymerase activity. 3D^{pol} becomes an active polymerase only after

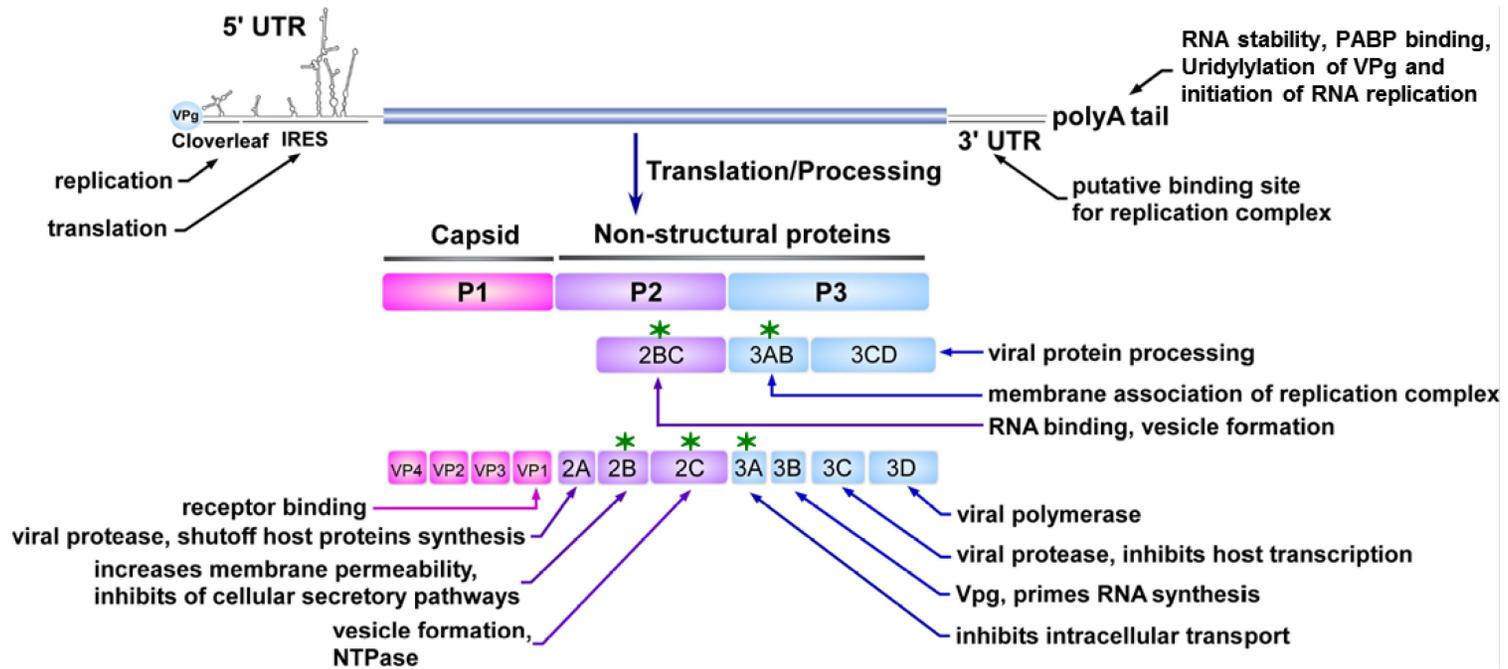


Fig. 1.5. The poliovirus genome and polyprotein processing.

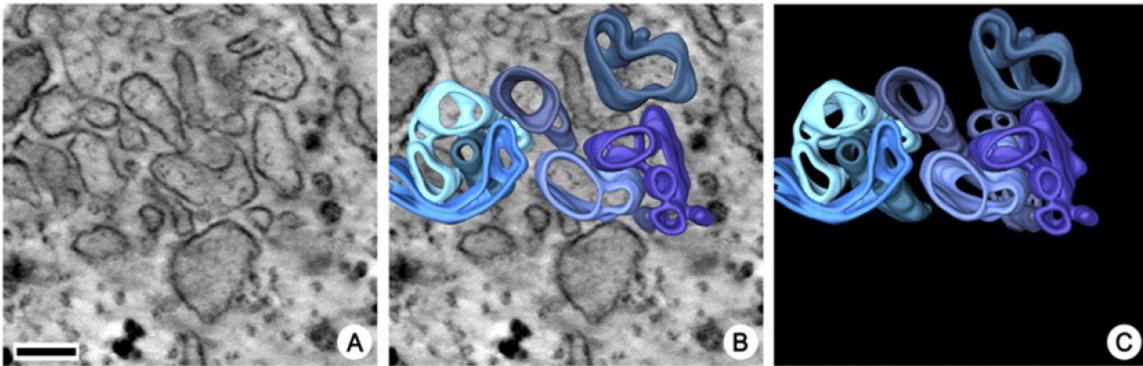
Poliovirus genome showing the small VPg (3B) protein primer at the 5'-end, the 5' UTR (untranslated region) containing the IRES, the open reading frame encoding the polyprotein, followed by the 3' UTR and a poly(A) tail. Processing of the polyprotein yields 11 fully processed proteins and several functional intermediates. The known functions are listed for the various cleavage products. Viral proteins that bind membranes are indicated by green stars. Adapted from (40).

cleavage from the 3CD^{pro} precursor protein and burial of the newly created N-terminal glycine in a pocket at the back of its palm domain (41). Our laboratory has shown that adding or deleting even a single residue to the N-terminus of 3D^{pol} abolishes activity, indicating that precise cleavage of the N-terminus is key in regulating 3D^{pol} function (41, 42). In Chapter 2 of this dissertation I will discuss my investigations of how proteolytic processing can also modulate 2C ATPase activity.

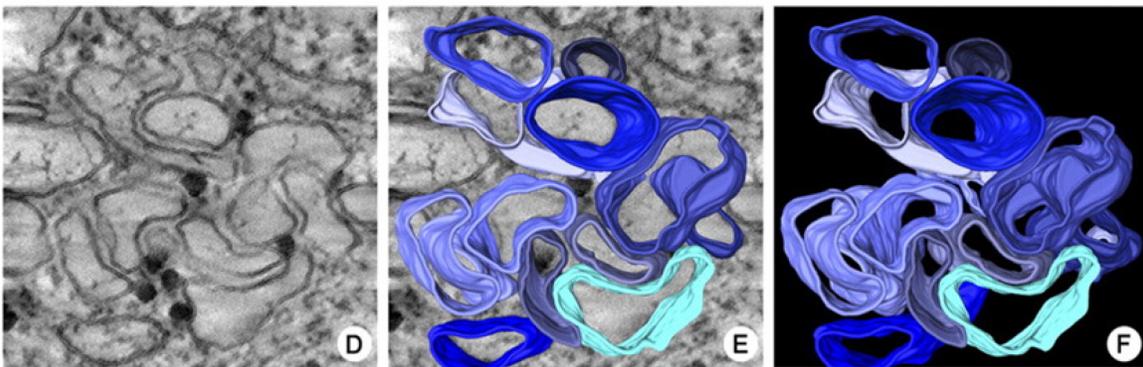
1.2.4 Replication Centers

Poliovirus, as well as other positive-strand RNA viruses, require membranes in order to carry out viral replication (43). PV infection induces massive membrane rearrangements to form what was originally thought to be clusters of vesicles ranging from 70-400 nm in diameter (44). The viral proteins assemble on these membranes to form replication centers/complexes (RCs) where viral RNA replication and packaging is carried out (34, 45-47). Recent electron tomography images reveal that these “vesicles” are actually an interconnected “web” of tubular membrane structures (Fig. 1.6 and Fig. 1.7C) (36). In this study, Belov et al. observed that these structures evolved from single- to double-walled membranes over the course of the early- to late-stage of PV infection, with the late-stage, double-walled structures exhibiting very little branching and appearing to be sealed. The authors found viral proteins and RNA associated in patches on the membrane surfaces throughout infection, although the highest amount of viral RNA replication was associated with the single membrane structures seen at the early and intermediate stages of infection. This report also supports several recent studies implicating the Golgi/Trans Golgi Network (TGN) and as the immediate source

Early Structures (3 hpi)



Intermediate Structures (4 hpi)



Late Structures (7 hpi)

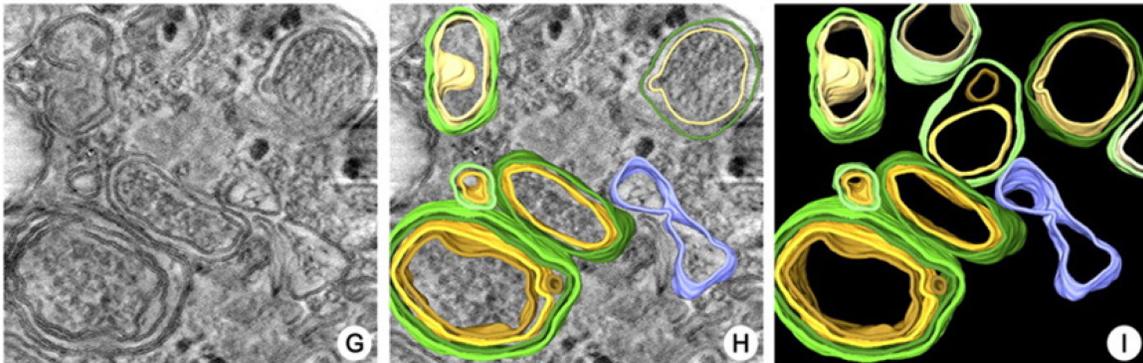


Fig. 1.6. Poliovirus forms and replicates on membranous structures.

3D reconstructions of poliovirus membranous replication complexes at the early (3 hpi), intermediate (4 hpi), and late (7 hpi) stages of the replication cycle. (A, D, and G) Central slices in tomographic volumes. (B, E, and H) Central slices with segmented overlays. (C, F, and I) Segmented volumes, with blue indicating single-membrane structures and yellow and green indicating inner and outer membranes of double-membrane structures, respectively. The bar in (A) represents 100 nm. Figure from reference (36).

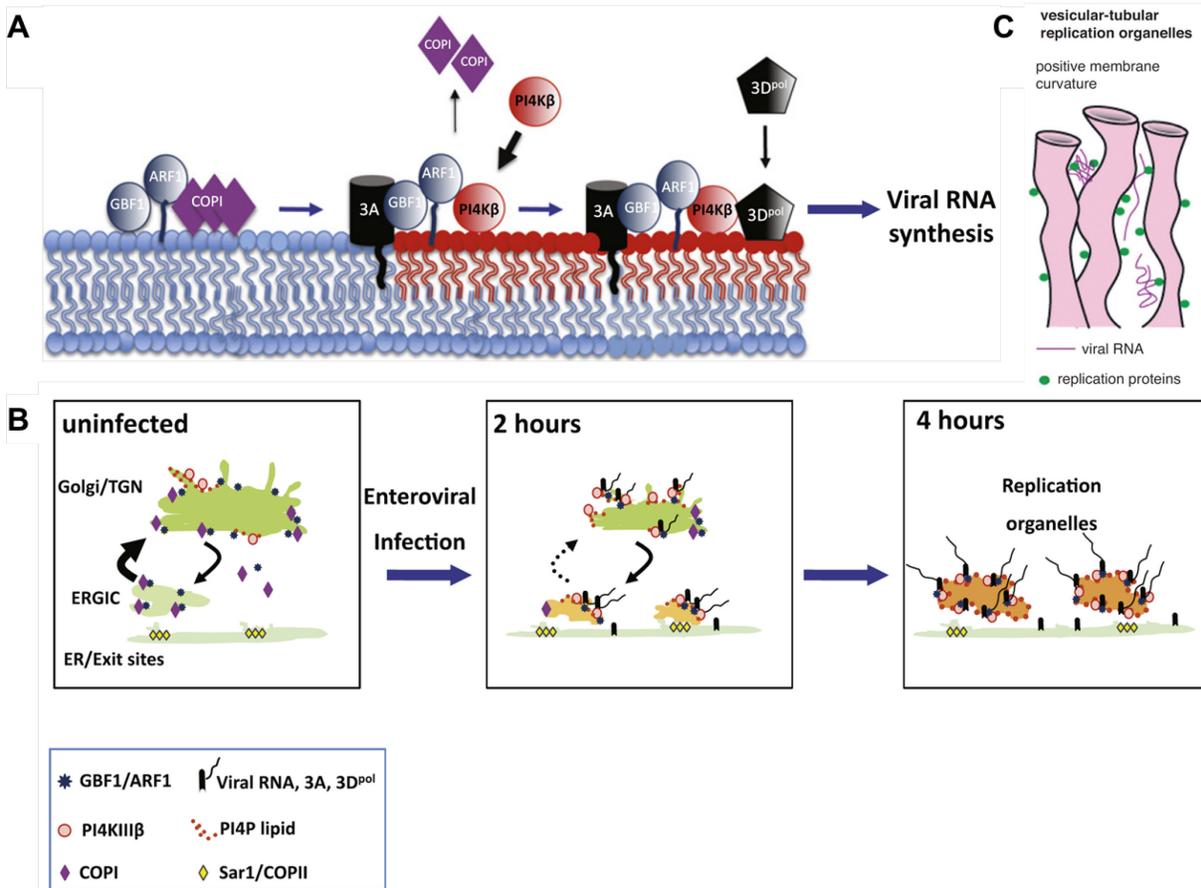


Fig. 1.7. Model for assembly of replication centers.

A) Viral 3A binds to host proteins GBF1/Arf1 to recruit PI4KIIIβ and the removal of COPI proteins. PI4KIIIβ then catalyzes the production of PI4P lipids (red lipids) that act to promote the association of 3D^{pol} where it can begin synthesizing the viral RNA. B) The typical exchange of membranes between the Golgi/TGN and the ER-Golgi intermediate compartment (ERGIC) in uninfected cells is disrupted upon infection by 3A-mediated inhibition of anterograde transport from the ERGIC to the Golgi/TGN (2 hrs.). The Golgi/TGN, already enriched with PI4P lipids prior to infection, serves as the initial site of viral protein assembly and RNA replication (2hrs). Shutoff of anterograde trafficking eventually leads to the deterioration of the Golgi/TGN and the formation of RC organelles that localize near ER exit sites (denoted by yellow diamonds representing Sar1/COPII, 4 hr. panel). (48). C) Representation of the tubular membranes structures that comprise the RC organelles upon which the viral replication proteins localize on the cytoplasmic side of the membrane (49). Adapted from (48, 49).

of lipids for the viral replication centers (36, 48, 50), rather than the endoplasmic reticulum (ER) as has been previously suggested (reviewed in (43)). The model for formation of these membrane structures by the viral and host proteins involved is shown in Fig. 1.7. In this model, the viral 3A protein interacts with host proteins to facilitate the disruption of the cellular secretory pathway and divert membranes from the Golgi/TGN to provide material for the RC organelles (48). In addition to inhibition of ER-to-Golgi trafficking by 3A, RC assembly also involves the other viral membrane-binding proteins, 2BC and 2C. No individual viral protein is able to fully reproduce the membrane structures observed during PV infection, but expression of either 2BC or 2C alone results in the formation of single-membraned vesicles (37, 38, 51, 52), and co-expressed 2BC and 3A produce double-membraned vesicles (52).

1.2.5 RNA Replication

Once the replication centers have been assembled, replication of the viral RNA by 3D^{pol} is ready to commence. However, this process also involves several host proteins, including the poly(rC)-binding protein 2 (PCBP2) (53, 54) as well as poly(A)-binding proteins (PABP) (55), and one model for how this is thought to occur is shown in (Fig. 1.8) (56). After translation, the positive-strand RNA is bound at the 5' UTR by PCBP2 as well as the PV precursor protein 3CD, and is bound at the 3'-UTR poly(A) tail by PABP. PCBP2 and PABP interact, causing the two ends of the RNA to come together and placing 3CD in close proximity to the poly(A) tail (Fig. 1.8A). 3CD is then cleaved, generating an active 3D^{pol} that is able to initiate negative strand synthesis using VPg as a primer (Fig. 1.8B) and synthesize the double stranded replicative form

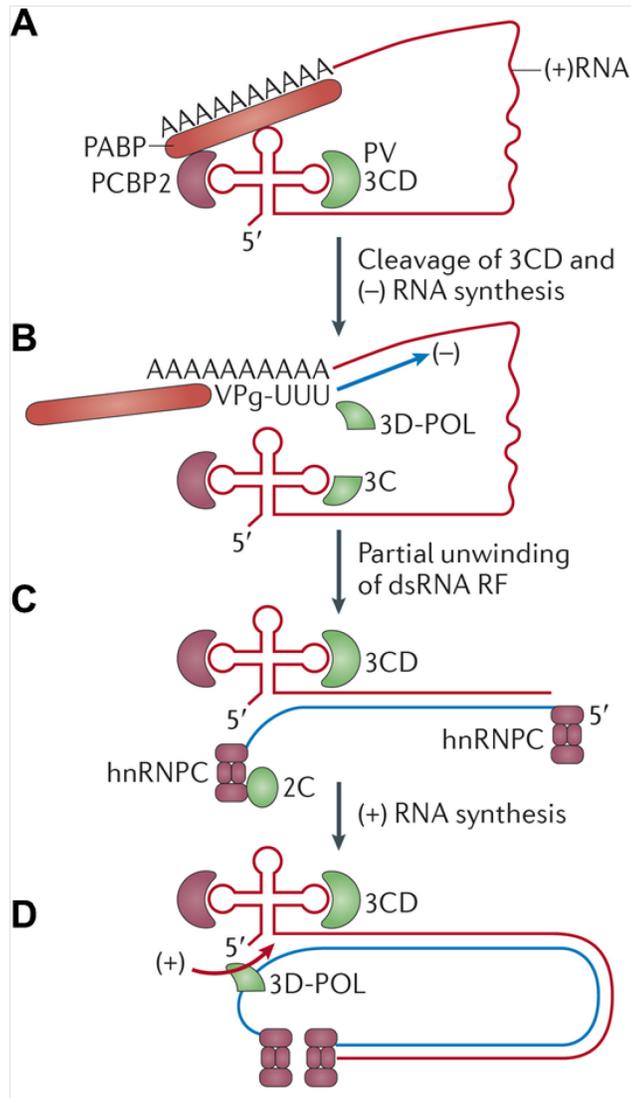


Fig. 1.8. Positive- and negative-RNA synthesis.

A) Initiation of negative-strand RNA synthesis by 3D^{pol} requires the circularization of the positive- strand RNA via contacts between the poly(A) binding protein (PABP) and the poly(C) binding protein (PCBP2), bound at the 5'- and 3'-end of the RNA, respectively. B) 3D^{pol} is liberated from 3CD and primes off of VPg to synthesize negative-strand RNA. (C-D) This results in the double stranded replicative form of RNA (dsRNA RF) from which multiple strands of positive-sense RNA are synthesized in a process that could involve viral protein 2C and/or the host protein hnRNPC. Figure from reference (56).

(RF) of the RNA. The newly synthesized negative-strand RNA is then thought to serve as a template for multiple rounds of positive-strand synthesis (Fig. 1.8C, D). The process by which the transition from positive to negative strand synthesis occurs is not well understood, but it has been suggested that several proteins may facilitate this switch, including the 2C protein that may help anchor the RNA to the replication centers, and the host heterogeneous nuclear ribonucleoprotein (hnRNP) that has been found to promote viral positive strand synthesis (35, 57-60). The resulting positive-strand RNA is then used to translate more viral proteins or packaged into virions, as shown in Fig. 1.3(7),(8).

1.3 Role of Poliovirus 2C in Viral Replication

1.3.1 Known functions of 2C

Poliovirus 2C is a 38 kDa protein that has been shown to play a role in RNA binding (61), membrane rearrangement (37, 38, 51), initiation of negative-strand RNA synthesis (62), and virion packaging (63-66). 2C is comprised of 329 amino acids and contains an N-terminal amphipathic helix, two predicted RNA binding regions, three NTP binding motifs, and a possible zinc finger motif (61, 67, 68). Two of the NTP binding motifs are Walker A and B motifs, while the third (motif C) shares homology to the helicase superfamily III (Fig. 1.9A) (69). 2C has been shown to have ATPase activity and there is some indication of GTPase activity, although there is conflicting evidence for this (68-70). 2C is known to associate with membranes via its N-terminal amphipathic helix and is involved in forming the membrane structures upon which viral replication occurs (37).

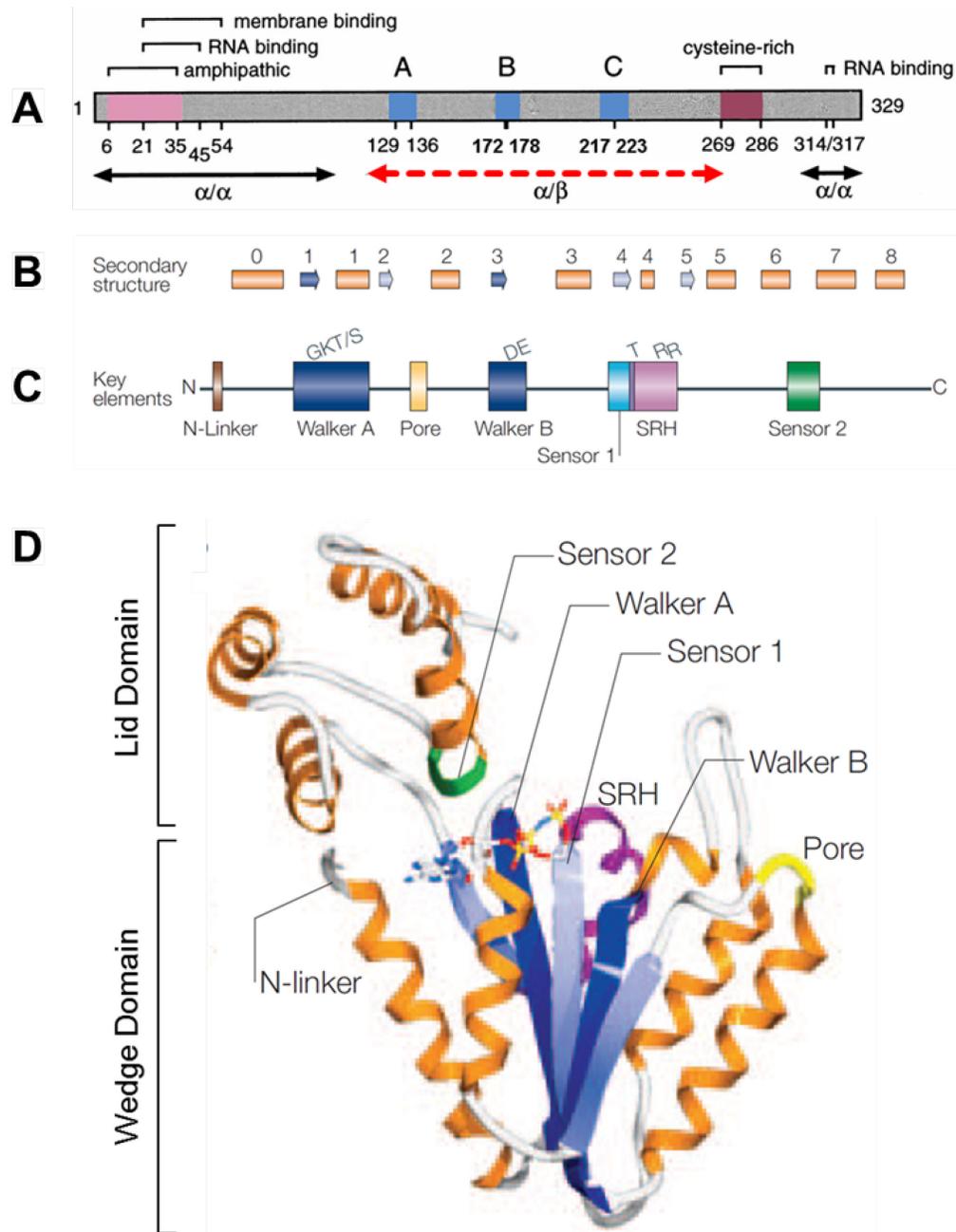


Fig. 1.9. 2C and the AAA⁺ ATPase domain.

A) Poliovirus protein 2C functional domains. The blue regions represent NTP binding Walker A and B motifs, and motif C. The red dashed arrow indicates the possible AAA⁺ ATPase domain. B) Secondary structural features of a typical AAA⁺ domain (α -helices are rectangles and β -strands are arrows) in alignment with C) key elements within the AAA⁺ domain. D) The crystal structure the NSF¹⁰³ AAA⁺ domain with the location of key elements. Figure adapted from references (71, 72).

1.3.2 2C Shares Homology to AAA+ ATPases

The function of 2C ATPase activity is not yet clear and sequence homology searches do not reveal any convincing matches for 2C other than other picornaviral proteins. However, further investigation using Phyre (73), which uses profile-profile sequence matching and secondary structure prediction to do homology modeling, suggests that residues ~100-260 are similar to the domain characteristically found in the ATPases associated with diverse cellular activities (AAA⁺) protein family. The AAA⁺ ATPase superfamily of enzymes are often considered protein “workhorses” because they display a multitude of functions, including protein degradation, membrane rearrangement, DNA replication and disassembly of protein complexes (72, 74). All are characterized by Walker A and B motifs and tend to oligomerize, usually into defined hexamers or extended fibers. Crystal structures reveal there are two regions within the AAA⁺ domain that are the hallmark of the AAA⁺ family: a wedged-shaped N-terminal region composed of α -helices and parallel β -strands, and a C-terminal domain composed of several α -helices that form a partial lid over the nucleotide-binding site (Fig. 1.9D). 2C is predicted to share this secondary structure and folding arrangement. In addition to structural homology, 2C also shares several key elements that are characteristic of AAA⁺ ATPases and are necessary for ATP binding and/or hydrolysis, including the Walker A and B motifs and the sensor 1 and Arginine fingers located in the second region of homology (SRH) (Fig. 1.9C). Indeed, previous studies have demonstrated that mutating residues in the Walker A, Walker B, and Sensor 1 motifs prevent the ability of 2C to hydrolyze ATP (68, 69). Several studies also indicate that

PV and FMDV 2C both form hexameric rings and that this oligomeric state may be important for coordinated ATP hydrolysis (75, 76).

There is still much that remains unclear as to the role of protein 2C in the viral replication process and it is the last of the poliovirus proteins whose structure remains unsolved. 2C has several known functions that are vital to the viral life cycle, including formation of replication center vesicles and binding viral RNA. Indeed, the inhibition of poliovirus RNA replication by guanidine (77, 78) has been linked to 2C function via resistance mutations (79, 80). 2C has also recently been shown to interact directly with the VP3 capsid protein, suggesting a role in viral RNA encapsidation (65). These data, in combination with its homology to the helicase superfamily III, have led to the hypothesis that 2C may have helicase or translocase activity, although neither has yet been definitively shown. Interestingly, a recent study of the closely related *Ectoropsis obiqua* picorna-like virus (EoV) shows that the EoV 2C protein is able to destabilize nucleic acid helices in an ATP-independent manner, suggesting that this protein may be more similar to an RNA chaperone rather than a helicase (81).

1.4 Structure and Mechanism of 3D^{pol}

1.4.1 Crystal Structure of 3D^{pol}

As mentioned above, 3D^{pol} is the RNA-dependent RNA polymerase (RdRp) that carries out all viral RNA replication. The crystal structure of 3D^{pol} (Fig. 1.10A) reveals that the protein shares the typical “right hand” structure that is observed in many polymerases, in which the active site residues in the palm domain is encircled by the index, middle, ring, pinky, and thumb domains (41). The ring finger serves as the roof

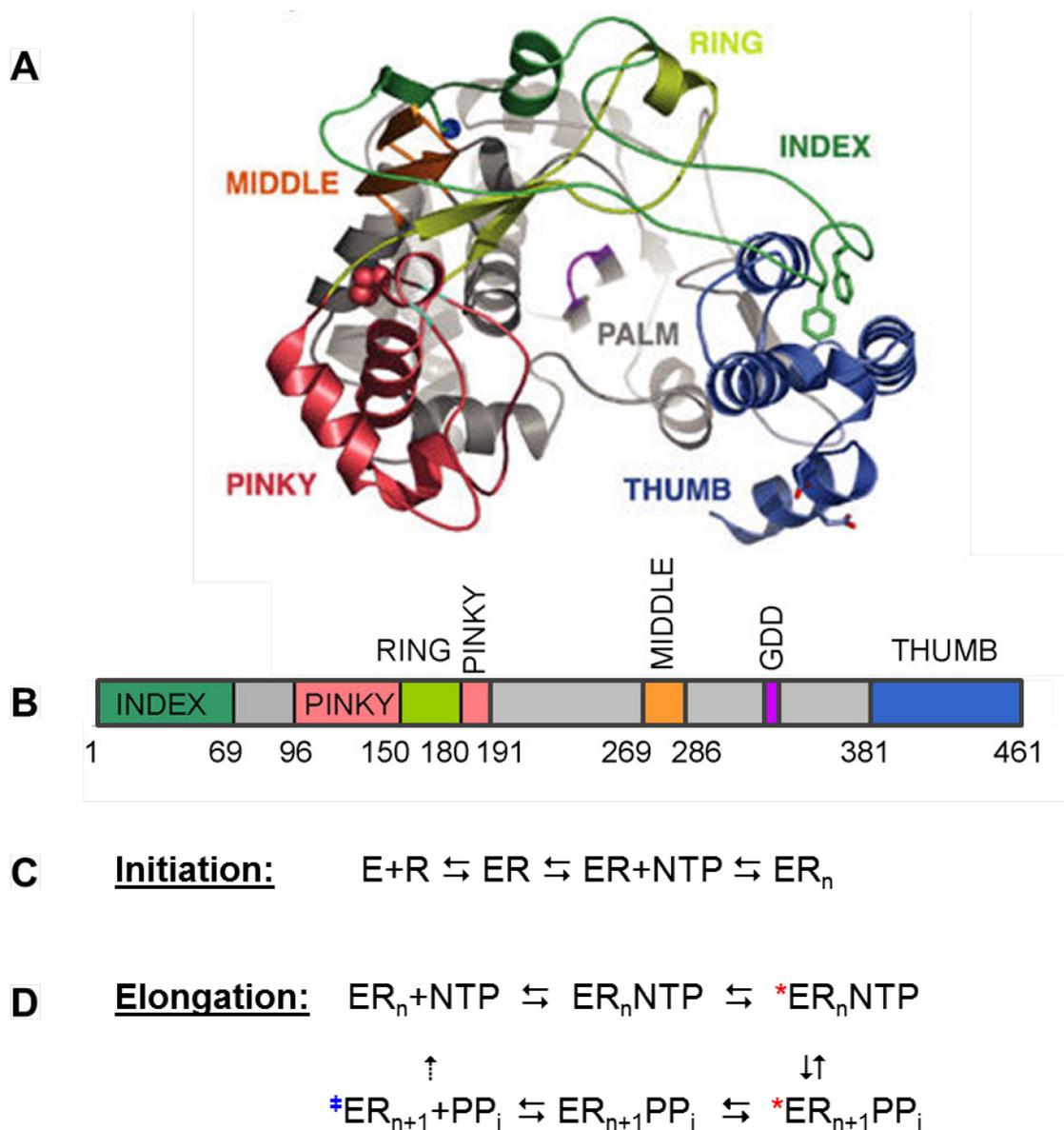


Fig. 1.10. Structure of PV 3D^{pol} and kinetic mechanism of RNA replication.

A) Top view of the crystal structure of poliovirus 3D^{pol}. Regions are color coded according to their “right hand” fingers domain. PDB code 1RA6 (41). These regions are also shown on the bar in (B) indicating their location in the 3D^{pol} amino acid sequence. GDD (magenta) corresponds to the active site sequence. C-D) Kinetic mechanism of 3D^{pol} initiation (C) and elongation (D). ‘E’ represents the enzyme 3D^{pol}, R represents the RNA and _n represents the incorporated NTP. The red asterisk (*) indicates the active site is in the closed conformation and (*) indicates the step at which a conformational change associated with translocation is thought to occur (82).

NTP entry channel on the back of the polymerase and is held in place by looping under the index finger. The index finger (residues 1-69, green) then extends over to the thumb domain to form a circle around the active site (GDD motif from residue 237-239, magenta).

1.4.2 Initiation of RNA Replication

Biochemical evidence indicates that there are two conformational changes that take place in order for $3D^{pol}$ to initiate RNA synthesis (Fig. 1.10C). Upon binding of $3D^{pol}$ to the RNA (E+R), it is thought to undergo a slow conformational change ($t_{1/2} \approx 12$ sec) that results in a $3D^{pol}$ -RNA complex (ER) that has as a half-life ranging from 20 min to 2 hrs at 22°C and 30°C, respectively (83, 84). Incorporation of the first nucleotide is slow compared to subsequent rounds of NTP incorporation, suggesting it undergoes another conformational change, but the resulting complex (ER_n) is significantly more stable ($t_{1/2} \approx 2.5$ hrs at 30°C, ≈ 8.5 hrs at 22°C) than the $3D^{pol}$ -RNA complex alone (74). This state in which $3D^{pol}$ has incorporated the first nucleotide and is poised for elongation is referred to as the elongation complex (EC).

1.4.3 Crystal Structure of the Elongation Complex and Mechanism of Elongation

In 2010, the Peersen laboratory published the first crystal structure of $3D^{pol}$ in complex with RNA that had already undergone several rounds of nucleotide incorporation (Fig. 1.11) (82). In addition to obtaining the structure of the $3D^{pol}$ EC, several different states of the elongation catalytic cycle were also captured by soaking NTPs or NTP analogs into the crystals, and these steps are outlined in Fig. 1.10D (82).

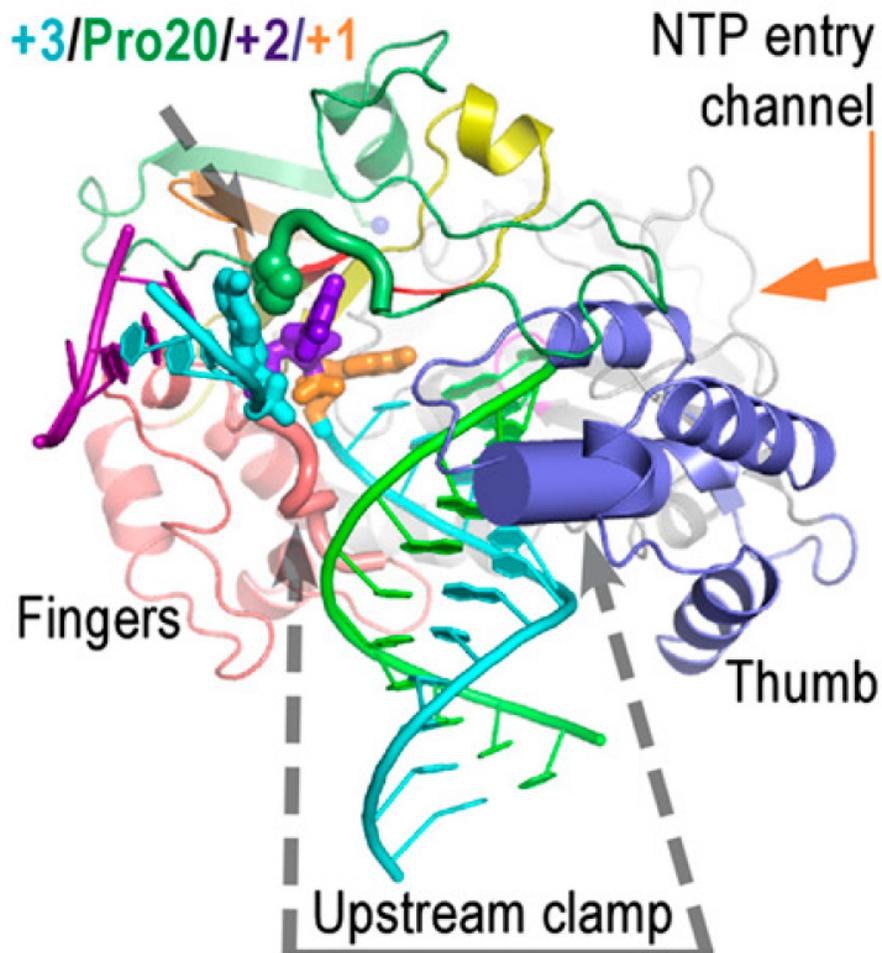


Fig. 1.11. Crystal structure of the 3D^{pol} elongation complex.

3D^{pol} is colored using the same coloring scheme shown in Fig. 1.10. The grey arrow shows the trajectory of the RNA, with the template strand shown in cyan and the product strand in green. The three, single stranded, downstream template nucleotides (+1, +2, +3) are colored as indicated in the top left corner. The clamp formed by the pinky (red) and thumb (blue) domains around the exiting RNA is indicated by the dashed grey arrows. PDB code 3OL6, figure from reference (82)

Briefly, prior to NTP binding, the active site of 3D^{pol} is in the open conformation; once 3D^{pol} binds the NTP, the active site goes from the open to closed position to enable catalysis, and this is followed by reopening of the active site. It is thought that translocation then occurs, although no structure was obtained for a translocation intermediate state.

Further structures obtained for PV, coxsackievirus, and rhinovirus ECs reveal highly conserved 3D^{pol}-RNA interactions at ≈ 10 base positions: two non-paired bases in the downstream template, i.e. +1 and +2 with the +3 base being unstacked from the +2 base by Pro20 and held outside the active site, and $\approx 7-8$ base pairs of the upstream template-product duplex that is held in place via interactions with the thumb and pinky domain (85). Approximately six of these bases are located within the interior of the protein: the +1 and +2 downstream bases and the four newly incorporated nucleotides (-1 to -4, with the -4 position being four positions upstream from the active site) that are duplexed with the template. At this point the duplex begins threading out of the polymerase between a loop in the pinky finger and an α -helix in the thumb domain, which pack into the major and minor groove, respectively, to form an “upstream clamp” on the exiting RNA. This α -helix in the thumb that forms part of the clamp is the topic of Chapter 3 of this dissertation.

1.4.4 3D^{pol} is a Low Fidelity Polymerase

An important characteristic of viral RdRPs is their high rate of nucleotide misincorporation, i.e. low fidelity, with $\approx 10^{-3}$ to 10^{-5} mutations per nucleotide copied, which is $\approx 10^6$ times higher than that seen for DNA replication in cells (86, 87). This high

error rate generates a diverse and quickly evolving group of viral populations referred to as quasispecies (88). This feature allows viruses to quickly adapt to selective environmental pressure, including antiviral drugs (86, 87). Identification of a mutation in 3D^{pol} that increases polymerase fidelity, G64S (89, 90), has provided the basis for researching new vaccine strategies targeting 3D^{pol} fidelity (91, 92).

1.5 Scope of this Dissertation

During the course of my graduate work, I have carried out research of the poliovirus 2C^{ATPase} as well as the poliovirus RdRP 3D^{pol}. In Chapter 2, I describe the characterization of 2C enzymatic activity and how that activity is regulated by the presence of the other PV membrane-binding polyprotein neighbors. Importantly, the proteins used for these studies were expressed using an *in vitro* transcription/translation system in the presence of small lipid bilayers, or nanodiscs, which resulted in highly active 2C ATPase activity compared to previous studies. In Chapter 3, I describe the additional preliminary or inconclusive characterization of some other 2C functions, including attempting to identify 2C helicase activity. Investigation of the role of the 3D^{pol} thumb helix “protein clamp” in mediating poly(A) tail length is the topic of Chapter 4. This chapter is followed by Appendix I, which relays the preliminary *in silico* molecular dynamics simulations I carried out to investigate the structural basis for 3D^{pol} thermal stability and the mechanism by which G64S is able to impart a higher fidelity phenotype.

CHAPTER 2

POLYPROTEIN CONTEXT REGULATES THE ACTIVITY OF POLIOVIRUS 2C^{ATPase}

BOUND TO BILAYER NANODISCS¹

Authors contributing to this chapter: Author Huntoon was an undergraduate researcher in the laboratory and assisted in the cloning, in vitro expression, and batch purification of several of the 2C constructs under my supervision. The experiments, figures, and tables in this chapter were completed by me, Courtney L. Springer.

2.1 Introduction

The picornaviruses are a family of human and mammalian pathogens that include hepatitis A virus, the rhinoviruses, foot-and-mouth disease virus, coxsackie virus, and poliovirus (PV). PV remains a health threat in many parts of the world, where it is being kept in check by extensive immunization efforts aimed at its eventual eradication (93). Poliovirus has a 7.5 kb single-stranded, positive-sense RNA genome that encodes a single ≈250 kDa polypeptide. Upon translation, this viral polyprotein undergoes a series of proteolytic cleavages to generate a total of eleven different proteins plus several important functional intermediates (Fig. 2.1). The four proteins in the P1 region (VP1-VP4) are the structural proteins that make up the virion capsid while

¹ This chapter has been published as **Springer CL, Huntoon HP, Peersen OB**. 2013. Polyprotein Context Regulates the Activity of Poliovirus 2CATPase Bound to Bilayer Nanodiscs. *J. Virol.* **87**:5994-6004.

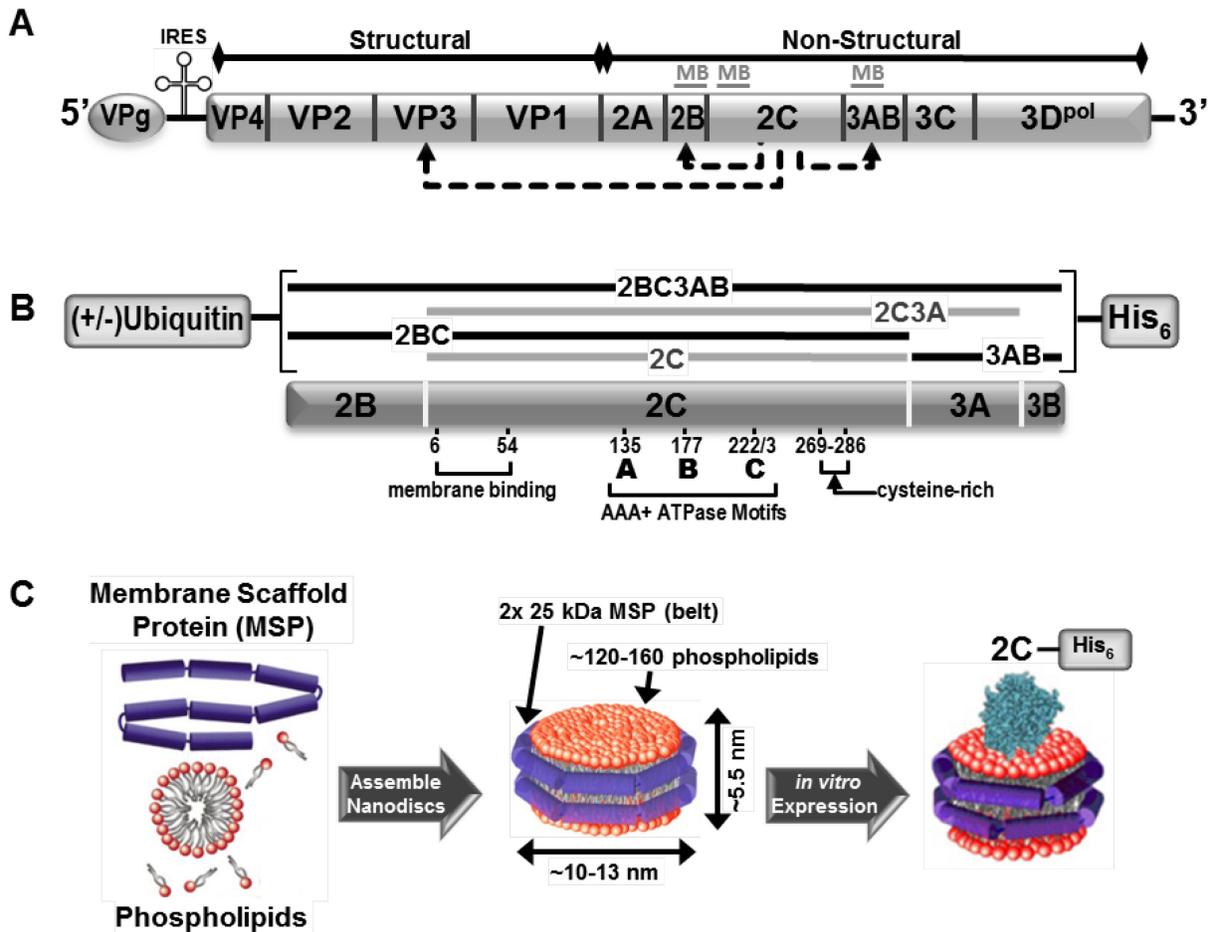


Fig. 2.1. Poliovirus polyprotein processing and nanodisc chaperoned expression.

(A) Organization of the poliovirus single-stranded positive-sense RNA genome encoding a single open reading frame for the viral polyprotein. Dashed arrows indicate viral proteins known to interact with the 2C^{ATPase} and gray bars identify membrane-binding (MB) regions. (B) Schematic representation of the five viral proteins that were generated for the present study. All constructs contained a C-terminal hexahistidine tag and were made both with and without a N-terminal ubiquitin domain (\pm Ub). Diagrammed below are known 2C functional domains, including the NTP interaction regions that comprise the core AAA+ ATPase motifs. (C) Depiction of the nanodisc-based *in vitro* expression protocol showing the initial assembly of discs from pure phospholipids and two copies of the amphipathic helix-based membrane scaffold protein (MSP) that are then added to bacterially based *in vitro* transcription/translation expression reactions (derived from 94, 95).

the remaining seven non-structural proteins in the P2 and P3 regions (2A-2C and 3A-3D, respectively) are involved in viral replication biochemistry (14). Processing of the polyprotein plays an important role in mediating protein function as several viral precursor proteins have different functional roles and biochemical activities than the fully processed proteins. For example, the cleavage of poliovirus 3CD^{pro} generates and activates 3D^{pol}, the RNA-dependent RNA polymerase that carries out replication of the viral RNA (41, 96, 97). Similarly, the 2C and precursor 2BC proteins act in concert to trigger a massive rearrangement of intracellular membranes that results in the formation of extensive “membranous web” structures (36, 98, 99), on the surfaces of which these proteins form the viral replication centers that are the sites of RNA synthesis and virion assembly.

The viral 2B, 2C, and 3A proteins all contain membrane-binding regions. Both 2BC and 2C are able to induce membrane rearrangement in living cells (37, 38, 51); however, co-expression of a combination of these membrane binding proteins is needed to recapitulate the double walled membranes observed during virus infection (52). 2B contains a pair of predicted integral membrane helices and is thought to form a viroporin structure (100-102), 2C is anchored to the membrane via a N-terminal amphipathic helix (38, 67, 103, 104), and 3A contains an amphipathic sequence within its C-terminal 22 residues (105, 106). There is evidence that these three membrane-anchored proteins interact with one another both *in vivo* and *in vitro*; substitution of the 2C amphipathic helix from other picornaviruses into a poliovirus background results in compensatory mutations in the membrane anchoring sequences of both 2B and 3A (39, 107). In addition to its roles in membrane rearrangements, 2C has also been implicated

in initiation of negative-strand RNA synthesis (62), interacting with the VP3 capsid protein (65), and virion packaging (63).

2C is comprised of 329 amino acids and contains an N-terminal amphipathic helix, two predicted RNA binding regions, three motifs implicated in NTP binding, and a cysteine rich putative zinc finger motif (Fig. 2.1B) (61, 67, 68). Two of the NTP binding motifs are Walker A and B, common to AAA+ ATPases, while the third is motif C (also referred to as Sensor-1) that shares homology with the superfamily III helicases (69). 2C has been shown to have ATPase activity and may also have low level GTPase activity (68-70). Electron microscopy images of 2C proteins from poliovirus and foot-and-mouth disease virus (FMDV) show hexameric ring structures and activity assays indicate that oligomerization of FMDV 2C is important for ATP hydrolysis (75, 76). There are no atomic level structures for the picornaviral 2C proteins and sequence homology searches do not reveal convincing matches for proteins other than those from closely related viruses. Structure prediction using the *Phyre* engine (73) that utilizes profile-profile sequence matching and structure based threading to do homology modeling shows a high probability that residues \approx 100-260 fold into a AAA+ family ATPase. The AAA+ (ATPases associated with diverse cellular activities) superfamily of enzymes are often considered protein workhorses because they are involved in a multitude of basic cellular functions, including protein degradation, membrane rearrangement, DNA replication, and disassembly of protein complexes (72, 74). These enzymes tend to oligomerize (74), forming structures that are consistent with the aforementioned electron microscopy data showing 2C ring assemblies.

The propensity of 2C to associate with membranes has made it difficult to study and much remains unknown about both its structure and its function during viral replication. Thus far, poliovirus 2C has been generated using GST (69) or maltose binding protein (75) fusions to improve expression and solubility, but those proteins remained difficult to work with. FMDV 2C has been purified in a soluble form by deleting the N-terminal membrane binding region to generate a soluble and biochemically active core ATPase fragment (76). Attempts by our laboratory to express poliovirus 2C using similar approaches in both *E. coli* and baculovirus systems resulted in protein that was largely insoluble and had little or no ATPase activity. Additionally, our GST-2C fusion construct could not be efficiently cleaved using PreScission protease, a derivative of the picornaviral rhinovirus 3C protease, suggesting the protein may not have been natively folded.

To circumvent these difficulties with expression and purification, we explored a new approach for isolating 2C by expressing it *in vitro* on small lipid bilayers known as nanodiscs (Fig. 2.1C). Developed by the Sligar laboratory, nanodiscs are soluble nanolipoprotein particles comprised of a lipid bilayer held together by an amphipathic helix membrane scaffold protein (MSP) that forms a belt around the lipid acyl chains (108, 109). Nanodiscs provide a native bilayer surface to which membrane proteins can adhere, thus reducing the likelihood of aggregation, while being small enough to remain fully soluble and tumble rapidly enough to enable a variety of spectroscopy experiments (94, 95, 110). Most importantly, this system has allowed us to biochemically characterize 2C activity when bound to true membrane bilayers instead of detergent

micelles. Additionally, this allowed us to explore 2C activity in the context of its membrane-associated polyprotein neighbors to examine how processing may affect 2C function.

In this work we demonstrate the *in vitro* expression and subsequent purification of poliovirus 2C, 2BC, 2C3A, 2BC3AB, and 3AB proteins on nanodiscs composed of dimyristoyl phospholipid bilayers with choline, glycine, and serine headgroups. Using a NADH coupled ATPase assay, we show that the 2C^{ATPase} is active over a broad range of pH values and its activity is stimulated by the presence of acetate while being slightly inhibited by monovalent salts. Notably, we observe a ≈60-fold activation of 2C ATPase activity by the fully processed 2C as compared to the 2BC3AB polyprotein precursor, suggesting that proteolytic processing of the viral polyprotein plays a major role in activating 2C.

2.2 Materials and Methods

2.2.1 Cloning

Constructs were cloned from pUC57-2BC3AB, a bacterial codon optimized synthetic gene made by Genscript Inc. (www.genscript.com) encoding for poliovirus 2BC3AB followed by a C-terminal GSSS-His₆ tag sequence. The construct includes a SacII restriction site immediately preceding the start codon and a NotI site following the His₆ sequence. Initially, the 2C-containing constructs were cloned into a pET26 expression vector developed by the Cameron laboratory for expression of poliovirus 3D^{pol} (111). With this vector the protein is initially translated as a fusion with an N-terminal ubiquitin domain (Ub), but this domain is cleaved off by the ubiquitin-specific

carboxyl-terminal protease Ubp1 that is co-expressed in the *E. coli* pCG1 strain, resulting in a processed polymerase with a N-terminal glycine residue that mimics the natural cleavage of the poliovirus polyprotein at Gln-Gly sites. The 3D^{pol} gene was removed from this vector and replaced with the gene for 2BC3AB-GSSS-His₆ via the SacII and NotI sites, resulting in pET26-Ub-2BC3AB-GSSS-His₆. 5'- and 3'- deletions within the open reading frame were then used to generate expression vectors for Ub2BC, Ub2C, Ub2C3A, and Ub3AB. A second set of these expression vectors where the N-terminal ubiquitin fusion was replaced by a single methionine start codon were also generated, and the integrity of all constructs was verified by sequencing.

For the *in vitro* translation experiments we found that the pET26-based constructs resulted in very low protein yield as compared to the Elongation Factor Ts produced from the pEFT control plasmid provided in the EasyXpress kit (Qiagen Inc.). Further analysis showed this to be the result of inefficient RNA transcription, which can have a major effect on *E. coli* based *in vitro translation* systems due to high endogenous RNase activity. To increase RNA levels in the reactions, we cloned all our constructs into the pEFT control vector by simply replacing the elongation factor Ts open reading frame. The subsequent mutations to the predicted AAA+ ATPase NTP binding motifs in 2C were made by the Quikchange mutagenesis protocol (Stratagene) to produce K135A (Walker A), D177A (Walker B), and T222A and N223A (Motif C/Sensor-1) in both the pEFT-Ub2C and pEFT-Ub2BC vectors.

2.2.2 Nanodisc Assembly

Nanodiscs were expressed and assembled with the MSP1E3D1 protein (112) and DMPC lipids (unless noted) using 1:150 molar ratio of MSP1E3D1:DMPC as per the protocols provided by the Sligar laboratory (sligarlab.life.uiuc.edu). Briefly, MSP1E3D1 was expressed from the vector distributed by Addgene as plasmid number 20066 (www.addgene.org/20066) and purified by nickel affinity chromatography. The N-terminal His₆ tag on the purified MSPs was removed using TEV protease that also contained a His₆ tag. After cleavage, the TEV/MSP mixture was run over a second nickel affinity column, from which the only the MSPs lacking the His₆ tag were collected in the unbound fraction. MSP and DMPC were then co-solubilized using cholate detergent, cholate was removed using hydrophobic gel-permeation BioBeads (Biorad, Inc.) to initiate nanodisc assembly, and the resulting discs were purified over a Superdex S-200 gel filtration column where elution also confirmed that the nanodiscs were of the expected size with a Stokes diameter of ≈ 12 nm (113). Nanodiscs containing DMPG or DMPS phospholipids were assembled using the same MSP protein and assembly protocols, with the difference that lipid solubilization and assembly were carried out just above the appropriate lipid transition temperatures of 23°C for DMPC and DMPG and 35°C for DMPS. We attempted to assemble nanodiscs using DMPE, but they did not remain soluble when brought below their ≈ 50 °C transition temperature. All lipids were purchased from Avanti Polar Lipids (www.avantilipids.com).

2.2.3 *In vitro* Expression and Purification

$2C^{ATPase}$ and the other 2BC3AB-based proteins were produced using the Invitrogen Expressway *in vitro* bacterial expression system. 0.1-2 mL protein expression reactions were incubated in an Eppendorf Thermomixer at 30°C and 1200 rpm for three hours using 1 µg of DNA per 100 µL final reaction volume (RV) and nanodiscs were added to a final concentration of 15 µM. Half of the required small molecule “feed mix” was added 30 minutes after starting the reaction and the other half was added after two hours. The completed reactions were then centrifuged for 10 minutes at room temperature. The supernatant was batch purified with one-quarter RV of settled Ni-NTA resin (Qiagen) pre-equilibrated in low imidazole buffer (50 mM Tris, pH 8.0, 300 mM NaCl, 10 mM Imidazole, 0.02% NaN₃, and 10-20% glycerol) by incubation at 4°C for at least 30 min. After mixing, the supernatant was removed and the resin washed with three RVs of low imidazole buffer. Protein was batch eluted in multiple sequential quarter-RV fractions of high-imidazole (500 mM) buffer. TCEP reducing agent (Pierce) was added to each elution to a final concentration of 10 mM, samples were analyzed by SDS-PAGE, and elution fractions containing significant amounts of protein were combined, aliquoted, and stored at -80°C for future use. Note that since the His₆ tag had previously been cleaved from the MSPs, only nanodiscs bound to the 2C constructs are co-purified in the elution fractions and this represents a small portion of the 15 µM nanodiscs that were added to the translation reaction.

Several constructs were later subjected to more rigorous rounds of purification in order to reduce the amount of contaminant proteins that co-eluted with 2C during the

batch purification process. In this purification process, 1-2 mL (final RV) of the Expressway reaction were purified by liquid chromatography using an AKTA Purifier System first over a 1 mL HisTrap HP Nickel resin column (GE Life Sciences) in the same low and high imidazole buffers used for batch purification. Peak fractions were pooled, diluted 5-fold and further purified over a 1 mL HiTrap Q HP ion exchange column (GE Life Sciences) with a gradient of low to high salt Q buffer (25 mM Tris, pH 8.5, 50 mM/1 M NaCl or KCl, 20% glycerol). 2C protein typically eluted in \approx 300 mM salt and was then stored as stated above.

2.2.4 Protein Yield and Concentration Determination

Determination of the specific activity of each ATPase construct requires knowing the precise concentration of ATPase in the purified samples. Since the purified samples contained the 2C construct, the nanodisc MSP, and some contaminant proteins from the batch-based purification scheme, we could not accurately determine the 2C concentration by simple 280 nm absorbance or other overall protein content assays. We therefore determined the concentrations of both the viral proteins and MSP by separating them on 12-15% SDS-PAGE gels alongside known concentrations of the ProSieve Unstained Protein Marker II standards (Lonza Group, Ltd.). The gel was stained with IRDye Blue Protein Stain (LI-COR Bioscience) and imaged on an Odyssey Infrared Imaging System at 700 nm. The intensities of the protein standard bands were quantified and plotted against total amount of protein in each band to yield a linear relationship between staining intensity and protein content (Fig. 2.2). The protein

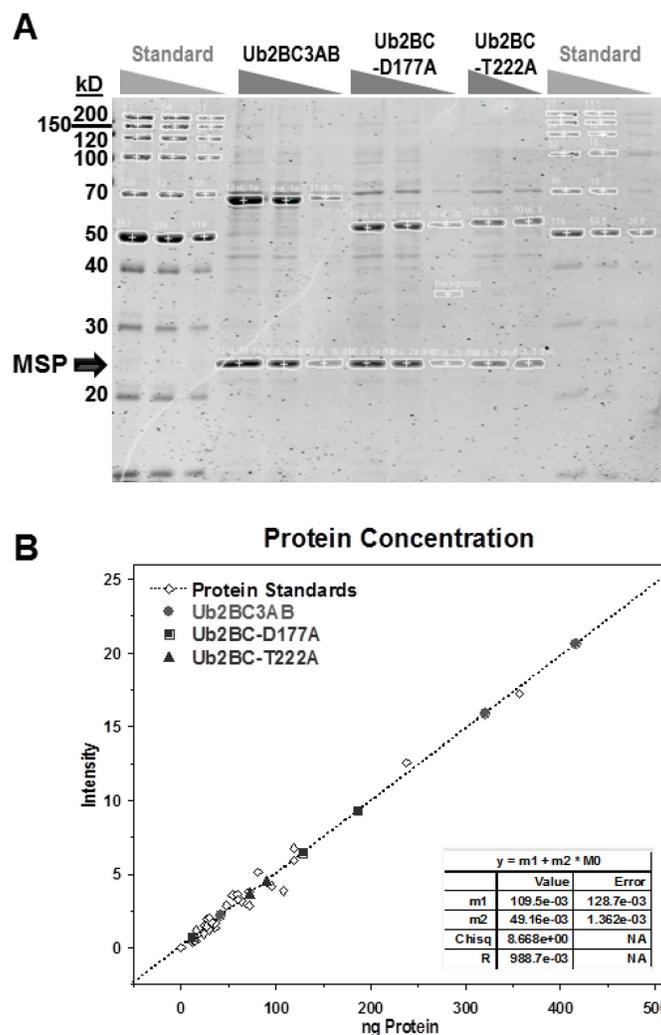


Fig. 2.2. Protein concentration determination.

(A) Protein concentrations were determined by running various dilutions of the ProSieve Unstained Protein Marker II standard alongside dilutions of 2C protein constructs on reducing acrylamide SDS gels. The gels were stained with LI-COR IRDye blue protein stain, imaged by scanning on a LI-COR Odyssey infrared fluorescence imaging system, and multiple bands from the standards, the 2C constructs, and nanodisc MSP were outlined for quantitation as shown. (B) Example standard curve generated from the correlation of the protein standards intensities (open circles) with the amount of protein loaded and then fit to a linear curve whose parameters are listed in the inset box. The concentrations of the various 2C construct dilutions were then interpolated on this line, as per the three proteins shown here (solid symbols), and averaged to obtain the concentration in the original sample. Independent gels with standards were analyzed for every purified protein, allowing us to calculate specific activities in the ATPase assays based on the amount of 2C protein added.

concentration of the in vitro expressed His-tagged viral proteins and the nanodisc MSPs that copurified with them were then obtained by interpolating their band intensities on this standard curve.

2.2.5 ATPase Activity Assay

The rate of ATP hydrolysis was determined using a ATP/NADH coupled assay (Fig. 2.3A) combined with an ATP regeneration system that allows us to assay steady state $2C^{\text{ATPase}}$ activity over a wide range of ATP concentrations (114). In this assay the conversion of $\text{ATP} \rightarrow \text{ADP} + \text{P}_i$ by an ATPase is coupled to ATP regeneration via the pyruvate kinase (PK) catalyzed conversion of phosphoenolpyruvate ($\text{PEP} + \text{ADP} \rightarrow \text{pyruvate} + \text{ATP}$) and the lactate dehydrogenase (LDH) catalyzed conversion of pyruvate + NADH \rightarrow lactate + NAD^+ . The rate of ATP hydrolysis is then monitored by the loss of NADH absorbance at 340 nm (Fig. 2.3B) with a molar extinction coefficient of $6.22 \text{ mM}^{-1} \text{ cm}^{-1}$. Coupled reactions consisted of 5 mM PEP, 0.024 U/ μL PK, 0.036 U/ μL linearly reduced to ≈ 0.25 upon complete oxidation of the NADH as a result of ATPase turnover.

The ATP hydrolysis assays were carried out at 37°C in a Varian Bio50 spectrophotometer with a thermostatted multi-cell holder using 0.5-40 nM $2C^{\text{ATPase}}$ domain in a base buffer containing 75 mM KCl, 50 mM magnesium acetate (MgOAc), 20 mM HEPES (pH 7.0), 400 μM NADH, and 5 mM DTT. Steady state data were collected for 30-60 minutes with points measured every ≈ 30 -45 seconds, depending on how many assays were being conducted in the 18-cell multicell holder at the time. Experiments to characterize the monovalent salt, pH, and magnesium dependence of

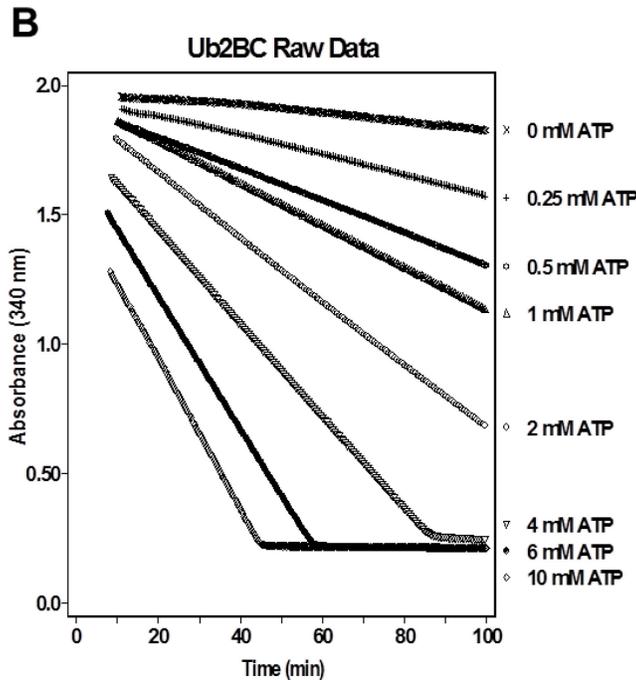
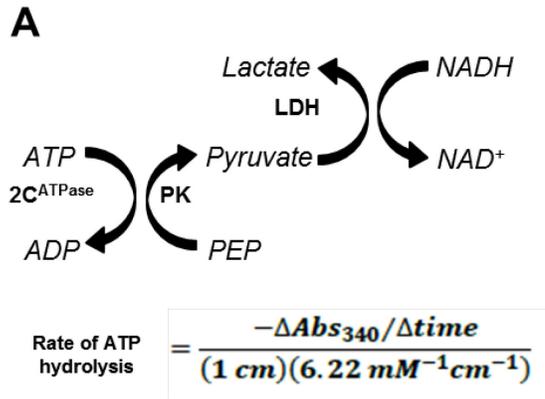


Fig. 2.3. Coupled ATP/NADH ATPase assays.

(A) Reaction scheme of ATPase assay in which ATP hydrolysis is coupled to oxidation of NADH via pyruvate kinase (PK) and lactate dehydrogenase (LDH). Loss of absorbance at 340 nm reflects the loss of NADH due to ATP hydrolysis. (B) Raw absorbance data for a set of Ub2BC assays with various ATP concentrations that were performed simultaneously using a Cary 50 Bio spectrophotometer with a thermostatted (37°C) multicell holder.

2C^{ATPase} activity were done by varying individual components of this buffer as needed. To verify that the buffer conditions tested did not adversely affect the coupled reporter system assay, 0.2 mM ADP was directly added to control reactions and the subsequent loss of NADH absorbance was instantaneous and always significantly faster than the rate of ADP production by 2C.

2.3 Results

2.3.1 Lipid Dependence of Nanodisc Based *in vitro* Expression

Nanodiscs were critical for viral protein synthesis and we could not detect poliovirus proteins from expression reactions done in the absence of nanodiscs (Fig. 2.4). We expressed both Ub2BC and Ub2C in the presence of nanodiscs made with dimyristoyl phospholipids attached to choline (DMPC), serine (DMPS) or glycerol (DMPG) head groups and found there were marked differences in the expression levels of the different proteins in the presence of the different lipids. Ub2BC expressed the best with DMPC and DMPS nanodiscs, and expression levels were reduced about 2-fold in the presence of DMPG nanodiscs (Fig. 2.4A). Alternatively, the best Ub2C expression was seen with DMPG discs and yield was reduced \approx 2-3-fold when in the presence of DMPS or DMPC (Fig. 2.4B). Activity assays showed that the headgroup type had up to \approx 3-fold effects on the ATPase k_{cat} values and the lipid giving maximal activity varied depending on the polyprotein being assayed; Ub2C shows the highest activity with DMPC, Ub2BC prefers DMPG, and both proteins exhibit half of their respective maximal activities with DMPS nanodiscs (Fig. 2.4C). Ultimately, we decided to use DMPC nanodiscs for further experiments as Ub2C demonstrated the highest

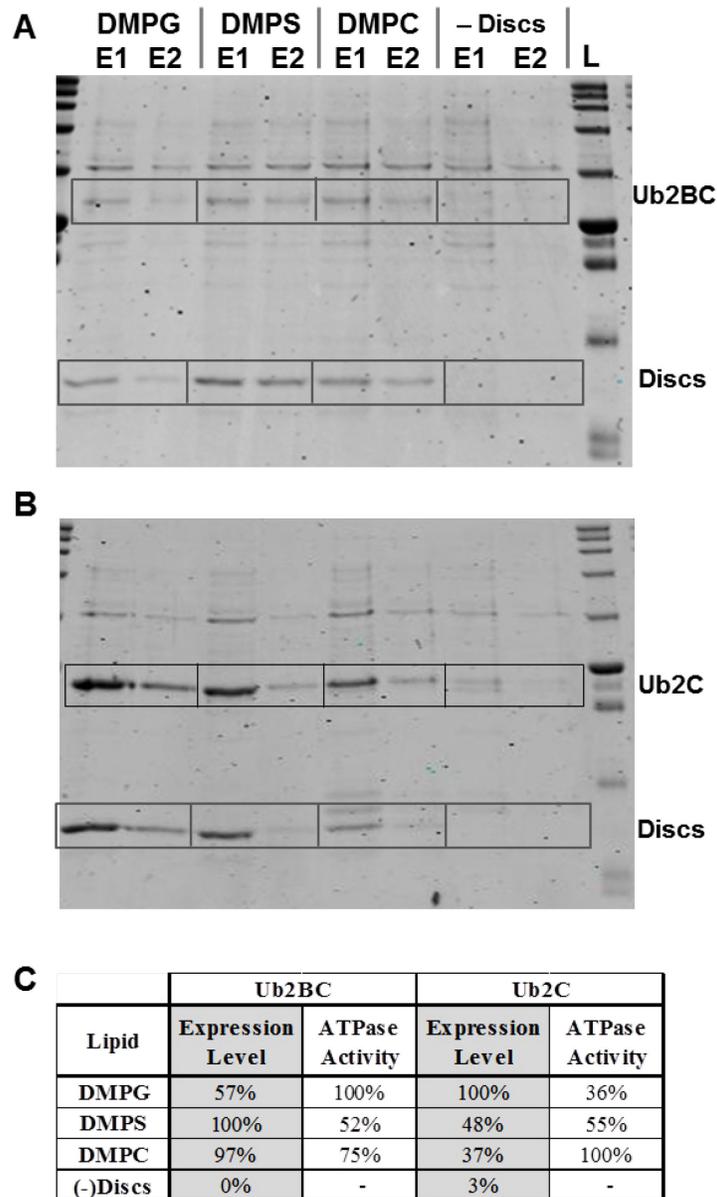


Fig. 2.4. Phospholipid dependence of Ub2C and Ub2BC expression and activity.

(A and B) Batch nickel purification elution fractions E1 and E2 from 100 μ L in vitro expression of Ub2BC (A) and Ub2C (B) in the presence of nanodiscs assembled with DMPG, DMPS or DMPC lipids or in the absence of nanodiscs entirely. Lane L contains ProSieve Unstained Protein Markers II. (C) Relative expression levels and ATPase activities obtained with the different nanodiscs for both Ub2C and Ub2BC proteins. ATPase assays were carried out at 37°C in the presence of 1 mM ATP.

ATPase activity on these discs and Ub2BC activity was only reduced by $\approx 25\%$ on DMPC discs as compared to its maximal activity observed on DMPG discs.

1.1.1 Expression of 2C Constructs

In the presence of DMPC nanodiscs, the average final batch purification yield for the various protein constructs was $\approx 11 \mu\text{g}$ per mL of translation reaction, with the different constructs varying from 1–40 $\mu\text{g}/\text{mL}$ in their yields. The lowest yield was observed for 2C alone and this protein had a strong tendency to aggregate into a pelleting fraction, but yield was improved by including polyprotein neighbor proteins, i.e. 2B and 3A (Fig. 2.5A). N-terminal Ub fusion proteins generally resulted in higher expression levels and better yield of soluble protein (Fig. 2.5B). Indeed, the soluble protein yield was typically lowest for 2C at $\approx 1 \mu\text{g}$ per mL of in vitro translation reaction, but this was increased ≈ 5 -fold for Ub2C, and even further by the addition of viral polyprotein neighbors, resulting in up to $\approx 40 \mu\text{g}$ per mL of reaction mixture for Ub2BC3AB. Thus, we primarily used the Ub fusion constructs for our ATPase activity assays. We also made alanine point mutations to the AAA⁺ ATPase motifs predicted to exist within the 2C domain: Walker A (K135A), Walker B (D177A), and motif C (T222A, N223A). These mutants demonstrated a 2-30 fold increase in protein yield when compared to the wild type, with the most significant increase seen for the Walker A mutant. Furthermore, Ub3AB expressed at higher levels than any of the 2C-containing constructs, resulting in up to 60 μg of protein per mL of in vitro expression reaction.

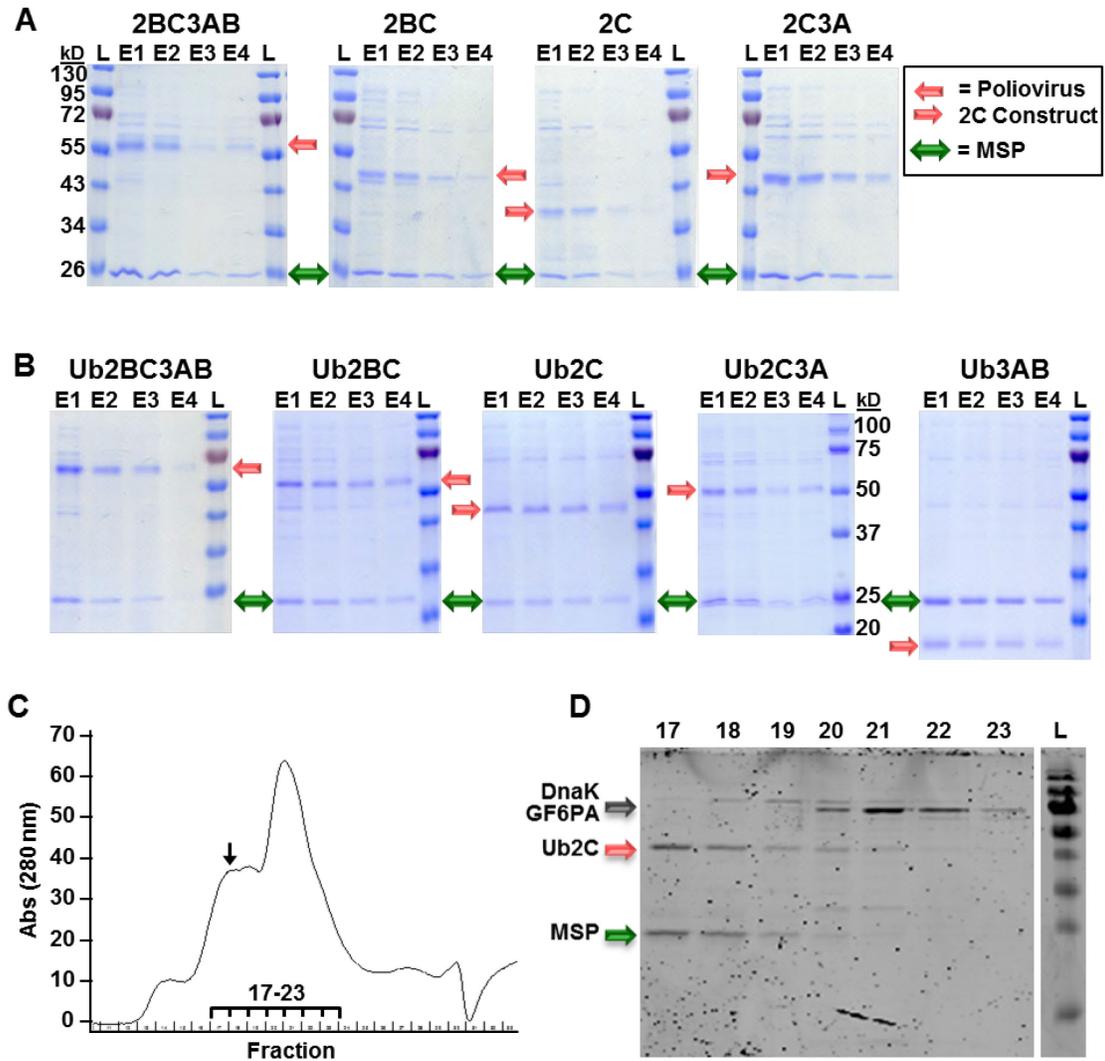


Fig. 2.5. Purification of *in vitro* translated 2C proteins.

(A and B) Batch nickel resin elution fractions (E1-E4) from DMPC nanodisc chaperoned *in vitro* expression of wild-type 2C constructs (A) and N-terminal ubiquitin fusion constructs (B). The bands for the desired 2C protein are marked by single-sided arrows, and the nanodisc MSP protein is denoted by double-sided arrows. Constructs were expressed in 100 μ L reactions, purified in batch with Ni^{2+} affinity resin, and 5 μ L samples of each 25 μ L elution fraction were analyzed on 12% acrylamide-SDS gels. The ladder molecular weight shown beside the 2BC3AB is the same for all gels except for Ub2C3A. (C and D) FPLC purification of Ub2C via HisTrap nickel affinity (not shown) and HiTrap Q anion exchange columns. The primary 2C elution position is shown by the arrow on the HiTrap Q chromatogram (C), and SDS-PAGE analysis of fractions 17 to 23 revealed two major proteins eluting immediately after 2C (D). Mass spectrometry was used to identify these proteins as bacterial DnaK and glucosamine fructose-6-phosphate aminotransferase (GF6PA).

To reduce the probability of having multiple viral proteins associated with a single disc, the in vitro translation reactions were carried out in the presence of a large excess nanodiscs (15 μ M), only 2-5% of the which were recovered by co-purification with the expressed hexahistidine tagged 2C constructs. Quantitation of the purified protein bands from multiple expressions showed overall mole ratios of one to two 2C constructs per pair of MSPs needed for each nanodisc. The protein-nanodisc complexes were soluble and little sample loss was observed upon centrifugation at 4°C, except for the 2C protein that was prone to precipitation. This was exacerbated in the FPLC purified samples that had a clear tendency to aggregate upon storage, but was improved somewhat by the addition of the N-terminal ubiquitin fusion. Note that we were able to stoichiometrically cleave the Ub domain from Ub2C using exogenously added Ubp2 protease (generously provided by Dr. Robert Cohen), consistent with the fusion protein being properly folded (data not shown). This is in contrast to our earlier attempts to express 2C as a GST fusion in insect cells where the resulting protein was not very soluble and could not be cleaved from GST, suggesting those proteins may have been misfolded or aggregated.

Initially, all the 2C constructs were only batch purified using nickel affinity resin; however, we consistently noticed a \approx 70 KDa contaminant protein co-eluting with the purified protein. This band was excised from a SDS-PAGE gel, analyzed by liquid chromatography tandem mass spectrometry (LC-MS/MS) at the Proteomics and Metabolomics Facility at Colorado State University, and identified as the bacterial chaperone protein DnaK (Hsp70), a known protein folding chaperone and ATPase (115). Concerned that DnaK could be contributing to the observed ATPase activity in

our assays, we subjected several constructs to a more rigorous FPLC-based purification scheme that included both a nickel affinity column and an ion exchange column (Fig. 2.5C, D). Ub2C, Ub2BC, Ub2BC3AB, Ub3AB, and Ub2C-K135A were all purified in this manner and used for the final Michaelis-Menten kinetic analyses (Fig. 2.7B, Table 2.2). Batch purified proteins were used to determine ATPase assay conditions and for initial kinetic comparisons (Fig. 2.7A, C, D; Table 2.2)

2.3.2 ATPase Assay

Enzymatic activity was determined using a coupled assay in which the rate of ATP hydrolysis is linked to a stoichiometric conversion of NADH to NAD⁺ (Fig. 2.3A). Fig. 2.3B shows the time dependence of NADH depletion measured by absorbance at 340 nm in an assay containing Ub2BC with ATP concentrations ranging from 0 to 10 mM. The lowest ATP concentration tested in this experiment, 0.25 mM, results in a NADH oxidation rate that is about twice that of background, and the rates then increase with increasing ATP concentration. At 4 mM ATP and above the total amount of ATP hydrolyzed in 30-60 minutes often exceeded the 0.4 mM NADH added to the assay and the absorbance data therefore reached a minimal signal plateau at ≈ 0.25 absorbance units. The raw data curves do not have the same Y-intercept because there is contaminating ADP in the ATP solution that is immediately converted back to ATP at the expense of NADH upon assay assembly. The concentration of this contaminating ADP increases along with that of the ATP, resulting in lower starting absorbance values at higher ATP concentrations, as shown in Fig. 2.3B.

2.3.3 Salt, pH, and Temperature Dependence

To determine the optimal conditions for 2C ATPase activity assays we examined the rate of ATP hydrolysis by batch-purified Ub2C under different monovalent salt (KCl and NaCl), magnesium, acetate, pH, and temperature conditions (Fig. 2.6A-D). To verify that the temperature and pH conditions being tested were not affecting the enzymatic activity of the PK and LDH components of the coupled assay, ADP was added to a final concentration of 0.2 mM, corresponding to half of the starting NADH concentration. In every case, the loss of NADH absorbance at 340 nm in the control reactions was immediate, indicating that turnover within the reporter system was not rate limiting.

The activity of Ub2C bound to nanodiscs showed a gradual but weak monovalent salt dependence with a fairly flat plateau between 25 and 100 mM KCl and a drop in activity at salt concentrations higher than that (Fig. 2.6A). The effect was small in the presence of KCl, with activity only reduced by $\approx 10\%$ in the presence of 200 mM salt, while NaCl exhibited a greater inhibitory effect with activity decreasing linearly with salt concentration to reach $\approx 60\%$ activity in 200 mM NaCl. We opted to use 75 mM KCl as a standard condition for further ATPase assays as it demonstrated only a $\approx 5\%$ decrease in activity and is fairly close to physiological salt concentrations. We examined the pH dependence of batch purified Ub2C activity at eleven different values and found a fairly flat activity profile from pH 6 to 7 followed by a decrease in activity as the pH was increased to 9 (Fig. 2.6B).

The activity of membrane proteins is often affected by the fluidity of the bilayer and can therefore be influenced by lipid phase transitions, and conversely the presence

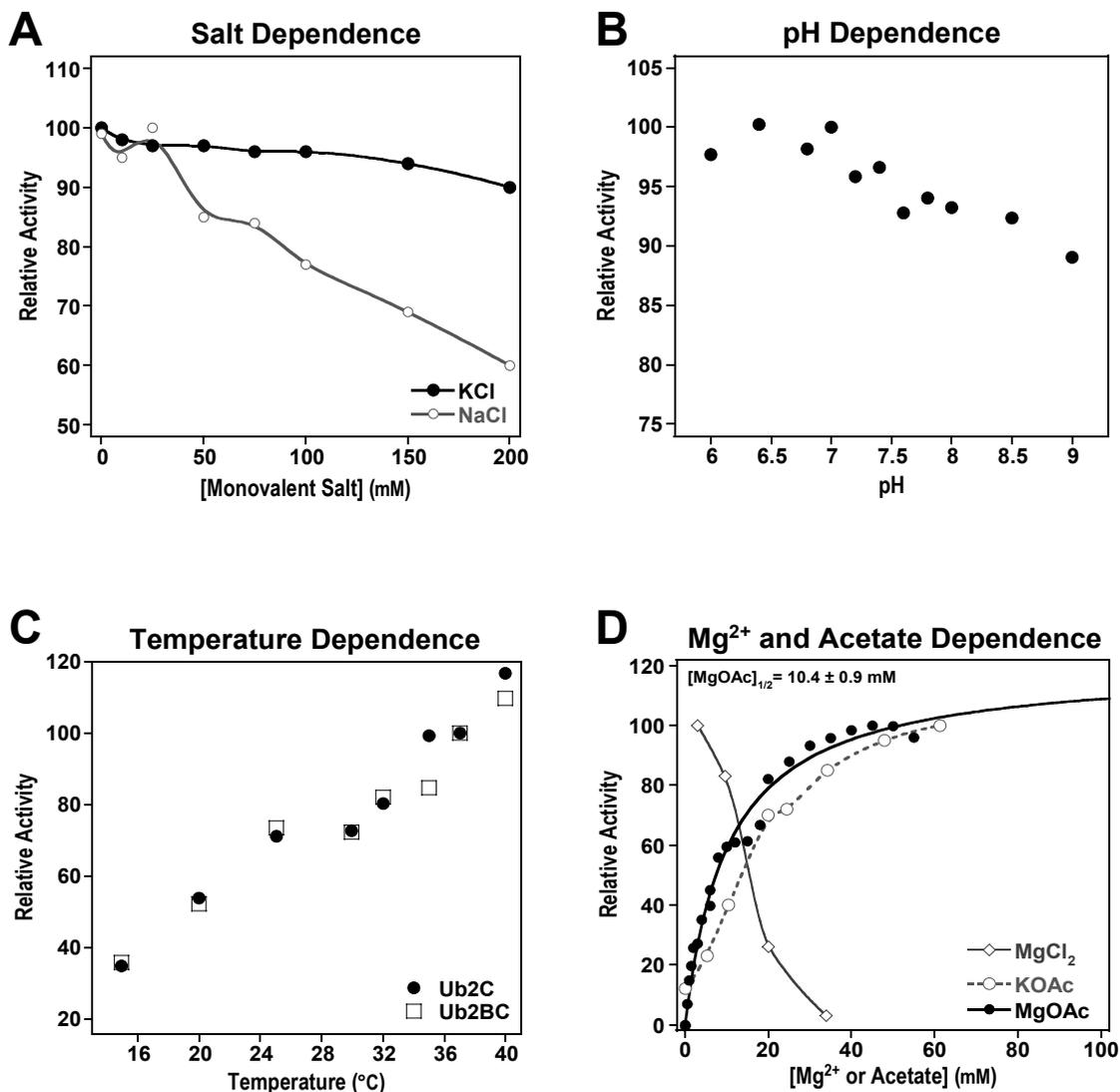


Fig. 2.6. ATPase activity of 2C proteins bound to nanodisc bilayers.

(A) Monovalent salt concentration dependence of Ub2C ATPase activity in the presence of NaCl or KCl. Reactions were carried out with 3 mM ATP and 1 nM batch purified Ub2C protein and normalized to the activity in the absence of added salt. (B) pH dependence of Ub2C ATPase activity examined with 2 mM ATP and 5 nM batch purified Ub2C, normalized to activity at pH 7. (C) Temperature dependence of Ub2C and Ub2BC ATPase activity carried out with 2 mM ATP and 5 nM batch purified protein. (D) Magnesium and acetate dependence of Ub2C ATPase activity. The magnesium acetate data are a combination of three sets of assays that were carried out with 1 mM ATP and 2-4 nM batch purified Ub2C. MgCl₂ and potassium acetate assays started with 75 μ L reactions containing 2 mM ATP and 3 mM MgCl₂ into which either MgCl₂ or potassium acetate were titrated with the total added volume added not exceeding 5 μ L.

of proteins can affect phase transition temperatures. The normal phase transition temperature of DMPC is $\approx 23^{\circ}\text{C}$, but it is increased by $3\text{--}4^{\circ}\text{C}$ in the context of a nanodisc (116). To examine if a lipid phase transition near 27°C affected 2C function we measured ATPase activity at a range of temperatures between 15 and 40°C . This was done using both Ub2C that is anchored to the membrane via an amphipathic surface helix and Ub2BC that is a more integral membrane protein due to the addition of a pair of predicted integral membrane segments. The data show an essentially linear temperature dependent increase in activity for both proteins with a slight inflection point in the vicinity of 30°C that may correspond to a weak effect from the DMPC phase transition of a nanodisc bilayer with a bound 2BC or 2C protein (Fig. 2.6C). The ≈ 3 -fold increase in ATPase activity between $15\text{--}40^{\circ}\text{C}$ is in contrast with prior studies of GST-2C that have shown a drastic loss of $\approx 25\%$ activity between 37°C and $\approx 40^{\circ}\text{C}$ (69). To ensure that the rates were not reflecting assay reagent instability (e.g. PEP, NADH, ATP), the temperature dependence experiments were repeated without the addition of Ub2C and these background rates were subtracted from the observed Ub2C rates to yield the data shown in Fig. 2.6C (correction was 5-10% of observed signal).

2.3.4 Magnesium and Acetate Dependence

The presence of magnesium acetate had a substantial stimulatory effect on the batch-purified Ub2C ATPase activity and this effect was further attributed to the addition of acetate. We initially carried out experiments using MgOAc in a fairly standard manner of maintaining the Mg^{2+} concentration 1 mM higher than the ATP concentration, but discovered that we could not obtain a reliable V_{max} for ATP hydrolysis without going

to unusually high ATP concentrations in the millimolar range. We then independently examined the magnesium and the acetate dependence of ATP turnover by titrating magnesium acetate, magnesium chloride, or potassium acetate in the reaction. The results show a clear stimulation by both magnesium and potassium acetate with activation midpoints of 10-15 mM and saturation at \approx 50 mM concentrations; in contrast, magnesium delivered as MgCl_2 strongly inhibited activity (Fig. 2.6D). We also verified that the magnesium acetate activation was not due to low level contamination by another divalent metal, such as Zn^{2+} , that could perhaps be binding to the Cys-rich zinc finger-like motif in 2C. To test for this possibility we added a low concentration of EDTA (1 mM) that would preferentially chelate Zn^{2+} or other divalent ions even in the presence of 50 mM Mg^{2+} (117, 118) and observed that this did not significantly change the ATPase rates (data not shown). High concentrations of magnesium chloride, magnesium acetate, or potassium acetate did not affect the response from the coupled assay reporter system.

2.3.5 Michaelis-Menten Kinetics of 2C Constructs

To examine the extent to which proteolytic processing modulates the ATPase activity of the core AAA+ domain within 2C, we determined the Michaelis-Menten enzyme kinetic parameters of ubiquitin fusion proteins for 2BC, 2C, 2C3A, 2BC3AB and 3AB and the non-ubiquitin versions of 2BC and 2C (Fig. 2.7A). We also made a series of mutations to the putative AAA+ ATPase and analyzed their effects on ATPase activity in the context of both Ub2C and Ub2BC (Fig. 2.7C, D). Based on the characterization experiments described above, these assays were carried out using the coupled

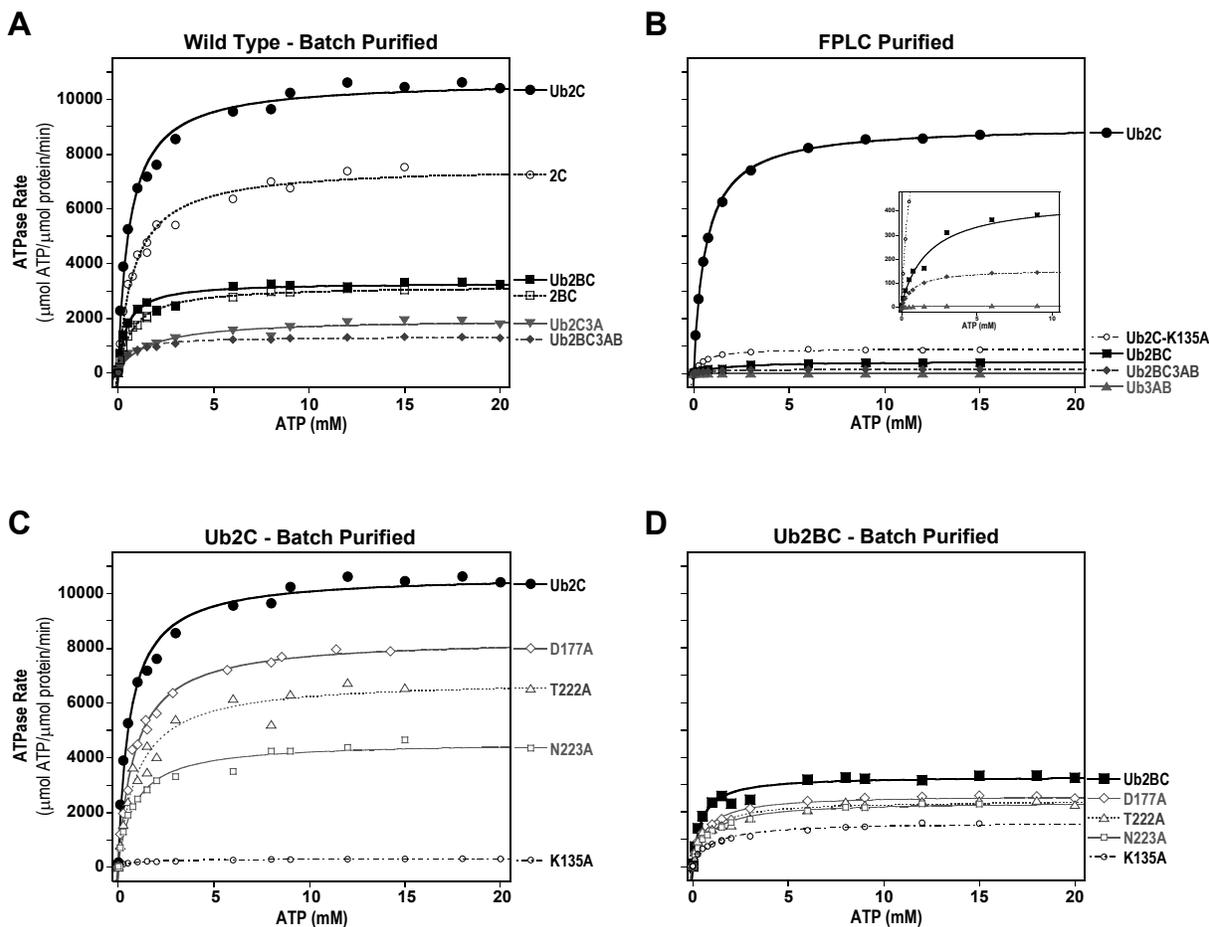


Fig. 2.7. Michaelis-Menten analysis of nanodisc-bound 2C ATPase activity.

All data were obtained under standardized conditions of 37°C with 20 mM HEPES (pH 7.0), 75 mM KCl, 50 mM magnesium acetate, and 5mMDTT. ATP hydrolysis rates were converted to specific activities using gel-based quantitation of protein concentrations and plotted with a uniform y axis scale for ease of comparison. (A) Data from batch-purified proteins with the wild-type polyprotein sequences. (B) Data from FPLC-purified ubiquitin fusion constructs of 2C, 2BC, 2BC3AB, and 2C-K135A, with an expanded-view inset showing details of the data for 2BC, 2BC3AB, and 3AB proteins. (C and D) Data from batch-purified Ub2C (C) and Ub2BC (D) proteins with mutations in AAA⁺ ATPase Walker A (K135A), Walker B (D177A), and motif C (T222A, N223A) sequences. The K_m and k_{cat} values resulting from the curve fits shown here for batch- and FPLC-purified proteins are listed in Tables 2.1 and 2.2, respectively.

ATP/NADH assay containing 0.5–6 nM 2C protein, 0.1–20 mM ATP, 50 mM MgOAc, and 75 mM KCl at pH 7.0 and 37°C. The observed ATP hydrolysis rates were converted to specific activities ($\mu\text{mol ATP}/\mu\text{mol protein/minute}$, hereafter referred to as min^{-1}) by normalizing to the protein concentrations determined by quantitative measurements of band intensities on SDS-PAGE gels (Fig. 2.2).

Initial ATPase assays were carried out using batch purified protein; however, these samples contained the contaminant protein ATPase DnaK, a bacterial chaperone. DnaK has a reported K_m of $\approx 1 \mu\text{M}$ and a k_{cat} ranging from $0.03\text{--}3 \text{ min}^{-1}$ depending on the type of peptide substrate and whether cofactor proteins are present (115). Since our ATPase assays were carried out at 0.1–15 mM ATP concentrations that are well above the K_m for DnaK, we chose to semi-quantitatively background correct the data by using a floating offset for the Michaelis-Menten curve fits, thus making the assumption that this offset represents a V_{max} rate for any contaminant DnaK in the reaction. This background correction represented only $\approx 3\%$ of the maximal 2C activity for the batch purified proteins and the resulting data and kinetic parameters are shown in Fig. 2.7A,C&D and Table 2.1. When the same curve fit was used to determine the Michaelis-Menten parameters for the FPLC purified proteins the fitted offset values were within error of zero, indicating that these protein preparations did not contain contaminant ATPase activity. The data obtained using the FPLC purified proteins are shown in Fig. 2.7B and the resulting k_{cat} and K_m values for Ub2C, Ub2BC, Ub2BC3AB, Ub3AB, and Ub2C-K135A are reported in Table 2.2.

Table 2.1. Michaelis-Menten parameters of batch purified protein.

Constructs	K_m (mM ATP)	k_{cat} (μ mol ATP/ μ mol Protein/min)	Fold Activation from:
Wild Type Constructs			Ub2BC3AB
Ub2BC3AB	0.55 \pm 0.06	1,290 \pm 30	1.00
Ub2C3A	1.8 \pm 0.4	1,850 \pm 90	1.4
Ub2BC	0.47 \pm 0.08	3,200 \pm 120	2.5
2BC	0.78 \pm 0.06	3,120 \pm 60	2.4
Ub2C	0.63 \pm 0.07	10,200 \pm 300	7.9
2C	0.9 \pm 0.1	7,300 \pm 200	5.7
Ub2BC AAA+ Mutants			Ub2BC
Ub2BC-K135A	1.1 \pm 0.1	1,550 \pm 50	0.48
Ub2BC-D177A	0.7 \pm 0.1	2,520 \pm 80	0.79
Ub2BC-T222A	0.8 \pm 0.2	2,300 \pm 100	0.72
Ub2BC-N223A	0.90 \pm 0.08	2,290 \pm 50	0.72
Ub2C AAA+ Mutants			Ub2C
Ub2C-K135A	0.7 \pm 0.1	290 \pm 10	0.03
Ub2C-D177A	0.88 \pm 0.07	8,200 \pm 200	0.80
Ub2C-T222A	1.0 \pm 0.2	6,700 \pm 300	0.66
Ub2C-N223A	0.9 \pm 0.1	4,400 \pm 200	0.43

Table 2.2. Michaelis-Menten parameters of FPLC purified protein.

Constructs	K_m (mM ATP)	k_{cat} ($\mu\text{mol ATP}/\mu\text{mol Protein}/\text{min}$)	Fold Activation from Ub2BC3AB
Ub2BC3AB	0.83 \pm 0.05	158 \pm 5	1.0
Ub2BC	1.8 \pm 0.4	440 \pm 20	2.8
Ub2C	0.63 \pm 0.02	9,000 \pm 900	57
Ub2C-K135A	0.51 \pm 0.04	920 \pm 30	6
Ub3AB	1.1 \pm 0.3	5.0 \pm 0.3	0.03

Proteins obtained by both purification methods demonstrate very similar activity patterns among the various 2C constructs and mutants and these data show that proteolytic processing on both sides of the core 2C domain modulate its ATPase activity. For the more reliable FPLC purified proteins (Fig. 2.7B, Table 2.2), we found that Ub2BC3AB has the lowest turnover rate at $\approx 160 \text{ min}^{-1}$, this increased modestly to 440 min^{-1} upon removal of 3AB to yield Ub2BC, and then increased greatly upon removal of the 2B domain to give the highest k_{cat} of $\approx 9,000 \text{ min}^{-1}$ for Ub2C. This 2C rate corresponds to a ≈ 20 -fold activation compared to Ub2BC and ≈ 60 -fold increase compared to Ub2BC3AB. Only background activity of 5 min^{-1} is observed for Ub3AB, which lacks the 2C domain. The ATPase activity modulation is primarily due to changes in k_{cat} as the K_m values of the purified proteins are consistent at $\approx 1.1 \pm 0.2 \text{ mM}$ ATP for all the constructs. Note that low yield of soluble Ub2C from the FPLC based protocol made protein quantitation less reliable and this in turn increases the error of the specific activity measurements for the Ub2C protein to $\approx 10\%$.

Among the mutations of the predicted canonical AAA+ ATPase motifs, the Walker A mutation (K135A) has the most substantial effect on 2C ATPase activity with k_{cat} being reduced by 90-95%. Interestingly, the K135A only reduces activity by $\approx 50\%$ in the context of 2BC, suggesting that the active site geometries of 2C and 2BC may be slightly different. The remaining mutations in the Walker B (D177A) and motif C (T222A and N223A) regions were not as deleterious and their k_{cat} values are reduced to 40-80% of those observed for the wild type proteins in the batch purified samples.

2.4 Discussion

Among the membrane anchored poliovirus proteins, both 2C and its precursor 2BC are involved in forming the host membrane structures that serve as the site of viral replication (37, 38, 51, 52); however, the membrane binding, oligomerization, and aggregation propensities of the picornaviral 2C proteins generally make them very difficult to express and purify for biochemical and structural studies. In this work we have shown that nanodisc bilayers can be used to chaperone the expression of soluble poliovirus 2C by providing membrane surfaces that allow for proper folding and help reduce aggregation by sequestering the proteins on separate bilayers, allowing us to carry out basic biochemical characterization of the enzyme. Processing of the viral polyprotein can be important for regulating protein function, as has been shown for 3D^{pol} where processing from the 3CD^{pro} protease precursor protein activates the polymerase and even minor changes to its N-terminal residues can have significant deleterious effects on activity (42, 96, 97). To determine whether proteolytic processing plays a similar role in mediating 2C behavior, we also examined the ATPase activity of 2C when expressed as a known functional precursor protein (2BC), with a polyprotein neighbor (2C3A), or in the context of the entire surrounding polyprotein (2BC3AB).

2.4.1 Nanodisc-Bound 2C is a Highly Active ATPase

The *in vitro* expression results clearly show that nanodiscs effectively chaperone viral protein expression, resulting in highly active 2C that could be purified as a 2C-nanodisc complex whose maximal turnover rate \approx 200-fold higher than that previously reported while retaining an ATP K_m of \approx 0.6 mM that is quite similar to the 0.5 - 0.7 mM

values reported for both poliovirus GST-2C and FMDV 2C proteins (69, 76). Previous studies of poliovirus 2C expressed as a GST fusion (69) and FMDV 2C expressed with an N-terminal truncation (76) have reported 50% inhibition of ATPase activity in the presence of 20 and 50 mM NaCl, respectively. In contrast, we found a weaker salt dependence that resulted in only a $\approx 10\%$ reduction of ATPase activity in the presence of 200 mM KCl. Sodium had a stronger inhibitory effect, but even at 200 mM NaCl the nanodisc anchored proteins retained $\approx 60\%$ activity (Fig. 2.6A). When testing the magnesium dependence, we found that Mg^{2+} in the form of magnesium acetate stimulated ATPase activity while Mg^{2+} in the form of $MgCl_2$ inhibited 2C activity (Fig. 2.6D), the latter of which is consistent with prior findings of GST-2C fusion construct activity (69). Further analysis showed that the activation was due to acetate stimulation of 2C, where saturable maximal 2C activity was observed at ≈ 50 mM concentrations of either potassium acetate or magnesium acetate (Fig. 2.6D). The origin of this stimulatory effect is not yet known, but ligand-induced activation is common among AAA+ ATPases where the presence of substrates is often necessary to promote oligomerization and activity (74), suggesting that high concentrations acetate may mimic a natural ligand.

We also found that phospholipid composition can modulate expression levels and ATPase activities of both 2C and 2BC, although the 2- to 3-fold effects on each are fairly minor (Fig. 2.4). Expression of the peripherally membrane anchored 2C was highest with DMPG lipids while 2BC preferred DMPS or DMPC. While subtle, the combined ≈ 10 -fold effects of phospholipid composition on expression levels and ATPase activity suggest that lipid headgroups can play a role in modulating host

membrane interactions with 2C and 2BC. This could perhaps influence the membrane rearrangements that take place during poliovirus infection, and the shift toward DMPS for higher 2BC expression is consistent with prior studies showing that 2B preferably associates with anionic lipids such as phosphatidylserine and phosphatidylinositol that are enriched in the Golgi and ER membranes targeted by the replication proteins (101).

Electron microscopy data of both poliovirus and FMDV 2C proteins showing hexameric ring structures characteristic of this class of enzymes (75, 76) and structural modeling using the Phyre protein fold recognition server (73) indicates with high confidence that 2C has homology to the superfamily III AAA+ ATPases. We mutated canonical residues within the characteristic Walker A, Walker B and motif C sequences (72) and found that they all reduced activity (Fig. 2.7, Tables 2.1 and 2.2). In particular, K135 in the Walker A motif plays an important role in positioning the ATP γ phosphate at an inter-subunit contact, linking oligomerization and ATP hydrolysis (119). Mutating K135 to an alanine not only abolished activity, but also significantly increased Ub2C solubility and resulted in \approx 30-fold increased protein yield from the FPLC purification protocol, all of which are consistent with AAA+ domain oligomerization. Interestingly, the same mutation was not as deleterious in the context of 2BC, where it resulted in only a \approx 50% activity loss, suggesting that the added 2B domain may not only prevent oligomerization but also cause the ATPase domain to utilize a slightly different active site geometry that accounts for its lower turnover rate. Furthermore, Walker B and motif C/Sensor-1 mutations generally affect ATP hydrolysis rates significantly more than ATP binding (72, 119), which is consistent with the relatively small changes we observe in the ATP K_m values for the various mutants.

Based on our findings, we believe that the nanodisc bound 2C protein is indeed forming multimeric structures that are necessary for AAA+ ATPase activity, although we cannot say if these are discrete hexamers or larger oligomers based on our data. Nanodiscs are quite effective at preventing oligomerization of integral membrane proteins by sequestering the transmembrane segments into individual discs, but it is important to note that 2C is a peripheral membrane protein and that its oligomerization is likely based on protein-protein interactions outside the lipid bilayer. As a result, the discs may chaperone proper protein folding, but do not necessarily prevent oligomerization. Our expression and purification data indicate that the isolated 2C protein is significantly more prone to aggregation or oligomerization than the larger 2BC and 2BC3AB constructs. We postulate that the propensity of 2C to oligomerize will bring multiple nanodiscs together in close proximity, creating a high local concentration that facilitates fusion of the disc membranes into larger bilayer structures and reducing solubility, in effect defeating the advantage of using the nanodiscs. Consistent with this, the K135A mutation that is predicted to disrupt oligomerization results in more soluble material and a significant increase in protein yield. It is also possible that 2C itself may be acting to change the structure of the nanodiscs as it is known to be involved in membrane rearrangement (37, 38, 51).

2.4.2 Proteolytic Control of 2C ATPase Activity

Comparative analysis of enzyme rates by the different protein constructs shows that poliovirus 2C ATPase activity is strongly modulated by the presence of its polyprotein neighbors. Starting with the 2BC-3AB protein, ATPase activity is increased

≈3-fold by removal of 3AB to generate 2BC and then another ≈20-fold upon removal of 2B to generate the fully processed 2C protein. This reflects an overall ≈60-fold increase in activity for the fully processed 2C as compared to its fully embedded precursor form and indicates that polyprotein neighbors exert allosteric effects on the core 2C domain that can modulate the rate of ATP hydrolysis. It is also notable that both the 2C and Ub2C proteins have higher activities than 2BC, showing that the allosteric effects are specific to having a 2B domain. This is in contrast to the 3D^{pol} proteolytic activation event where any N-terminal modification, even adding a single amino acid residue, will inhibit activity (41).

Overall, the results from this study lead to a model in which precursor forms of 2C, such as 2BC and 2BC3AB that contain all the membrane binding elements of the viral polyprotein, assemble into the viral replication complex in a precursor form. During proteolytic processing, the viral proteases carry out multiple cleavages of the polyprotein, the first of which is 2A^{pro} mediated cleavage of the P1 region from the polyprotein and this is followed by the 3C^{pro}/3CD^{pro} mediated cleavages of the P2 and P3 regions (120). The exact order in which the P2 and P3 proteins are cleaved is not well established, but the intact P2-P3 protein is rarely observed during infection and it is thought that the P2-P3 junction (i.e. 2C-3A linkage) is cleaved early during processing. This would make 2BC the predominant precursor form, wherein 2C ATPase activity is kept at a low basal level because the presence of the 2B domain sterically prevents 2C from forming the oligomeric structures required for maximal ATP hydrolysis. This pre-assembly of 2BC proteins on the membrane surface is then followed by processing of the 2B-2C junction, triggering 2C to assemble into larger oligomeric structures that

increase ATPase activity \approx 20-fold. The true purpose of this ATPase activity during viral replication is not yet known and not addressed by our data, but it may play a role in RNA helicase or translocation activity, capsid assembly, RNA packaging, or perhaps a combination of these functions.

In conclusion, we have shown that the membrane associated poliovirus proteins can be synthesized in an active and soluble form when anchored to small nanodisc membrane bilayers. The core domain within these proteins is the 2C ATPase, which is shown to exhibit high ATPase activity over a broad range of monovalent salt concentrations, pH values, and temperatures. The activity of this core domain can be modulated by the addition of either N- or C-terminal polyprotein neighbors and the different polyprotein constructs exhibit slight activity preferences for different lipid headgroups. Although issues with oligomerization and solubility continue to present a clear obstacle to structural studies of poliovirus 2C that would require large amounts of purified protein, our nanodisc based expression data do set the stage for experiments to further understand the interactions of the picornaviral 2C and 2BC proteins with RNA and to study their effects on membrane dynamics by biophysical methods.

CHAPTER 3

ASSESSMENT OF 2C HELICASE ACTIVITY AND ADDITIONAL PRELIMINARY CHARACTERIZATION

This chapter contains preliminary and/or inconclusive data obtained during characterization of the PV 2C protein that was not reported in Chapter 2. This data includes assessing 2C helicase activity, oligomerization state, and secondary structural analysis.

3.1 Introduction

As discussed in sections 1.3 and 2.1, 2C contains the NTP binding motif C that is characteristically found in the helicase/translocase Super Family 3 (SF3) proteins (69, 121, 122). 2C also binds RNA (61) and demonstrates ATPase activity (68-70, 75), but the purpose of this activity has not yet been identified. Thus, it has been postulated that 2C may function as a RNA helicase or translocase during viral replication. To date, there have been no reports of 2C exhibiting helicase activity (69, 70). Since our *in vitro* expression system in the presence of nanodiscs resulted in such highly active 2C protein, which we attribute to the proteins being in the more biologically relevant context of membranes, we speculated that we may be able to observe 2C helicase activity with our nanodisc-bound protein.

Translocases are enzymes that utilize the energy from ATP hydrolysis to move along strands of nucleic acids and are further classified as helicases if they are also able to unwind duplexed nucleic acid strands. Currently, these enzymes are classified

into six superfamilies (SF), SF1-SF6, and are further classified based on their rate, direction, step size, and processivity of translocation (121, 123). First identified in the genomes of small RNA and DNA viruses, SF3 helicases are distinguished by the presence of a N-terminal non-helicase domain followed by a modified AAA+ ATPase core domain containing the standard Walker A and Walker B motifs as well as the SF3-specific motif C. Similarly, the central portion of 2C shares homology to the AAA+ proteins and contains motif C whereas the N-terminal domain of 2C is involved in membrane and RNA binding but does not demonstrate intrinsic ATPase activity. Members of the SF3 helicases oligomerize into hexameric or double hexameric rings, a feature also shared by 2C (75), and translocate in the 3'-5' direction (121). Examples of SF3 helicases include the Rep40 protein of adeno-associated virus type 2 (124), the simian virus 40 helicase (125), and the papillomavirus E1 protein (126).

3.2 Materials and Methods

3.2.1 Design and Preparation of RNA for Helicase Assays

I designed several RNA constructs to test the helicase/translocase activity of 2C. The final sequences (Fig. 3.1A) are a slightly modified version of a PETE (polymerase elongation template element) RNA previously used in our laboratory for polymerase binding and elongation studies (84, 127-130) and were ordered from Integrated DNA Technologies. The unlabeled strands contain a core 22 nt sequence (RNA-3) with a 5' 10mer polyA (RNA-4) or a 10mer 3' polyA (RNA-1). When annealed to the labeled strand this creates a blunt end double stranded 22mer (RNA-B), or a double stranded section with a 5' overhang (RNA-A) or a 3' overhang (RNA-C). Complementary to the

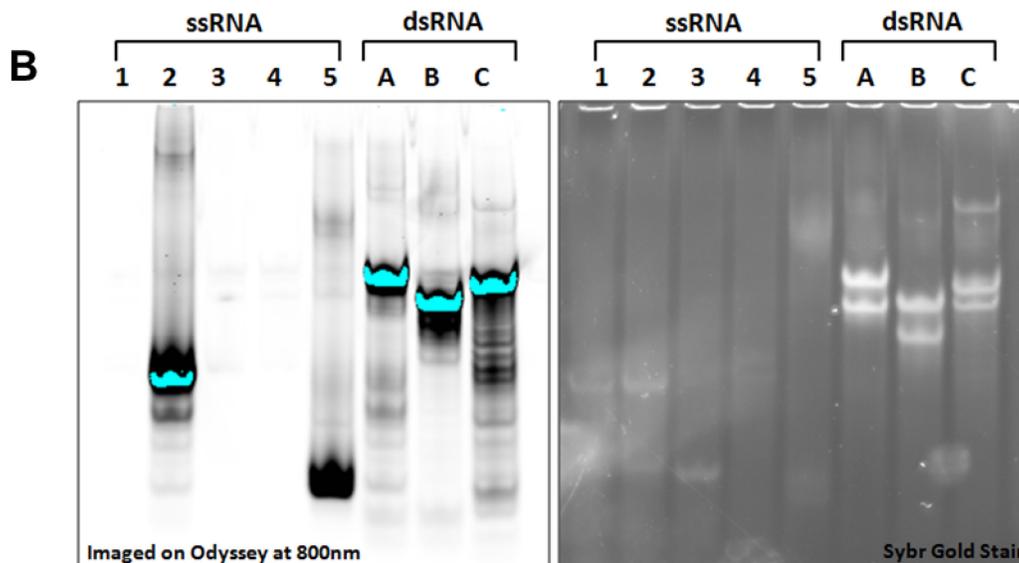
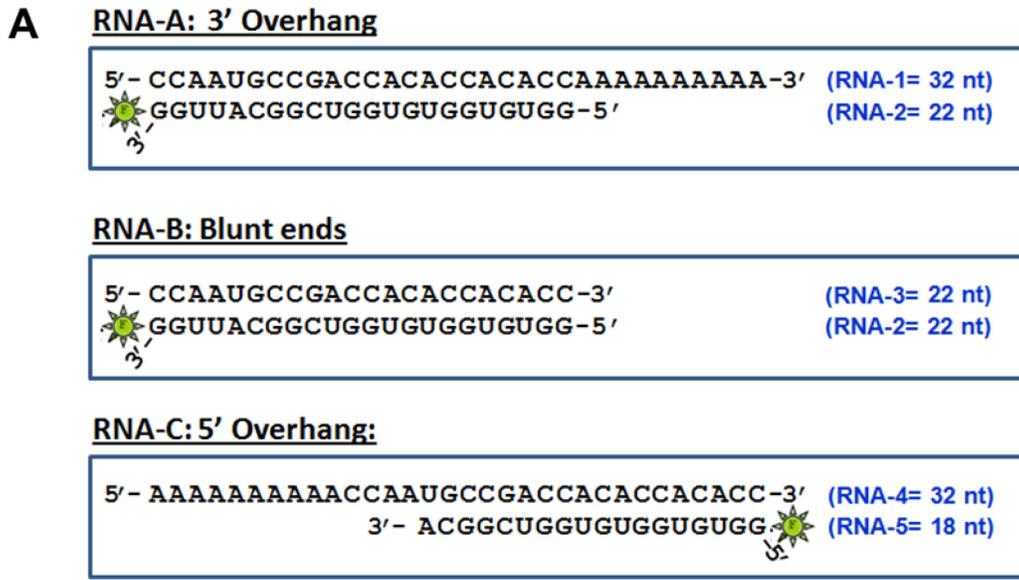


Fig. 3.1. Design of RNA constructs for 2C helicase assays.

A) ssRNAs labeled with LI-COR 800CW maleimide (green star) were annealed to unlabeled RNA strands to make dsRNA-A with a 3' overhang, dsRNA-B with blunt ends, and dsRNA -C with a 5' overhang. **B)** Non-denaturing 20% PAGE gel of the ssRNA and annealed dsRNA. The same gel was imaged on the Odyssey at 800 nm (left) and also photographed after being stained with SYBR Gold (right). Please note that the blue bands in the Odyssey image merely reflect saturation from overexposure during imaging. Doublet bands seen in the SYBR stained gel reflect the different mass of the labeled vs. unlabeled strands (800CW dye is 1166 g/mol).

core 22-mer sequence are the 22 nt RNA-2 containing a thiol modification at the 3' end, and the 18 nt RNA-5 with a thiol modification at the 5' end. Finally, a DNA strand complementary to the core 22-mer sequence (DNA-1) and 18mer (DNA-2), the DNA equivalent to RNA-2 and -5 respectively, was also ordered to be used to test RNA vs. DNA binding to 2C, or as a competitor strand.

RNA-2 and RNA-5 were labeled with IRDye 800CW maleimide from LI-COR Biosciences via the thiol modifications, as previously described (84). These labeled strands were then annealed to the unlabeled RNA strands 1, 3, and 4 to generate three dsRNA constructs: RNA-A with a 3' overhang, RNA-B with blunt ends, or RNA-C with a 5' overhang (shown in Fig. 3.1A). The strands were annealed by heating at 95°C for 2 min then gradually cooling to 25°C over a 45 minute period in buffer containing 5 mM Tris, pH 7.5 and 10 mM NaCl. To verify proper labeling and formation of the RNA duplex, all the single stranded and annealed RNA were run on a non-denaturing 20% polyacrylamide gel (Fig. 3.1B). Note that the blue bands that appear in the 800 nm Odyssey image correspond to RNA samples that have been overexposed during and merely reflect intensity saturation, and blue bands present in other gels in this chapter are similarly a result of overexposure.

merely reflect saturation from overexposure during imaging

3.2.2 Electrophoretic Mobility Shift Assays (EMSA)

2C protein containing a N-terminal ubiquitin tag (Ub2C) was FPCL (fast protein liquid chromatography) purified over a Nickel column and ion exchange column, as described in 2.2.3 ,and concentrated to $\approx 1 \mu\text{M}$. EMSA assays were carried out with

Ub2C and either dsRNA-A, ds RNA-B or dsRNA-C in ATPase buffer as described in section 2.2.5 (20 mM HEPES, pH 7.0, 5 mM DTT, 50 mM MgOAc, unless noted) with the exceptions that NADH was not included and 30 mM KCl instead of 75 mM KCl was used. The different concentration of KCl was used to account for the 300 mM NaCl in the HiTrap Q column elution buffer, so by using 30 mM KCl the final salt concentration in the pre-initiation mixture was ~100 mM. The dsRNA and Ub2C were pre-incubated at room temperature for 20 minutes to 1 hour to allow time for Ub2C to bind the RNA and then reactions were diluted into ATPase buffer containing the reagents as indicated in the figures. This was typically a mixture of 10-fold or 50-fold excess unlabeled ssRNA-2 or -5 (so that any labeled strand that was displaced would not rebind the unlabeled strand) and 12.5 mM ATP (near V_{max} concentration) in buffer. The final concentration of Ub2C was 100 nM and the final concentration of dsRNA was 20-100 nM (as noted in the figure legends). Reactions were incubated at room temperature (RT) or 37°C for 2-4 hours (as indicated in the figure legends) and reactions were terminated by the addition of non-denaturing loading buffer containing 200 mM EDTA (40% glycerol, 0.32% bromophenol blue, 200 mM EDTA) at a ratio of 1:1 sample:buffer. The samples resolved on non-denaturing 20% acrylamide gels and subsequently imaged at 800 nm on the Odyssey infrared scanner.

3.2.3 RNase T1 Helicase Assays

The RNase T1 helicase assays were done in ATPase buffer as described above for the EMSA experiments containing 3 mM MgCl₂ rather than 50 mM MgOAc (20 mM HEPES, pH 7.0, 5 mM DTT, 3 mM MgCl₂, 30 mM KCl), as Mg²⁺ inhibits RNase T1

activity. I opted to use magnesium in the form of $MgCl_2$ rather than $MgOAc$ in light of the acetate-dependent stimulation (discussed in sections 2.3.4 and 2.4.1) for which we hypothesize acetate may be mimicking a natural ligand, such as RNA. Since $MgCl_2$ was used in the reactions, when ATP was included in a reaction, it was added at a final concentration of 2 mM. Mixtures containing Ub2C and dsRNA were pre-incubated for ≈ 30 min at $37^\circ C$ and then diluted in ATPase buffer containing RNaseT1 and (+/-) ATP, so that the final concentrations of the reagents were 25-100 nM Ub2C, 25-50 nM dsRNA, 0.1-1 unit/ μL RNase T1, and 3 mM ATP (actual concentrations are noted in the figure legends). Reactions were incubated at $37^\circ C$ and samples were taken at the indicated times and mixed at a 1:1 ratio with denaturing loading buffer (95% formamide, 0.025% bromophenol blue, 0.025% xylene cyanol, 0.025% SDS, 5 mM EDTA) and boiled for 5 minutes. The samples were then resolved on a denaturing 15-20% polyacrylamide/7M Urea gel and imaged at 800 nm on the Odyssey infrared scanner.

3.3 Results of 2C Helicase Assays

3.3.1 Assessment of 2C Helicase Activity via EMSA

Initial characterization of 2C helicase activity was done using non-denaturing electrophoretic mobility shift assays (EMSA, Fig. 3.2A). The goal of my preliminary helicase assay was to determine whether Ub2C was able to separate any of the three dsRNAs and try to get a feel for the appropriate dsRNA and ssRNA trap to use. This experiment, shown in Fig. 3.2B, was carried out with 100 nM Ub2C, either 20 nM or 100 nM dsRNA-A, -B, or -C in the presence of 10-fold or 50-fold excess ssRNA-2 (over the dsRNA concentration) and 12.5 mM ATP. Control reactions were also set up with 50-

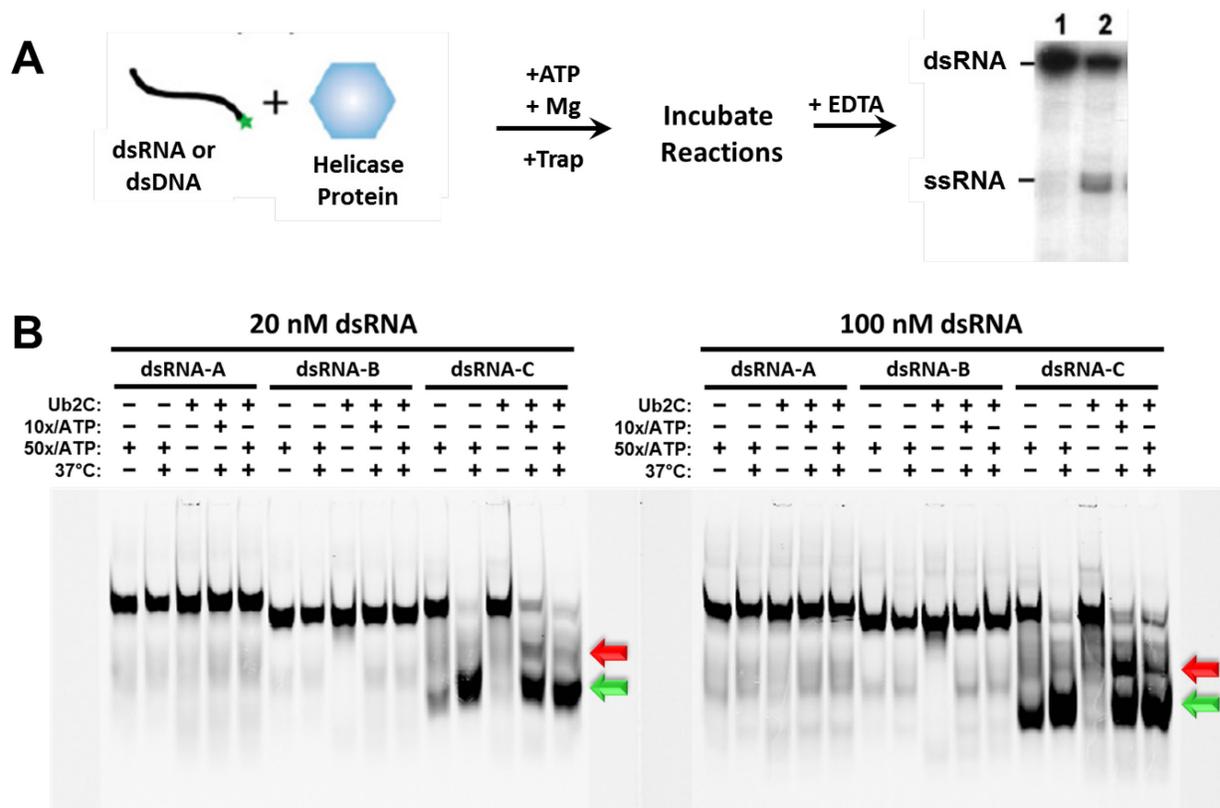


Fig. 3.2 Ub2C EMSA helicase assay with RNA-A/B/C.

A) Overview of the EMSA assay. ATP, magnesium, and a ssRNA trap are added to the pre-incubated dsRNA and protein to initiate the reaction. Reactions are terminated with EDTA and resolved on a native gel. Lane 1 is representative of the dsRNA prior to initiating the reaction and lane 2 shows the separated ssRNA containing the label along with some dsRNA substrate that has not been unwound. Figure adapted from reference (131) for illustrative purposes. B) Results from 2C helicase assays containing 100 nM final Ub2C concentration with either 20 nM or 100 nM dsRNA-A/B/C and 10x or 50x labeled RNA-2 trap. The green arrow indicates the location of the free labeled 18-mer (RNA-5) and the red arrow indicates the location of the RNA band that appears in the presence of Ub2C. 0.2 pmol of dsRNA were loaded per well and run on a non-denaturing 20% polyacrylamide gel. 10x/ATP = 0.2 μ M or 1 μ M unlabeled ssRNA-2 + 12.5 mM ATP; 50x/ATP = 1.0 μ M or 5 μ M unlabeled ssRNA-2 + 12.5 mM ATP; 37°C = sample taken after 2.5 hr. incubation at 37°C (otherwise taken immediately after mixing). Reactions were carried out in ATPase buffer containing 50 mM MgOAc.

fold ssRNA-2 trap except no protein was added in order to verify that the trap would not displace the labeled strand in the absence of protein. After incubating the (+/-)Ub2C and dsRNA mixture at room temperature for 1 hour, a sample was taken (denoted as “-“ 37°C in Fig. 3.2B) and then the remaining volume was mixed with ssRNA-2 and ATP and incubated at 37°C for 2.5 hours. The resulting gels do not reveal labeled strand separation for dsRNA-A or dsRNA-B. The labeled 18mer (ssRNA-5) in the dsRNA-C did appear to dissociate, although we suspected this dissociation may be the result of the unlabeled ssRNA-2 competing off the 18-mer, since free 18-mer is seen even in the(-)Ub2C control reactions. Interesting, however, was the appearance of a band approximately halfway between the free 18-mer and the dsRNA-C duplex that is only seen in the (+) Ub2C reactions (as indicated by the red arrow in Fig. 3.2B) perhaps indicating that 2C has some affinity for the 3' end of the non-overhanging strand at a 5' overhang.

To determine whether the separation seen for RNA-C was as a result of 2C helicase activity or whether ssRNA-2 was displacing the labeled ssRNA-5, I repeated the reaction explained above, but with no Ub2C protein (Fig. 3.3). In addition to doing this experiment in the presence of 10- and 50-fold excess unlabeled RNA-2 (RNA-2U), I repeated this experiment using labeled RNA-2 (RNA-2L, Fig. 3.3), unlabeled RNA-5 (RNA-5U, Fig. 3.3), DNA-1 or DNA-2 (data not shown for DNA) to see if they could compete off the annealed RNA-5 in RNA-C. ATP was left in the reactions despite the absence of Ub2C in order to fully mimic the reaction conditions. Indeed, only the unlabeled RNA-2 appeared to be displacing RNA-5 from the duplex, a phenomenon that

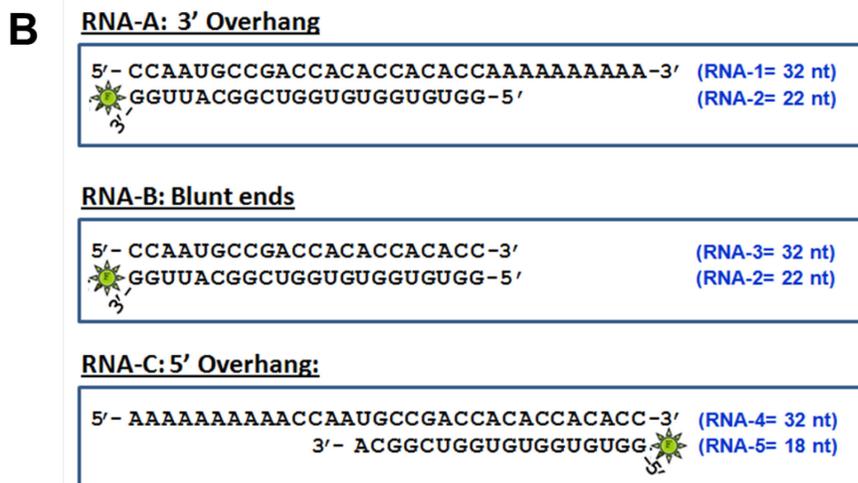
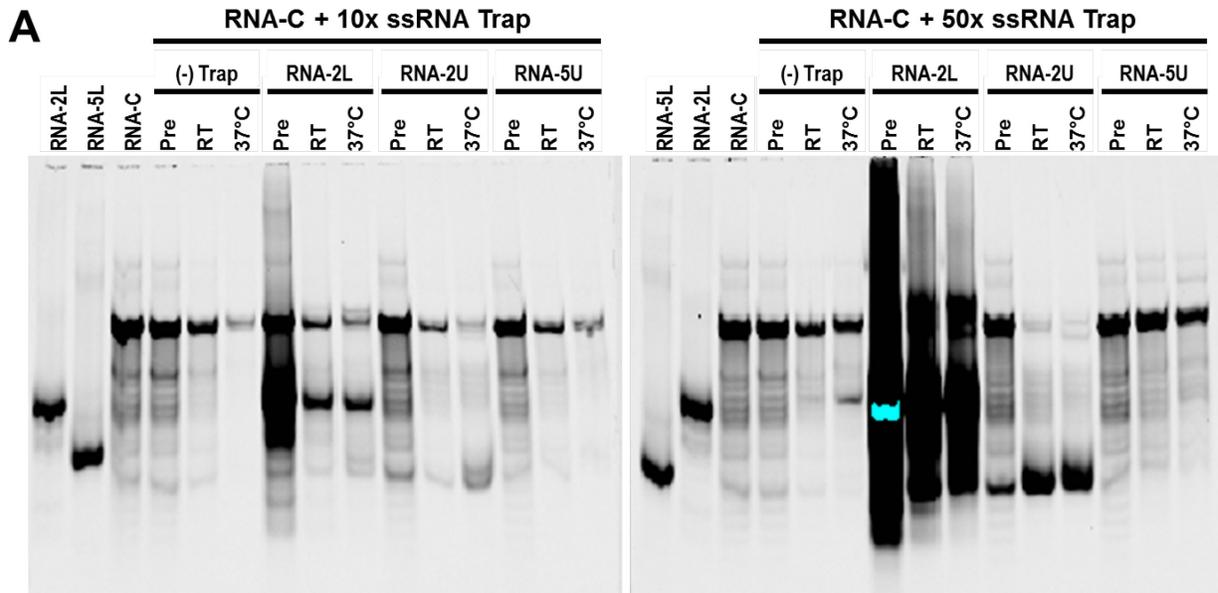


Fig. 3.3. EMSA competition assay with RNA-C and no protein.

A) Competition assay containing 100 nM RNA-C with 10x or 50x single stranded labeled RNA-2 (RNA-2L), unlabeled RNA-2 (RNA-2U), unlabeled RNA-5 (RNA-5U), or no additional ssRNA ((-)Trap). 10x/ATP = 1 μ M ssRNA + 12.5 mM ATP; 50x/ATP = 1.0 μ M unlabeled ssRNA + 12.5 mM ATP. Mixtures were incubated at either room temperature (RT) or 37°C for 3 hours and 45 min. Reactions were carried out in ATPase buffer containing 50 mM MgOAc. Stock RNA-2L, RNA-5L and RNA-C were run for markers. Note that the labeled version of RNA-5 was not used in this competition experiment since it was the same as the labeled strand in RNA-C. The blue band appearing in the gel to the right is a result of overexposure during imaging. B) Image of RNAs used in this assay from Fig. 3.1A for reference.

was much more pronounced when used at 50-fold excess over RNA-C. However, the middle band seen in the previous helicase assay with Ub2C (Fig. 3.2B, red arrow) was not present in this experiment. Since RNA-2 was competing off RNA-5, I tested whether the DNA oligos could be used as traps. I did this by subjecting RNA-C and unlabeled ssRNA or ssDNA mixtures (at the same concentrations/conditions as used in the competition assay) to a round of annealing (95°C for 2 min, cool to 25°C over 45 min). As there was excess unlabeled single stranded nucleic acid, the annealing should cause most of the labeled oligos to be replaced by the unlabeled complements. Indeed, the unlabeled ssRNA-2 and -5 did anneal to the 32mer RNA-1; however, the ssDNA oligos did not anneal to RNA-1 (data not shown). Therefore, we decided to carry out the subsequent assays using RNA-5 as well as RNA-2 as a trap in only 10-fold excess for comparison.

The putative helicase activity of Ub2C was again tested with RNA-C this time using RNA-5 and RNA-2 at 10-fold excess over RNA-C and in the presence or absence of ATP, to see if the third band observed in Fig. 3.2B was unique to reactions containing RNA-2 and whether it was ATP dependent. Mixtures containing RNA-C and (+/-) Ub2C in the standard ATPase buffer were pre-incubated at RT for ≈20 min before adding buffer containing the ssRNA trap and (+/-)ATP and the reactions were then incubated at RT or 37°C for 3.5 hours. From the resulting gel (Fig. 3.4A), it appears that this middle band is unique to the presence of the RNA-2 trap, albeit it is much more pronounced in the presence of 2C but does not seem to be ATP-dependent.

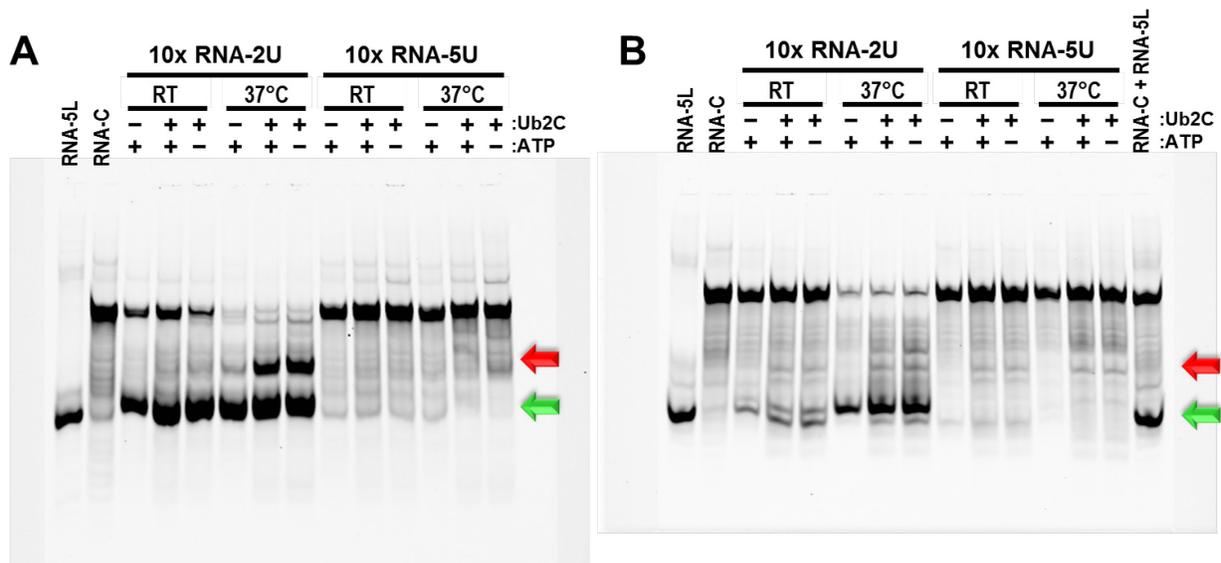


Fig. 3.4. Ub2C EMSA helicase assay with RNA-C in MgOAc vs. MgCl₂.

Helicase assays containing (+/-)100 nM final Ub2C concentration, 100 nM dsRNA-C, and 10x (1 μ M) unlabeled RNA-2 or RNA-5 trap in ATPase buffer containing either A) 50 mM MgOAc and (+/-) 12.5 mM ATP or B) 3 mM MgCl₂ and (+/-) 2 mM ATP. The green arrows indicate the location of the free labeled RNA-5 18-mer and the red arrows indicate the unknown middle band species. The Ub2C/RNA-C mixtures were incubated at room temperature (RT) for 20 min before initiating the reactions. Initiated reactions were incubated at RT or 37°C for A) 3.5 hours and B) 3 hours.

Up to this point the helicase assays were carried out in MgOAc and we knew that acetate had a significant stimulatory effect on 2C ATPase activity (see section 2.3.4). One theory as to why acetate had such an effect is that it could be acting as an RNA mimic, thus promoting oligomerization and activity. If this was indeed the case, then the presence of acetate could potentially be limiting the association of 2C and RNA. Therefore, I tested 2C helicase activity in the presence of magnesium in the form of MgCl₂ (Fig. 3.4B). This experiment was carried out as described for the previous assay the ATP concentration since we and other labs have observed that MgCl₂ inhibits ATPase activity as it exceeds a 1 mM concentration difference (discussed in section 2.3.4). The results (Fig. 3.4B) are similar to that seen in Fig. 3.4A wherein there the displacement of the 18mer RNA-5 from dsRNA-C only occurs only with the ssRNA-2 trap both in the presence and absence of 2C. Not as much 18mer was displaced when MgCl₂ was used as compared to MgOAc, which is likely why it is visible as a doublet band in the presence of MgCl₂ rather than a large smeared band (green arrow). Interestingly, a faint middle band similar to that previously seen only when the RNA-2 trap was used appeared in the reactions for both the RNA-2 and RNA-5 traps (red arrow). This band appears only in the presence of Ub2C and does not appear to be ATP dependent. Despite seeing some difference in reactions containing Ub2C, definitive conclusions could not be obtained from these experiments. This in addition to our concerns that the ssRNA traps could be displacing the 18mer in RNA-C and the fact that these differences did not appear to be ATP dependent led us to develop another method to assess 2C helicase activity.

3.3.2 Assessment of 2C Helicase Activity via RNase T1 Assays

In order to avoid the use of RNA traps, I designed a helicase assay using RNase T1. RNase T1 specifically cleaves single stranded RNA at the phosphodiester bond between the 3'-guanylic residue and the 5'-OH of the neighboring nucleotide. Since all three of my dsRNA constructs contained G residues in the labeled strand and there were no G residues in the single stranded overhang portions (see Fig. 3.2A), this enzyme was a promising means by which to monitor strand separation. Fig. 3.5A&B shows the various single stranded products predicted to result from RNase T1 degradation of ssRNA-2 (labeled at the 3'-end) and ssRNA-5 (labeled at the 5'-end) separated ssRNA strands from the dsRNA-A, -B, or -C and still retain the LI-COR label.

As proof of principle and to make sure that RNase T1 would not degrade the various duplexed RNA constructs, I tested the labeled ssRNAs (-2&-5) as well as the dsRNAs in the presence and absence of the T1 enzyme and monitored the reactions over the course of an hour at 37°C (Fig. 3.5C). As anticipated, ssRNA-2 and ssRNA-5 show signs of degradation in the presence of T1 in as little as a minute and after 15 minutes a ladder of the predicted cleavage products is clearly visible, with the number of bands observed for each being very close to that predicted (11 for RNA-2 and 10 for RNA-5). The lack of similar bands in parallel reactions without the addition of T1 that were incubated for 60 minutes demonstrates that the banding is not due to unintended degradation. The dsRNAs showed very little T1-dependent degradation after 15 min and low levels of degradation after 60 min, but was still significantly less than that seen for the ssRNAs after 15 min. These results were encouraging that this was a viable means by which to test for potential 2C helicase activity.

Using this assay, I monitored the cleavage products produced from the three dsRNAs when incubated either in the presence of 2C and ATP or in the presence of Ub2C alone as well as control reactions in which no 2C enzyme was added (Fig. 3.6). Samples were taken immediately after starting the reactions and after incubating the reactions at 37°C for 75 minutes. Upon examination of the gel, one can immediately see that there is a sizable amount of material that remained in the wells only in the samples that contained 2C, suggesting that 2C-RNA complexes may be forming, but are precipitating out of solution. Comparison of the cleavage products for each RNA at 75 minutes reveals a distinct banding pattern for RNA-A and RNA-C only when 2C is present, but there does not appear to be any significant difference between the 75 minute RNA-B samples in the presence versus absence of 2C. The 75 minute samples containing 2C and RNA-A or RNA-C show a different laddering pattern in which the smaller cleavage products between \approx 1-10 nucleotides in length are no longer visible as they are in the equivalent reactions lacking 2C. Given that there was still a significant amount of labeled strands that was not degraded and the lack of a positive control, it is difficult to definitively interpret these results. It may be that this observed footprinting is merely indicative of 2C binding the RNA and protecting it from being degraded and not actually separating the strands. Alternatively, it could be that there was simply not enough 2C or RNase T1 to carry out unwinding or degradation of separated strands, respectively.

Since we could not concentrate the protein any further as it tended to precipitate very rapidly, nor could we add more 2C in the reactions, which would increase the salt concentration and could prevent 2C from binding the RNA, we opted to increase the

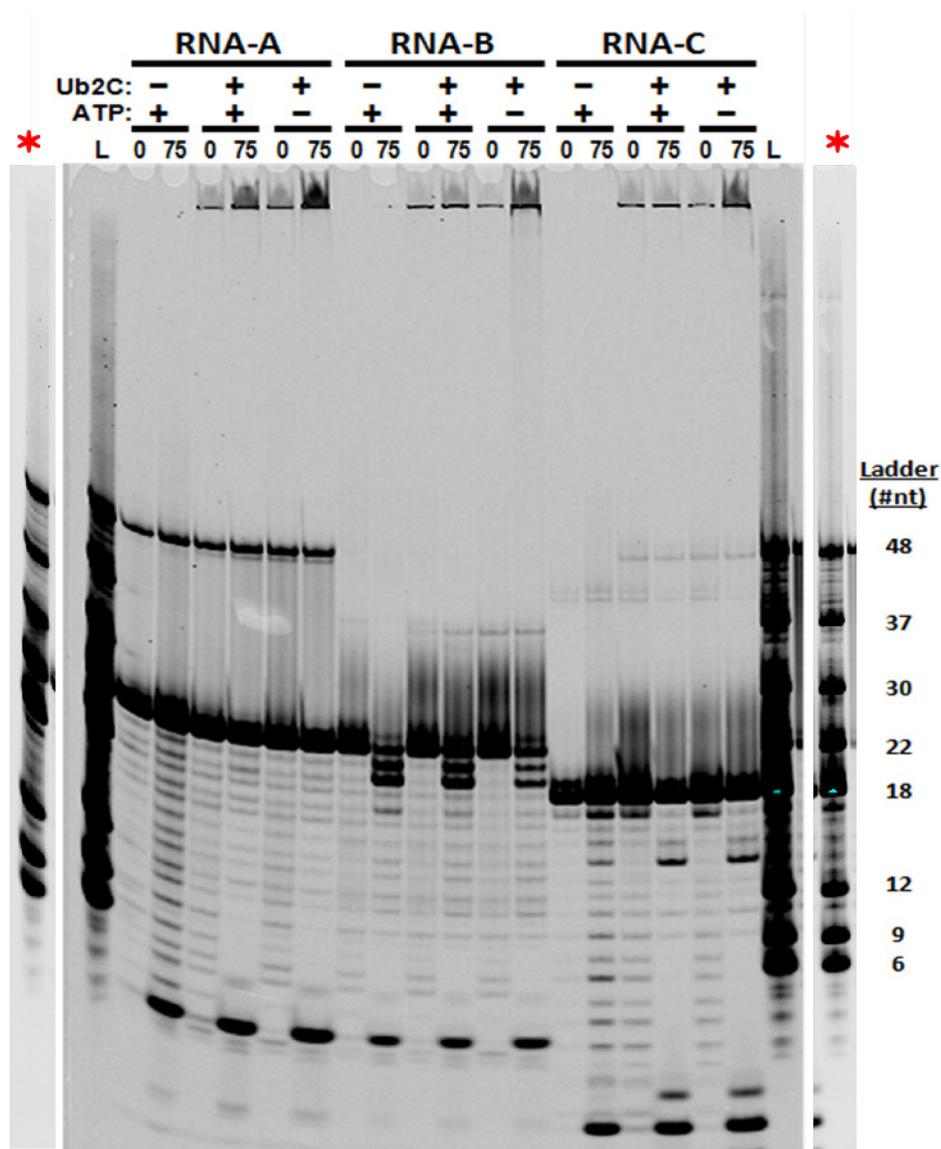


Fig. 3.6. RNaseT1 helicase assay with 0.1 U/ μ L T1.

Helicase assay of dsRNA strands mixed either (-)Ub2C/(+)ATP, (+)Ub2C/(+)ATP, or (+)Ub2C/(-)ATP in the presence of RNaseT1. There does not appear to be significant loss of dsRNA suggesting that Ub2C does not have helicase activity, although there does appear to be some protection in the RNA-A and -C reactions when 2C is present. Reactions were carried out with 0.1 Units/ μ L RNase T1 in 20 mM HEPES, pH 7, 60 mM KCl, 3 mM MgCl₂, 5 mM DTT and 25 nM dsRNA in the presence (or absence) of 100 nM Ub2C and 2 mM ATP and samples were resolved on a 15% polyacrylamide/7 M Urea gel. The lanes on each end marked by a red star are copies of the RNA ladders immediately next to them in which the intensities have been reduced so that the markers in the RNA ladder can be seen more clearly.

concentration of T1 in the reactions to a final concentration of 1 unit/ μ L. Unfortunately, this had the effect of degrading too much of the RNA in the reactions, again making it difficult to draw any conclusions. An example of this is shown in Fig. 3.7 in which either ds- or ssRNA was used at a final concentration of 50 nM and RNase T1 was added at a final concentration of 1 unit/ μ L. Having run out of the concentrated FPLC purified aliquots, this reactions was carried out with FPLC stocks that had not been concentrated, making the final concentration 25 nM Ub2C in the (+) reactions in Fig. 3.7. In this experiment, I additionally wanted to test whether the denaturing loading buffer I was using was inhibiting RNaseT1 activity. Thus, I treated the (-) and (+) reaction samples for each timepoint with the protease proteinase K. Proteinase K was added to the sample at a final concentration of 200 μ g/ μ L and incubated at 50°C for 30 minutes before adding the denaturing loading buffer. To directly compare the two methods, I also took a 15 minute timepoint sample of the Ub2C-containing reaction prior to proteinase K treatment and only added denaturing buffer, which is marked as (\pm) in Fig. 3.7 (i.e. to denote (+)Ub2C/(-)proteinase K). As you can see from the gels, there was a significant amount of RNA degradation, making the results difficult to interpret, although comparison of the 15 minute samples containing Ub2C that either been treated with proteinase K (+) or added to denaturing buffer (\pm) to terminate the reaction. At this point, the stock of FPLC purified protein had nearly been exhausted and we decided to utilize the remaining material for several experiments that were pertinent to the studies described in Chapter 2, rather than use it to further optimize the helicase assays.

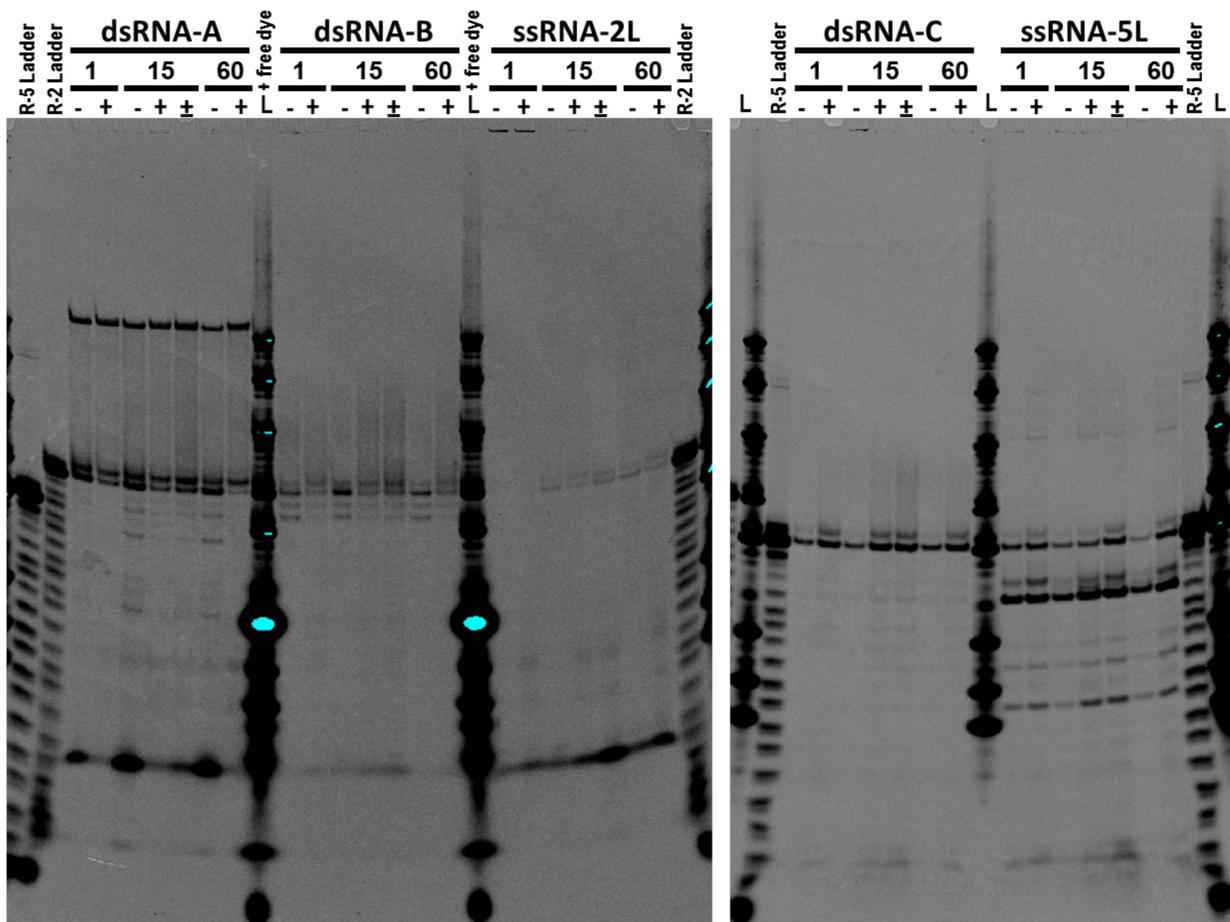


Fig. 3.7. RNaseT1 helicase assay with 1 U/ μ L T1 and proteinase K.

Helicase assay using 1 unit/ μ L RNase T1, 50 nM dsRNA, and 2 mM ATP in the presence(+) or absence (-) of 25 nM Ub2C with samples taken at 1, 15, and 60 minutes after initiation the reactions. The (+) and (-) reactions were terminated by proteinase K treatment. Lanes marked by (\pm) indicate that a sample from the (+)Ub2C reaction was taken before the addition of proteinase K and only the standard denaturing buffer was used to terminate the reaction. All reactions were carried out in 20 mM HEPES, pH 7, 60 mM KCl, 3 mM MgCl₂, 5 mM DT. Samples were resolved on a 20% polyacrylamide/7 M Urea gel. Lanes marked with “L” contain the same ladder that was used in Fig. 3.6, with the addition of the free 800CW maleimide dye in the gel on the left. The lanes marked “R-2 Ladder” or “R-5 Ladder” contains labeled, ssRNA-2 or -5 alkaline hydrolysis ladders.

3.4 Additional Preliminary Characterization

3.4.1 *Mixing Experiments*

It has been reported that FMDV 2C and PV 2C form oligomeric rings of 5-8 subunits that exhibit a coordinated ATP hydrolysis mechanism, in which the ability of each subunit to hydrolyze ATP is linked to the other subunits in the ring (75, 76). To determine whether Ub2C expressed on nanodiscs demonstrated this same behavior, I conducted a mixing experiment similar to that described for FMDV 2C (76). In this experiment, active 2C is mixed the inactive 2C Walker A mutant at molar ratios ranging from 6:0 to 0:6 and assessed for their ATPase activity in order to determine how the systematic replacement of a subunit in the 2C hexamer with an inactive enzyme will affect the ability of the complex to hydrolyze ATP. As was seen for FMDV 2C, a non-linear decrease in activity in the presence of higher ratios of the Walker A mutant is indicative of a coordinated hydrolysis mechanism, reflecting that loss of subunit function within the ring reduces the activity of the whole complex (76). The results of this mixing experiment carried out with batch purified poliovirus Ub2C and Ub2C-K135A are shown in Fig. 3.8A. Unexpectedly, I observed a linear relationship between the ATPase rate and concentration of wild type Ub2C, indicating that ATP was not being coordinately hydrolyzed by poliovirus 2C. I further verified that the ubiquitin tag was not preventing oligomerization or ATPase activity by repeating these experiments in the presence of a deubiquitinating enzyme (data not shown). However, based on the on the low solubility of our wild type 2C and its propensity to precipitate (as discussed in Chapter 2), I suspect that 2C has already oligomerized during the purification process and simply mixing two protein batches together at low concentrations was not enough to dissociate

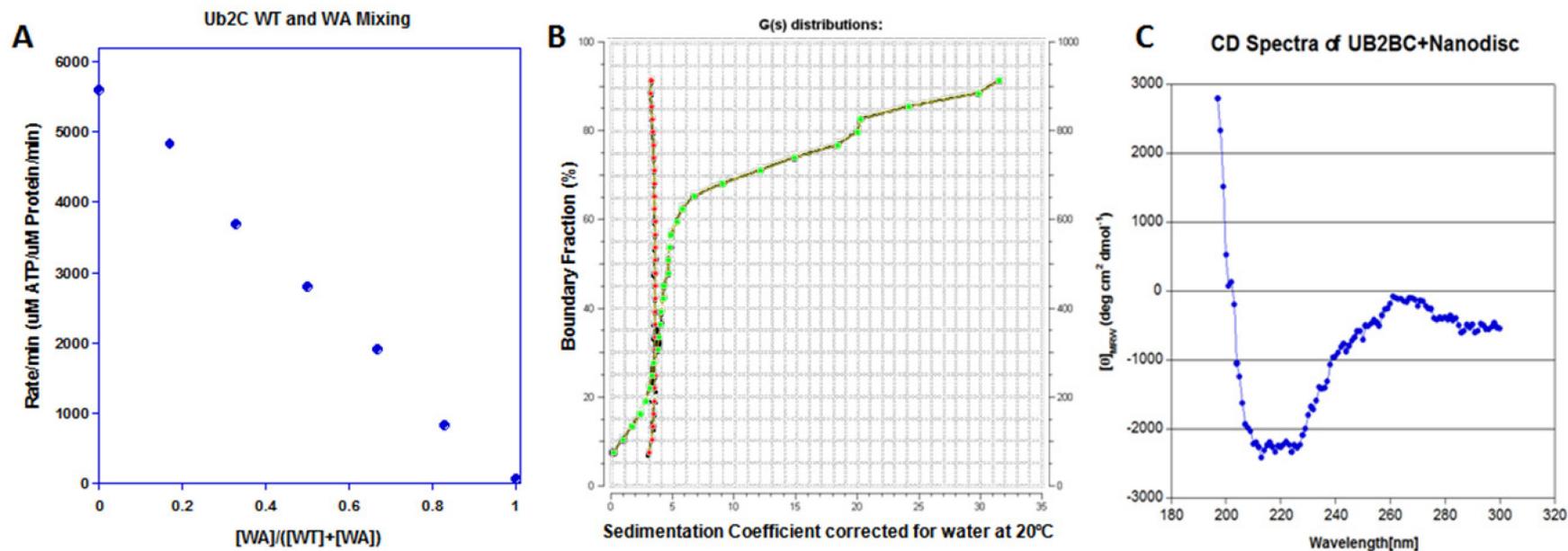


Fig. 3.8. 2C oligomerization state.

A) Mixing experiment with Ub2C and the Walker A mutant Ub2C-K135A. ATPase reactions were done in standard ATPase buffer conditions (50 mM Mg/75 mM KCl, 20 mM HEPES/KOH, pH 7.0, 5 mM DTT, 37°C), with 2 mM ATP and 1 nM total protein. B) van Holde-Weischet distribution plot of sedimentation velocity experiment showing nanodiscs alone (red) and 2C3A+nanodiscs (green). AUC experiments were carried out in 200 mM NaCl, 5 mM Tris, pH 7.5, 2% Glycerol and 5 mM DTT. C) Mean Residue CD ellipticity of UB2BC+Nanodiscs. The concentration of this sample was very low, resulting high background noise. CD experiments were carried out in 20 mM NaPhos, pH 7.5, 100 mM NaCl, and 1 mM DTT.

the preformed complexes. Perhaps future mixing experiments can be carried out using wild type and mutant 2C that has been co-expressed at different ratios in the same *in vitro* reaction to promote the formation of rings containing the two proteins immediately after translation.

3.4.2 Preliminary AUC and CD Data

I also carried out preliminary studies of 2C oligomerization state by analytical ultracentrifugation (AUC) and secondary structure analysis using circular dichroism (CD). These experiments were started shortly after I had the nanodisc *in vitro* expression system up and running. It was already evident that 2C had the lowest protein yield compared to the other constructs, thus our strategy was to first assess the oligomerization state and secondary structure of the more soluble constructs while I worked to optimize 2C expression. A sedimentation velocity experiment with 2C3A (Fig. 3.8B) revealed that nanodiscs alone are homogeneous at ~3S, whereas 2C3A+nanodiscs are heterogeneous, suggesting the presence of monomers at 4S and oligomers extending to ~30S. Preliminary CD experiments were done using Ub2BC. Fig. 3.8C shows the mean residue ellipticity of Ub2BC+Nanodiscs; however, the concentration of this sample was too low to obtain a strong signal above background noise. Ultimately, the high cost and relatively low protein yield associated with this expression system led us to focus primarily on characterization of the ATPase activity of the various constructs, assays which required only small amounts of protein (as described in Chapter 2), rather than continuing on with these studies.

CHAPTER 4

INVOLVEMENT OF THE POLIOVIRUS 3D^{POL} THUMB α -HELIX IN DETERMINING POLY(A) TAIL LENGTH ²

Authors contributing to this chapter: Fig. 4.1 and the data in Table 4.1 were the work of authors Kempf and Kelly, members of Dave Barton's laboratory at the University of Colorado School of Medicine in Aurora, CO. All other figures and tables in this chapter reflect experiments I carried out.

4.1 Introduction

Picornaviruses have a positive-sense RNA genome that contains a terminal 3'-end poly(A) tail ranging from \approx 20-150 nucleotides in length (132, 133). The poly(A) tail plays an important role in viral replication by serving as the initiation site of negative-strand synthesis (134). The poly(A) tail is also thought to promote circularization of the viral RNA (Fig. 1.8A) through its association with PABP (poly(A)-binding protein), which then interacts with PCBP (poly(rC)-binding protein) and 3CD, bound to the 5' end of the RNA, in order to initiate negative-strand synthesis (55) and has been found to stimulate cap-independent translation from the IRES (135, 136). Deletions or nucleotide substitutions within the poly(A) tail reduce the infectivity of the viral RNA (137, 138) and

² Elements in this chapter pertaining to Fig. 4.1, Table 4.1, Fig. 4.4, and summarized in section 4.3.1 were published as part of **Kempf BJ, Kelly MM, Springer CL, Peersen OB, Barton DJ**. 2013. Structural features of a picornavirus polymerase involved in the polyadenylation of viral RNA. *J. Virol.* **87**:5629-5644

shortening the poly(A) tail less than 12 nucleotides significantly reduces negative-strand RNA synthesis (139). Poly(A)₁₁ tails (consisting of 11 nts) reduces the amount of negative strand RNA synthesized to only ≈3% of that seen with poly(A)₈₀ tails and a poly(A) length of at least 20 nts is needed in order to achieve RNA levels comparable to wild type (139). PABP has been reported to need a minimum of 8 nucleotides in order to bind the poly(A) tail (55), suggesting part of this length requirement is to accommodate the association of PABP. Conversely, depletion of PABP has minimal effects on PV RNA synthesis or viral protein production in HeLa S10 extracts (140), leaving the purpose of the poly(A) tail length still unclear.

Another aspect of the poly(A) tail that has not been well understood is the means by which the viral genome is polyadenylated and how the length of the poly(A) tail is regulated. In 2010, Dave Barton's laboratory reported that the lengths of the negative-strand poly(U) tail and subsequent positive-strand poly(A) tail do not necessarily match that of the initial genomic RNA poly(A) tail and can result in poly(A) and poly(U) stretches longer than the templating sequence (141). To further investigate which regions of the polymerase could play a role in regulating polyadenylation, the Barton laboratory generated a series of point mutations based on the crystal structure of the poliovirus 3D^{pol} elongation complex, targeting residues in the thumb and fingers domain of the polymerase that are important for mediating interactions with the RNA, as shown in Fig. 4.1 (82, 142). The poly(A) tail length and fidelity of these mutants were assayed by sequencing virion RNA recovered from HeLa cells transfected with poliovirus RNA containing 3D^{pol} mutations (142). Interestingly, mutations located in an α -helix of the

thumb domain (N409A, D412A, S416A, L419A, and L420A), which packs into the minor groove of the product-template RNA duplex RNA as it exits the polymerase, had the most significant effects on poly(A) tail length (Fig. 4.1C, (142)). They also characterized the effects of these mutations on polymerase fidelity and observed that the S416A thumb mutant as well as a K133A lysine mutant in the fingers domain both had lower mutation rates than that seen for the previously known G64S high fidelity mutation (89, 90, 142).

Based on these findings, and in collaboration with the Barton laboratory, we wanted to determine whether the altered length of poly(A) transcripts made by these mutant polymerases could be recapitulated with purified polymerase *in vitro* and investigate mechanism by which these mutants are affecting the poly(A) tail length. Along with the five alanine mutants listed above (N409A, D412A, S416A, L419A, and L420A), we also generated seven other 3D^{pol} mutants targeting these sites or at sites that potentially affected replication fidelity (L419A+L420A, D412N, S416T, S416V, K133A, G64S, G64S+S416A). These mutants were tested for their impact on reiterative transcription of homopolymeric templates as well as characterized to determine their effect on 3D^{pol} initiation rate, elongation complex stability, elongation rate, and replication fidelity.

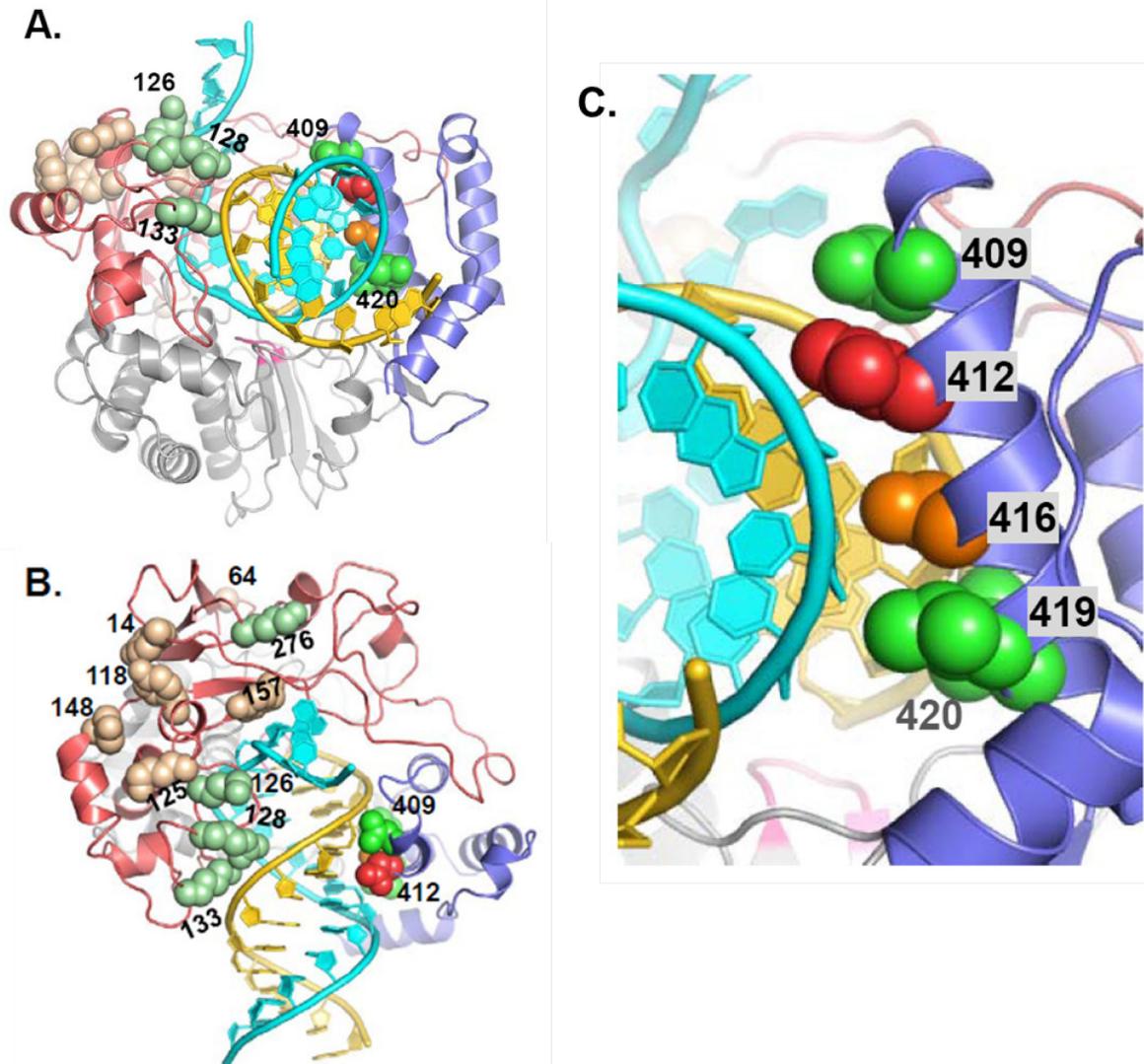


Fig. 4.1. Structural orientation of residues in PV 3D^{pol} involved in the polyadenylation of viral RNA.

Structure of the elongation complex showing the RNA template strand (light blue), the RNA product strand (gold), and the locations of the residues that were mutated by the Barton laboratory to determine their effect on poly(A) tail length. These mutated residues are shown as spheres and colored according to their effects on poly(A) tail length. Residues resulting in unchanged poly(A) tail lengths are shown in tan; those with modest effects are in light green; the thumb domain residues that resulted in longer poly(A) tails are in bright green. D412A, which yielded shorter poly(A) tails, is in red; and S416A, which broadened the poly(A) tail length distribution, is in orange. The three panels show front (A) and top (B) views of the 3D^{pol}-RNA complex and a close-up of the thumb domain helix (C) that inserts into the minor groove of the product RNA duplex. *This figure was published in (142).*

4.2 Materials and Methods

4.2.1 Protein Expression and Purification

3D^{pol} mutants were cloned by site directed mutagenesis into the pET26–UbdH expression vector developed by Craig E. Cameron (111) with the addition of an L446D mutation in 3D^{pol} and a C-terminal GSSS-His₆ tag, as previously described (84). Plasmids were transformed and proteins expressed in *E. coli* BL21 PCG1 cells similar to that described in (84). Briefly, cells were grown overnight at room temperature in LB medium containing 50 µg/mL kanamycin and 34 µg/mL chloramphenicol. 5 mL of the overnight culture was then used to inoculate 1 liter of 2XYT medium with 50 µg/mL kanamycin. The cells were grown at 37°C and shaken at 250 rpm until reaching an OD₆₀₀ of 0.6 to 0.8. The cultures were then transferred to room temperature, isopropyl-β-d-thiogalactopyranoside (IPTG) was added to a final concentration of 0.5 mM, and the cells were grown for an additional 12 to 18 hours with shaking at 250 rpm. The harvested pellet was either stored at -80°C or immediately resuspended in low imidazole buffer (50 mM Tris pH 8.0, 300 mM NaCl, 10 mM imidazole, 20% glycerol, and 0.02% NaN₃) and lysed at 18,000 lb/in² in a model M-110L Microfluidizer (Microfluidics, Newton, MA). The lysate was centrifuged for 40 min at 17,000 rpm in a Sorvall SS-34 rotor and loaded onto a nickel-charged chelating Sepharose fast-flow column (GE Healthcare), followed by step elution with high imidazole buffer (50 mM Tris pH 8.0, 300 mM NaCl, 350 mM Imidazole, 20% glycerol, and 0.02% NaN₃). Fractions containing 3D^{pol} were pooled and diluted 10 fold in low salt buffer (25 mM Tris pH 8.5, 50 mM NaCl, 20% glycerol, and 0.02% NaN₃) so that the final NaCl concentration was 75 mM. The diluted protein was then purified over a HiTrap Q HP column (GE Healthcare) and

eluted over a linear gradient to 1 M NaCl. The resulting 3D^{pol} fractions were combined, concentrated down to < 1 mL, and run over a Superdex 200 gel filtration column (GE Healthcare) equilibrated in 200 mM NaCl, 20% glycerol, 5 mM Tris (pH 7.5), and 0.02% NaN₃. TCEP was added to the purified 3D^{pol} at a final concentration of 10 mM, the protein was concentrated down to 200-600 μM in Amicon 30 kDa MWCO concentrators and stored at -80°C for future use. I would also like to acknowledge Alison Thurston, a rotation student in our lab in the Fall of 2012, for her assistance in the cloning and purification of the various α-helix alanine mutants.

4.2.2 RNA Oligonucleotides and Labeling

The various PETE (polymerase elongation template element) RNA (127) substrates used for the experiments in this chapter are shown in Fig. 4.2. These self-priming RNA templates were synthesized by Integrated DNA Technologies with a 3'-or 5'-fluorescein-label and/or with an amino modified deoxythymidine (dT), which we labeled with IRDye 800RS NHS ester (LI-COR Bioscience) as previously described (84). The RNA hairpins were annealed by heating at 95°C for 5 min followed by chilling on ice for 15 min. The RNA was then aliquoted and stored at -80°C until needed.

4.2.3 Poly(N) Slippage Assays

Unless noted otherwise, reaction mixtures containing 15 μM 3D^{pol}, 1 μM RNA template, 65 μM the individual nucleotide triphosphates (NTPs) (as indicated in the appropriate figures), 50 mM HEPES, pH 6.5, 87.5 mM NaCl, 4 mM MgCl₂, and 5 mM TCEP were mixed on ice for 5 min and then incubated at room temperature for the

elongation reaction. Samples (1 μL) were taken from the reaction mixtures at various time points and quenched with 19 μL of reaction buffer containing 300 mM NaCl and 25 mM EDTA. Gel loading dye (10 μL of a solution containing 95% formamide, 0.5% bromophenol blue, and 0.5% xylene cyanol) was added to the quenched samples. The RNA in each sample was separated by electrophoresis on a 15% polyacrylamide/7 M urea gel and imaged using a LI-COR Odyssey 9120 infrared imager system (142).

4.2.4 Initiation and Stability Assays

Initiation Assays

Initiation and stability assays were carried out similarly to that previously described (84). Briefly, for initiation experiments, 1 μM 9+2-24 RNA was mixed with 15 μM 3D^{pol} , 40 μM each of ATP and GTP in low salt reaction buffer (50 mM HEPES, pH 6.5, 75 mM NaCl, 4 mM MgCl_2 , and 5 mM TCEP) and incubated on ice for 25 minutes to allow the mixture to equilibrate. Next, the reactions were initiated by moving them to room temperature to allow the incorporation of the first two nucleotides. 1 μL samples were taken from the reaction mixtures at various time points over the course of 30 minutes and quenched with 19 μL of quench buffer (i.e. reaction buffer containing 300 mM NaCl and 25 mM EDTA), as demonstrated in Fig. 4.3. Gel loading dye (10 μL of a solution containing 95% formamide, 0.5% bromophenol blue, and 0.5% xylene cyanol) was then added to the quenched samples. The samples were resolved on 15% polyacrylamide–7 M urea PAGE gels and imaged using a LI-COR Odyssey 9120 infrared imager system. To calculate the percent of initiated material at each timepoint,

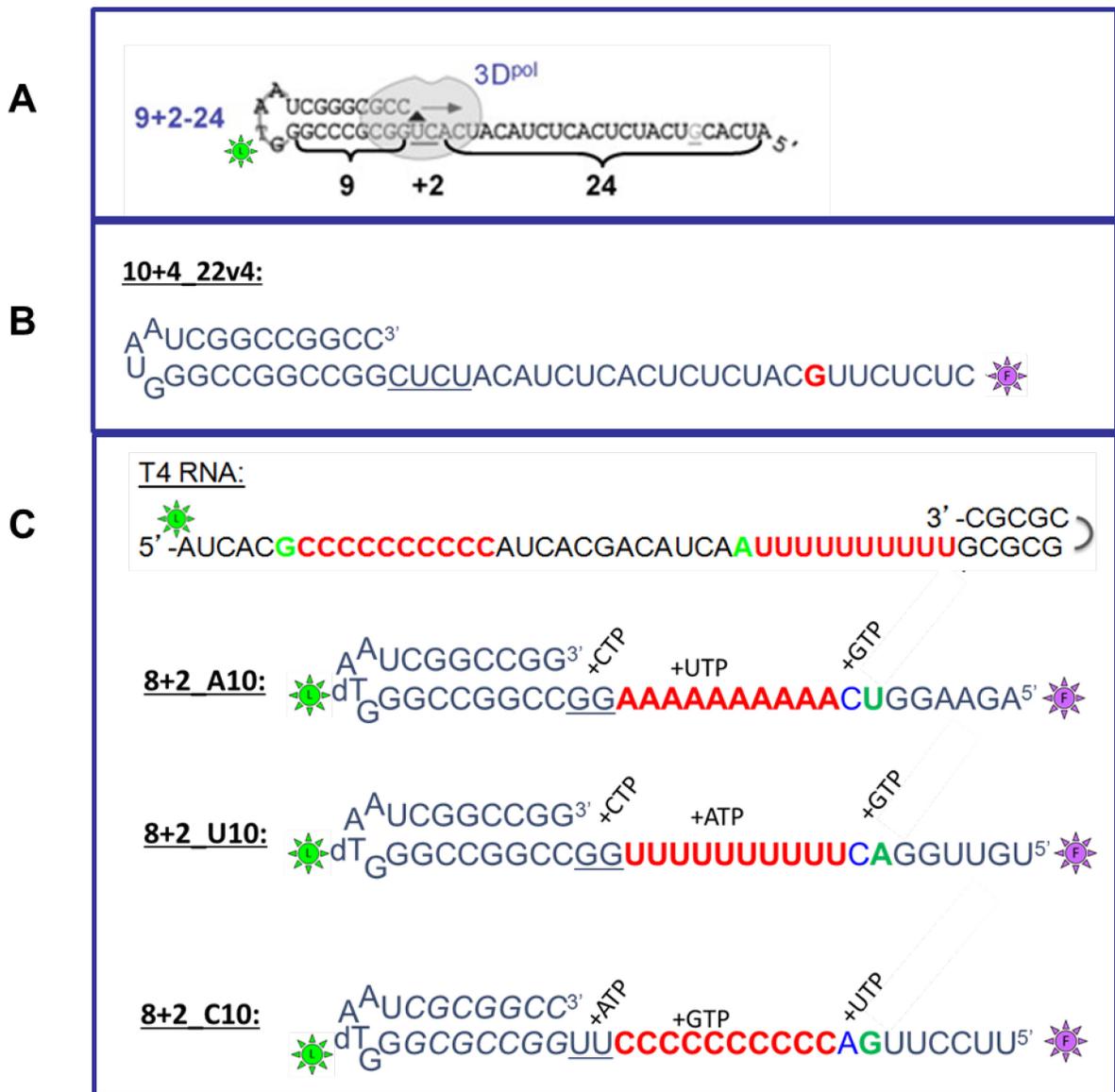


Fig. 4.2. RNA constructs.

A) RNA use for initiation and stability assays. B) RNA used for elongation and fidelity assays. The red G indicates the unique guanine 8 nucleotides from the 5'-end that is the site of 2' deoxyribose CTP misincorporation for fidelity assays. C) RNA constructs used for Poly(N) transcription assays. The location of the fluorescent label and/or Li-COR dye on each RNA is represented by a purple or green star, respectively.

the amount of +2 product was quantified as a ratio of the total RNA (both elongated [+2] and unelongated [U]) in each sample), i.e. $\% \textit{Initiated} = \frac{[+2]}{[U]+[+2]}$.

Stability Assays.

After 30 minutes at room temperature, the remaining reaction mixtures were used for testing complex stability. This was done by diluting the reactions 10-fold into in high salt reaction buffer (50 mM HEPES, pH 6.5, 300 mM NaCl, 4 mM MgCl₂, and 5 mM TCEP) to prevent any protein-RNA complexes that fall apart from re-binding. The dilution mixture also contained 0.1 μM of a different PETE RNA (10+1-12) that was included as a degradation control and to ensure that re-binding and initiation was not occurring (described in (84)). Over the course of 5 hours, samples were taken from the diluted reactions and were “chased” with ATP, GTP, and UTP to determine the fraction of ECs that remained intact and thus able to incorporate an additional 20 nucleotides (+20 products in Fig. 4.3). This was done by mixing 5 μL of the diluted reactions with 5 μL high salt reaction buffer containing 40 μM UTP, 36 μM each of ATP and GTP (so that the final concentration of each the three NTPs was 20 μM). Each chase was allowed to proceed for 3 minutes at which point 10 μL quench buffer was added, followed by 10 μL of gel loading dye. The bands were resolved and imaged in the same manner as described for the initiation assay samples. The stability of the elongation complex formed by each mutant was calculated from the percent of +2 product that was able to form the +20 product, i.e. $\% \textit{Elongated} = \frac{[+20]}{[+20]+[+2]}$. Initiation and stability values reported in this work reflect data obtained from single sets of experiments.

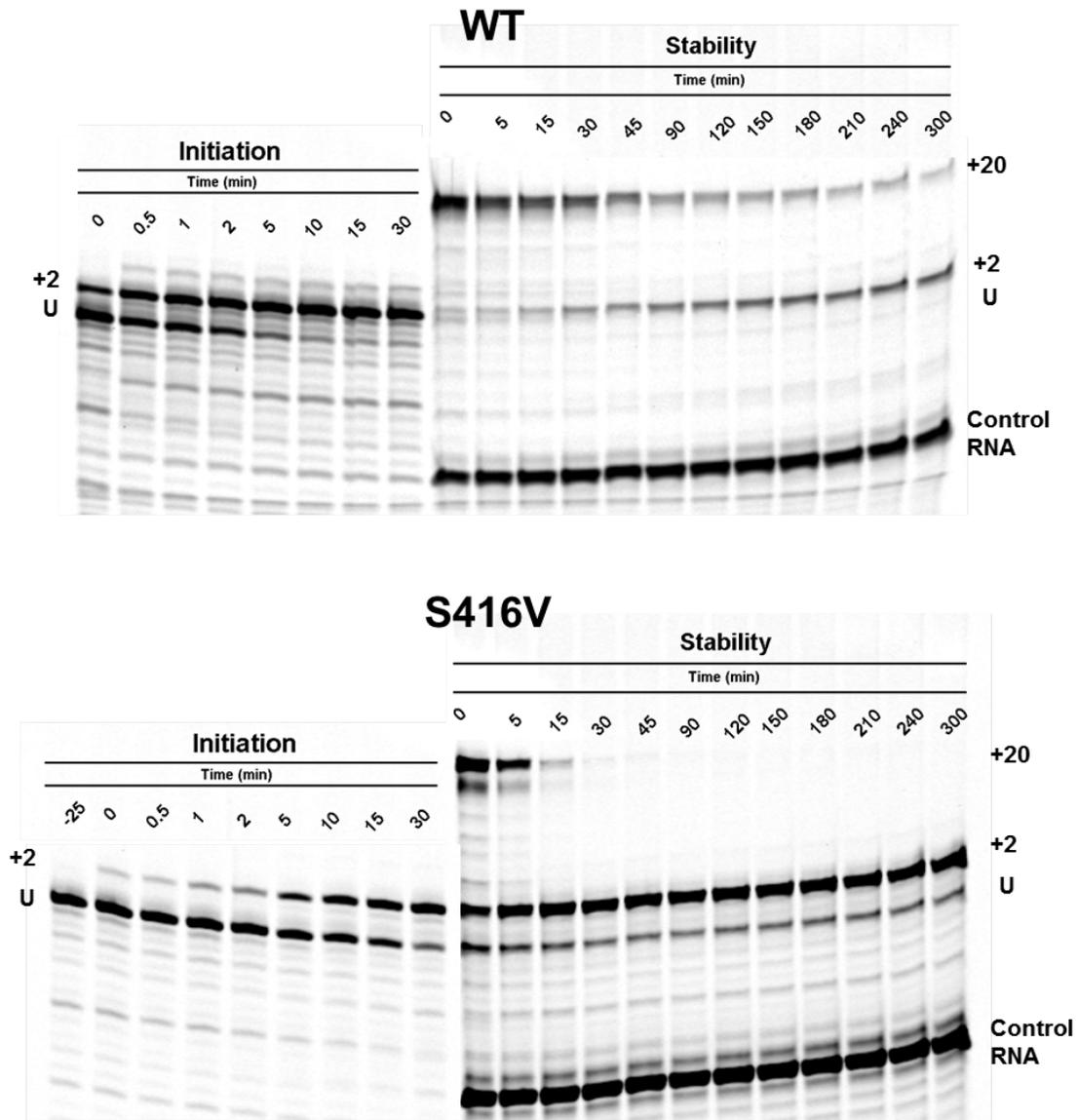


Fig. 4.3. Initiation and stability of the 3D^{pol} elongation complex.

Representative denaturing PAGE gels of initiation and stability assays showing data for wild type (WT) 3D^{pol} and the S416V mutant. To determine initiation rates (left panels), the unelongated (U) 9+2-24 RNA was incubated on ice with 3D^{pol}, ATP, and GTP for 25 minutes (-25 min) before putting at room temperature to initiate the reactions (0 min) and the formation of the +2 product was monitored over 30 minutes. Reactions were then diluted 10 fold into high salt buffer that also contained a degradation and re-initiation control RNA and kept at room temperature for several hours. Samples were periodically taken and chased for 3 minutes with ATP, GTP, UTP to determine the percentage of the +2 EC was still able to elongate to the +20 product.

4.2.5 Stopped-Flow Assays to Determine 3D^{pol} Elongation Rates and Fidelity

3D^{pol} elongation rates and fidelity were determined at 37°C on an Applied Photophysics SX-20 stopped flow instrument similar to that described in (129) using the 10+4_22_v4 5' fluorescein-labeled PETE RNA shown in Fig. 4.2B. Stable +4 elongation complexes were made prior to the stopped-flow assays by incubating 15 μM 3D^{pol}, 5 μM RNA, and 60 μM each of ATP and GTP at room temperature for 15 minutes then stored at -20°C for future use. Immediately prior to elongation or fidelity reactions, the preformed elongation complexes were typically diluted 500-fold (highly unstable mutants, such as S416V, were only diluted 250-fold in order to obtain measurable signal changes) in reaction buffer containing 50 mM HEPES, pH 6.5, 75 mM NaCl, 4 mM MgOAc, and 5 mM TCEP. The reactions were carried out at 37°C and were initiated by rapid mixing in the stopped flow instrument of the diluted elongation complex mixture with an equal volume of reaction buffer containing the appropriate NTPs as well as additional MgOAc (so the final concentration of Mg²⁺ would remain at 4 mM above the total amount of NTPs). The final elongation assay reactions contained 15-30 nM polymerase, 10-20 nM RNA, and 1 μM-2.5 mM NTPs. In the fidelity assays, the reactions contained the same final protein and RNA concentrations but contained a constant final concentration of 20 μM ATP, 1 μM UTP, and 1 μM GTP and either 1 to 300 μM 2'-deoxy CTP (2'-dCTP) or 5 nM to 3 μM CTP, similar to that described in (128). Elongation and fidelity results reported in this work reflect data obtained from single sets of experiments.

4.3 Results

4.3.1 Identification and effect of 3D^{pol} mutations on poly(A) length and fidelity in HeLa cells and in vitro

The Barton laboratory generated 28 mutants containing single or double point mutations, targeting 17 residues within the poliovirus 3D^{pol} that are closely associated in interactions with the RNA, in order to identify potential sites involved in polyadenylation. These sites included tyrosine residues along the template entry channel in the fingers domain; lysine and arginine residues in the fingers domain that are involved in interactions with the phosphate backbone of the RNA; residues within the α -helix of the thumb domain that interact with the product-template duplexed RNA as it exits the active site; and at residue G64 that is involved in replication fidelity, as a Ser mutation at this site is known to increase fidelity as demonstrated by a lower mutational frequency (89, 90, 142). Virion RNA produced by HeLa cells transfected with poliovirus RNA containing 3D^{pol} mutations were sequenced to establish which of the mutations were stably maintained in the viral populations. The sequences of virion RNA from the stable mutations were further analyzed to determine the effects of these mutations on poly(A) tail length. Similarly, the number of mutations that were introduced within the 3D^{pol} region of the virion RNA was also determined in order to assess whether these mutations could potentially be affecting the replication fidelity of the polymerase. The resulting data is summarized in Table 4.1.

16 of the 28 engineered 3D^{pol} mutations were stably maintained in the viral genome, 9 of which generated poly(A) tail phenotypes that differed from wild type when resolved by urea-PAGE (data not shown) (142). The mutants with altered poly(A) tail

Table 4.1. Impact of 3Dpol mutations on poly(A) length and fidelity *in vivo*.

3D ^{pol} group and genotype ³	Length of poly(A) tails ⁴					Replication fidelity ⁵	
	Phenotype	Range	Mean	Median	P value ⁶	No. of mutations	Frequency
Wild type	WT	25–93	52.9	51	NA	8	1.33
<i>Genetically stable with altered poly(A) tails:</i>							
K126A	Shorter	21–71	47.1	47.5	0.0224	5	0.83
R128A	Shorter	23–61	43.8	43	0.0001	10	1.67
K133A	Longer	29–96	58.6	58	0.0443	1	0.17
K276A	Shorter	21–85	46.6	44.5	0.0145	4	0.67
N409A	Longer	25–110	60.2	59	0.0272	3	0.5
D412A	Shorter	21–80	41.9	40.5	<0.0001	4	0.67
S416A	Broader	28–97	53.2	50	<i>0.9251</i>	0	0
L419A	Longer	39–113	65.7	58.5	0.0004	3	0.5
L420A	Longer	28–121	61.7	59	0.0061	9	1.5
<i>Genetically stable with wild type poly(A) tails:</i>							
G64S	WT	30–101	52.2	50	<i>0.7940</i>	2	0.33
K125A	WT	28–98	53.7	49.5	<i>0.7923</i>		
Y14A	WT						
Y118F	WT						
Y148F	WT						
Y157F	WT						
Y118F-Y148F	WT						
<i>Genetically unstable:</i>							
G64T	ND						
G64D	ND						
Y118A	ND						
Y148A	ND						
Y157A	ND						
Y118A-Y148A	ND						
Y118A-Y148F	ND						
Y118F-Y148A	WT ⁷						
K125A-K126A	ND						
K127A	ND						
K127A-R128A	ND						
R188A	ND						

Data presented in this table was the work of Kempf and Kelly published in (142). These results were determined by cDNA synthesis and sequencing of virion RNA purified from HeLa cells that had been transfected with poliovirus RNAs containing 3Dpol mutations.

³ The stability of the 3D^{pol} mutations engineered into the poliovirus genome was determined by cDNA synthesis and sequencing of the 3CD region of virion RNA.

⁴ The nucleotide lengths of poly(A) tails from virion RNA was determined by TOPO-TA cloning and cDNA sequencing.; WT, wild type; ND, not done.

⁵ The 3D^{pol} region (45,000 bases, nucleotides 5860 to 7360) of 30 TOPO-TA clones was sequenced for each virus. The total number of mutations is reported as well in terms of the frequency per genome equivalent (i.e. total number of mutations/7,500 bases).

⁶ P values were calculated in comparison to the result for the wild-type poliovirus using unpaired t test and values that are not statistically different from wild type are shown in red italics. NA, not applicable.

⁷ 118F-Y148A was an unstable mutation that reverted to Y118F-Y148V. The poly(A) tail for this pseudorevertant was wild type.

lengths all mapped to four of the lysine and arginine residues in the fingers domain (K126A, R128A, K133A, K276A) and the five residues in the α -helix in the thumb domain (N409A, D412A, S416A, L419A, L420A), whereas the tyrosine residues in the fingers domain and the G64S mutations did not alter poly(A) tail lengths compared to wild type.

Our laboratory also showed that the thumb helix mutations in purified polymerase generated variable poly(A) transcripts *in vitro*, not only confirming that 3D^{pol} is able to reiteratively transcribe a poly(U) template but also showing poly(A) tail length phenotypes similar to that seen by the Barton Laboratory (Fig. 4.4 and Table 4.1). This was done using the T4 RNA (Fig. 4.2C) containing a poly(U) sequence 10 bases in length, followed by a stretch of heteropolymeric sequence, a poly(C) sequence, and an additional heteropolymeric sequence (Fig. 4.4A). This RNA was initially tested with wild type 3D^{pol} in the presence of different NTP combinations (Fig. 4.4B). In the presence of all four NTPs, wild type 3D^{pol} produced a clearly defined, fully elongated RNA product that was 39 bases longer than the starting template (Fig. 4.4B, +39). When only ATP and UTP were available in the reaction, 3D^{pol} is able to incorporate 12 nucleotides onto the template, corresponding to templating by the U₍₁₀₎ sequence and the two adenosines immediately thereafter, resulting in a heteropolymeric end on the new transcript (Fig. 4.4B, +12). In the presence of ATP only, 3D^{pol} initially incorporates 10 nucleotides across from the U₍₁₀₎ template, but several minutes into the reaction this was followed by a progressively longer smear of poly(A) products ranging in size from 11 to more than 40 bases in length (Fig. 4.4B, ATP-only reaction). These poly(A) products, that are clearly longer than the poly(U)₁₀ template sequence, demonstrate that

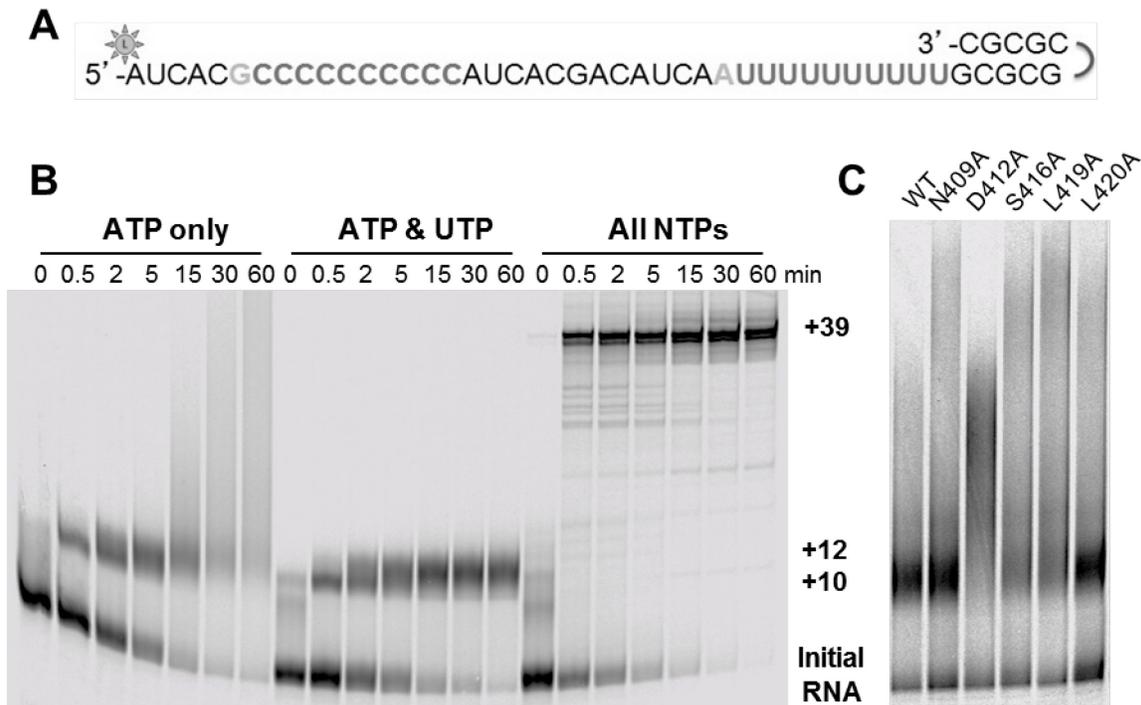


Fig. 4.4. Reiterative transcription of a poly(U) sequence by purified 3D^{pol}.

(A) Diagram of the RNA template. The RNA is labeled at its 5'-end with an infrared dye so that both starting material and products can be imaged on gels. A 5-base self-priming hairpin is present at the 3'-end, followed by a 10-base-long poly(U) sequence. Beyond the poly(U) sequence are 13 bases of heteropolymeric sequence, a 10-base-long poly(C) sequence, and 6 bases of additional heteropolymeric sequence. (B) RNA products synthesized by wild-type 3Dpol in reaction mixtures containing ATP only resulting in an initial +10 product followed by an increasingly longer product over the course of an hour, ATP and UTP to generate a +12 product, and all four NTPs to form a +39 product. Reaction mixtures containing purified 3Dpol (15 μ M), RNA template (1 μ M), and the indicated NTP(s) (65 μ M each) were incubated for 0 to 60 min. RNA products from the reactions were quenched by addition of EDTA and high salt, fractionated by denaturing gel electrophoresis, and visualized by scanning of the infrared label on the RNA. Numbers to the right of the gel are in bases. (C) RNA products synthesized by wild-type and mutant 3D^{pol} in reaction mixtures containing ATP (65 μ M) after 30 min of incubation. *This experiment was completed by me, Courtney Springer, and this figure is published in (142).*

3D^{pol} is able to carry out reiterative transcription when stalled at the end of a poly(U) homopolymer region due to lack of the next nucleotide needed. When this ATP-only reaction was repeated with the 3D^{pol} thumb α -helix mutants, N409A, D412A, S416A, L419A, and L420A, they clearly alter the amounts of poly(U)-dependent reiterative transcription compared to that for wild-type 3D^{pol} (Fig. 4.4C). The N409A, S416A, L419A, and L420A mutations resulted in poly(A) products longer than those synthesized by wild-type 3D^{pol}, whereas the D412A mutation resulted in a shorter poly(A) product distribution (Fig. 4.4C), closely reflecting the poly(A) tail phenotypes observed for these mutants *in vivo* (Table 4.1). These data show that 3D^{pol} can reiteratively transcribe poly(U) sequences *in vitro* and that the N409A, D412A, S416A, L419A, and L420A mutations influence the magnitudes of reiterative transcription on poly(U) sequences.

Sequencing of the poly(A) tails from virion RNA produced by these mutants by the Barton laboratory confirmed these findings, with alanine mutations at residues K126, R128, K133, K276, N409, D412, L419, and L420 generating poly(A) tail lengths that differed statistically significantly from the lengths observed by wild type (Table 4.1, size of poly(A) tail, range, mean, median, and p value). The one exception was S416A, which generated a broader range of poly(A) tail lengths when the products were resolved by PAGE (Fig. 4.4C, and Figure 5 as published in (142) but not included in this thesis), but did not generate poly(A) lengths of statistically significant difference from wild type.

To assess whether these mutations had an impact on replication fidelity, the Barton laboratory sequenced the 3D^{pol} region (nucleotides 5860 to 7360) of virion RNA from a total of 30 clones generated by each mutant (45,000 bases total per mutant), corresponding to an equivalent of six genomes (7,500 bases per PV genome). The

total number of mutations observed from wild type 3D^{pol}, the nine poly(A) mutants, and the high fidelity G64S mutant as well as the mutational frequency (mutations per genome equivalent) is listed in Table 4.1 (replication fidelity). As expected, G64S had a lower mutational frequency than wild type, of 0.33 for G64S and 1.33 for wild type. Interestingly, there were no mutations found in the sequencing data that was collected from the virions produced by S416A. The K133A mutation also produced fewer mutations than G64S, having a mutational frequency of only 0.17. These data suggest that the S416A and K133A mutations could also be affecting fidelity as well as poly(A) length.

4.3.2 Mutations Made in the Thumb Domain of Purified PV 3D^{pol}

Based on findings described above, we generated several point mutations in the helix of the PV 3D^{pol} thumb domain to assess how these mutations may be involved in mediating poly(A) transcription as well as at sites potentially affecting replication fidelity. As discussed in section 1.4, the crystal structure of the poliovirus 3D^{pol} elongation complex shows that this α -helix in the thumb and a loop in the pinky domain come together to form a “clamp” around the product-template duplexed RNA as it leaves the polymerase (Fig. 1.11) (82). Residues N409, D412, S416, L419, and L420 lie on the face of this α -helix in the thumb and pack into the minor groove of the dsRNA (82). We initially generated alanine point mutations at these 5 locations as well as a double L419/L420 mutation (N409A, D412A, S416A, L419A, L420A, and L419/420A). Based on our initial characterization of these mutants, we also made additional mutations at D412 (D412N) and S416 to (S416T, S416V) designed to introduce more systematic

changes in their amino acid character. Additionally, we cloned K133A that was identified by the Barton laboratory (along with S416A) as a potential fidelity mutant, G64S that is a known high fidelity mutant, and a double mutant containing both G64S and S416A to see if this would result in an additive fidelity effect, if S416A was indeed a high fidelity mutant.

4.3.3 Effect of Mutations on Poly(N) Transcripts

Design of PETE RNAs for poly(N) transcription assays

Our initial poly(A) assay using the T4 RNA and the original five alanine thumb mutants demonstrated that these mutations were able to reiteratively transcribe a poly(U) template, as seen in Fig. 4.4C. The T4 RNA was very useful for demonstrating phenotypic difference in the poly(A) products between the different mutants, as seen in Fig. 4.5 that shows the full 120 min time course poly(N) slippage experiment for which only the 30 min timepoints are present on the gel in Fig. 4.4C . However, it was very difficult to quantify the differences between the mutants as the T4 RNA appeared as a broad, ill-defined band on PAGE gels, most likely reflecting multiple energetically favorable conformations or significant variations in length causing it to appear as a smear (Fig. 4.5). We were also concerned this RNA did not have any unique “lock” sequence prior to the poly(U) stretch that would allow us to form a stable 3D^{pol}-RNA elongation complex (discussed in section 1.4.2) by incorporating the first several nucleotides. In other words, we were worried that if the polymerase immediately started elongating on the homopolymer sequence, we may observe product formation difference between the mutants that were a reflection of altered initiation rates rather

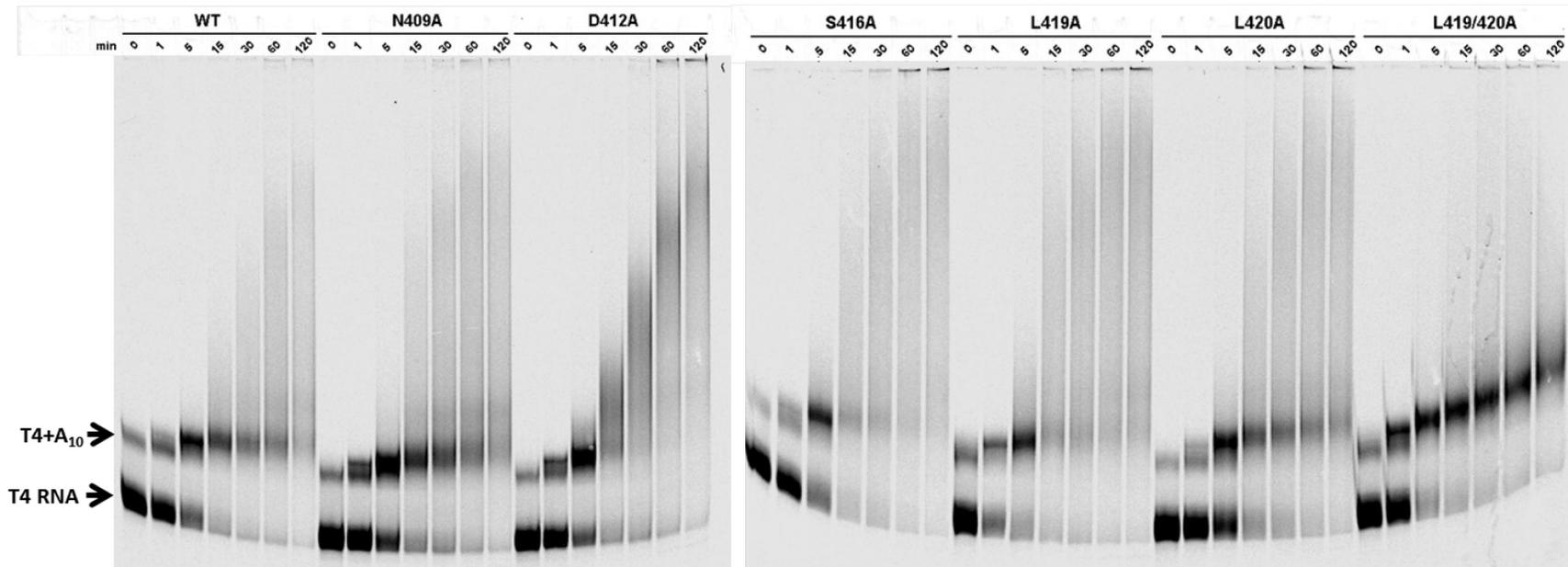
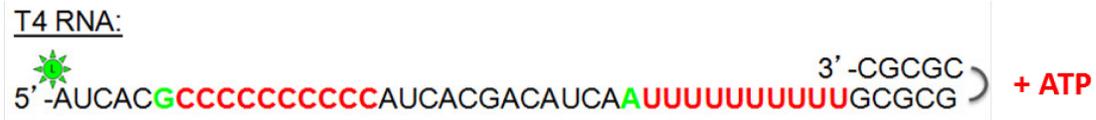


Fig. 4.5. Poly(N) slippage assay with alanine α -helix thumb 3D^{pol} mutants and the T4 RNA.

Denaturing PAGE gel showing the full 120 minute time course experiment that is described in Fig. 4.4C demonstrating the mobility of the T4 RNA that results in histogram-like smears. The unelongated T4 RNA is indicated on the gel as well as the T4+A₁₀ species in which only ten adenosines have been transcribed from the T4 poly(U)₁₀ template.

than a direct result of difference in poly(A) transcription. In addition to designing an RNA that would address these concerns, we wanted to further test these mutant polymerases on a poly(A) template, as previous studies have shown that poly(U) sequences longer than their poly(A) template can be generated during viral replication (141), as well as find out whether this phenomenon is specific to poly(U) and poly(A) sequences. Thus, we designed three PETE RNAs, shown in Fig. 4.2C, consisting of a 8 bp hairpin duplex at the 3'-end followed by a two nucleotide single stranded "lock" region, thus they are designated as '8+2'. Immediately 5' of this 8+2 portion is a 10 nucleotide A, U, or C homopolymer stretch (A10, U10, and C10), followed by a unique "post-lock" base, and finally a 7mer heteropolymer sequence at the 5'-end (Fig. 4.2C).

Wild type and mutant 3D^{pol} reiteratively transcribe poly(A) and poly(U) but not poly(C)

RNA templates

These RNAs were initially tested with wild type 3D^{pol} that had been pre-locked to form the stable elongation complex by incubating the RNA and polymerase with the first NTP on ice for 5 minutes, forming the +2 product. The reactions were then added to buffer alone, buffer containing only the poly(N)₁₀ cognate nucleotide, or buffer containing all NTPs. The reactions were then incubated at room temperature for 30 minutes, at which point they were quenched and run on a 7M Urea/15% polyacrylamide gel alongside samples of just the RNA with no protein added. The resulting gel is shown in Fig. 4.6 along with an image of the three RNAs for reference. When just the first NTP is included in the reaction, there is a shift from the unelongated template (+0) to the +2 product for all the RNAs. When ATP and GTP are present in the reactions

containing 3D^{pol} and C10 RNA, there is a clearly defined +12 product band corresponding to the incorporation of two ATP bases and ten GTP bases, indicating that wild type 3D^{pol} is not able to reiteratively transcribe a poly(C) template. In the presence of the all NTPs, there is a distinct band reflecting the fully elongated product in which 20 nucleotides have been incorporated (FE₊₂₀) on the C10 RNA. Similarly, there is a clear FE₊₂₀ product band that forms when the reactions containing wild type polymerase and A10 or U10 RNA are provided all NTPs. Notably, the fully elongated products from A10 and U10 have a slower mobility on the gel than that seen for the fully elongated C10 product, which we attribute to the high GC content of the C10 product acting to prevent the RNA from fully denaturing in the gel.

The most intriguing observation from this experiment is the formation of distinct products transcribed from the A10 and U10 RNAs in the presence of both the locking nucleotide and homopolymer nucleotide that are longer than the poly(N)₁₀ template. In these reactions (Fig. 4.6, A10 RNA +CU, U10 RNA +CA), the expected +12 product is present (supported by the presence of ≈10 bands in the lane between the +2 and +12 products), but several other distinct bands are evident as well. In the reaction containing 3D^{pol}, A10 RNA, CTP, and UTP there is a clearly discernible ladder above the +12 band, with each band presumably reflecting the change in size associated with the incorporation of an additional nucleotide. Thus, the two predominant bands above the +12 band would represent a +13 product and, counting 6 bands up, a +19 product, corresponding to transcription of a poly(U) sequence ranging from 11 to 17 bases long from the poly(A)₁₀ template. Faint laddering continues above the +19 product, showing a small amount of a +20 product and perhaps larger products that fade into a smear.

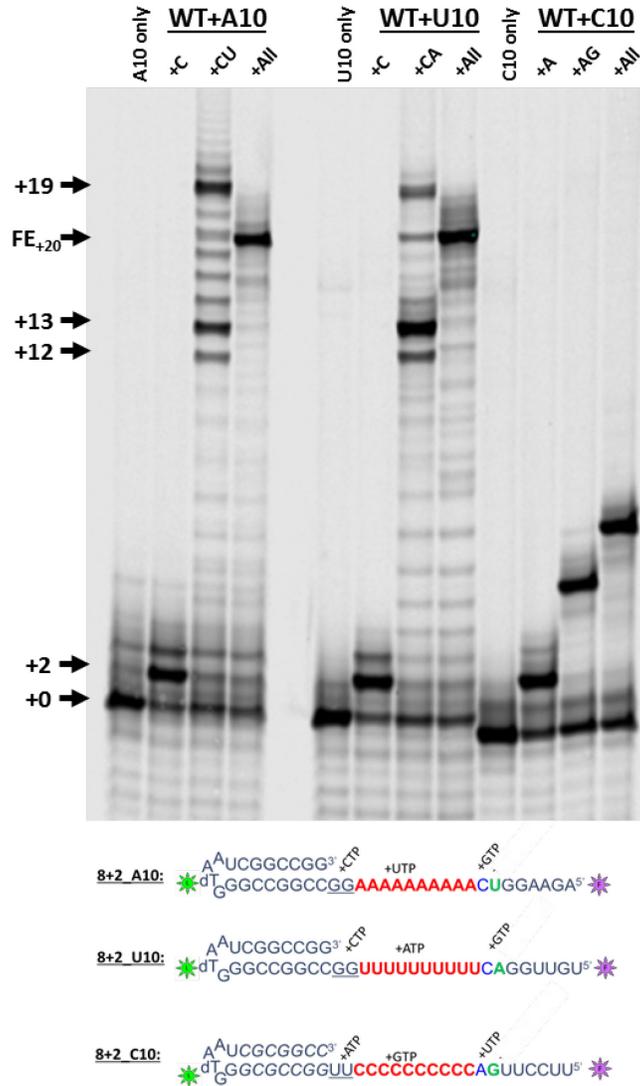


Fig. 4.6. Testing 8+2_A10, U10, and C10 RNA with wild type 3D^{pol}.

Wild type (WT) polymerase was mixed with either A10, U10, or C10 RNA (shown below gel) and pre-locked by including the appropriate nucleotide (+CTP for A10 and U10; +ATP for C10) in order to generate a stable elongation complex in which the first two nucleotides were incorporated. These were mixed with buffer containing either no additional NTPs (resulting in a +2 product), the poly(N)₁₀ homopolymer cognate nucleotide, or all the NTPs (generating the fully elongated, FE₊₂₀, product) and the reactions were incubated at room temperature for 30 minutes. Addition of just the homopolymer nucleotide resulted in the expected +12 product in all the reactions as well as a possible +13 to +19 product in the reactions containing the A10 and U10 RNA (for which the band can more easily be counted from the laddering in Fig. 4.7 and Fig. 4.9). To the left of these three samples for each RNA is lane containing the unelongated RNA (+0). The reactions contained a final concentration of 15 μM wild type 3D^{pol}, 1 μM RNA, and 65 μM of the each NTP present in the reaction.

Although there is not as clearly defined step-size laddering observed for the reaction containing the U10 RNA with ATP dependent elongation, the major products are very similar to those observed for the A10 reaction with a very prominent +13 band, an accumulated +19 product, several intermediate bands, and a faint +20 band. Again, we believe RNA secondary structural effects arising from the extended, non-cognate, homopolymeric sequence is responsible for the difference in mobility between the poly(U/A) products and the fully elongated +20 products (FE₊₂₀) products observed for the A10 and U10 RNAs.

To determine the impact of our mutants on reiterative transcription of the poly(N) sequences in the A10, U10, and C10 RNAs, we repeated the experiment described above in which the elongation complex is pre-formed in the presence of the first nucleotide on ice for 5 minutes and then mixed with the homopolymer cognate NTP and incubated at room temperature. We did this using eight of the mutant polymerases, D412A, S416A, G64S, K133A, D412N, S416T, S416V, and G64S/S416A, alongside wild type. Samples were taken at various time points after adding the second NTP at room temperature, and the samples taken at 15 and 30 minutes are shown in Fig. 4.7. As was seen with wild type, the reactions carried out with the C10 RNA template in the presence of ATP and GTP form a single band resulting from the formation of the expected +12 product (Fig. 4.7C) for all the proteins. When the A10 or U10 RNA template is used (Fig. 4.7A,B) there are again several products that are longer than the expected +12 product, with the most substantial bands corresponding to a +13 and +19 product, similar to that previously seen with wild type 3D^{pol} (panels to the left of the A10 and U10 gels are from Fig. 4.6 for reference).

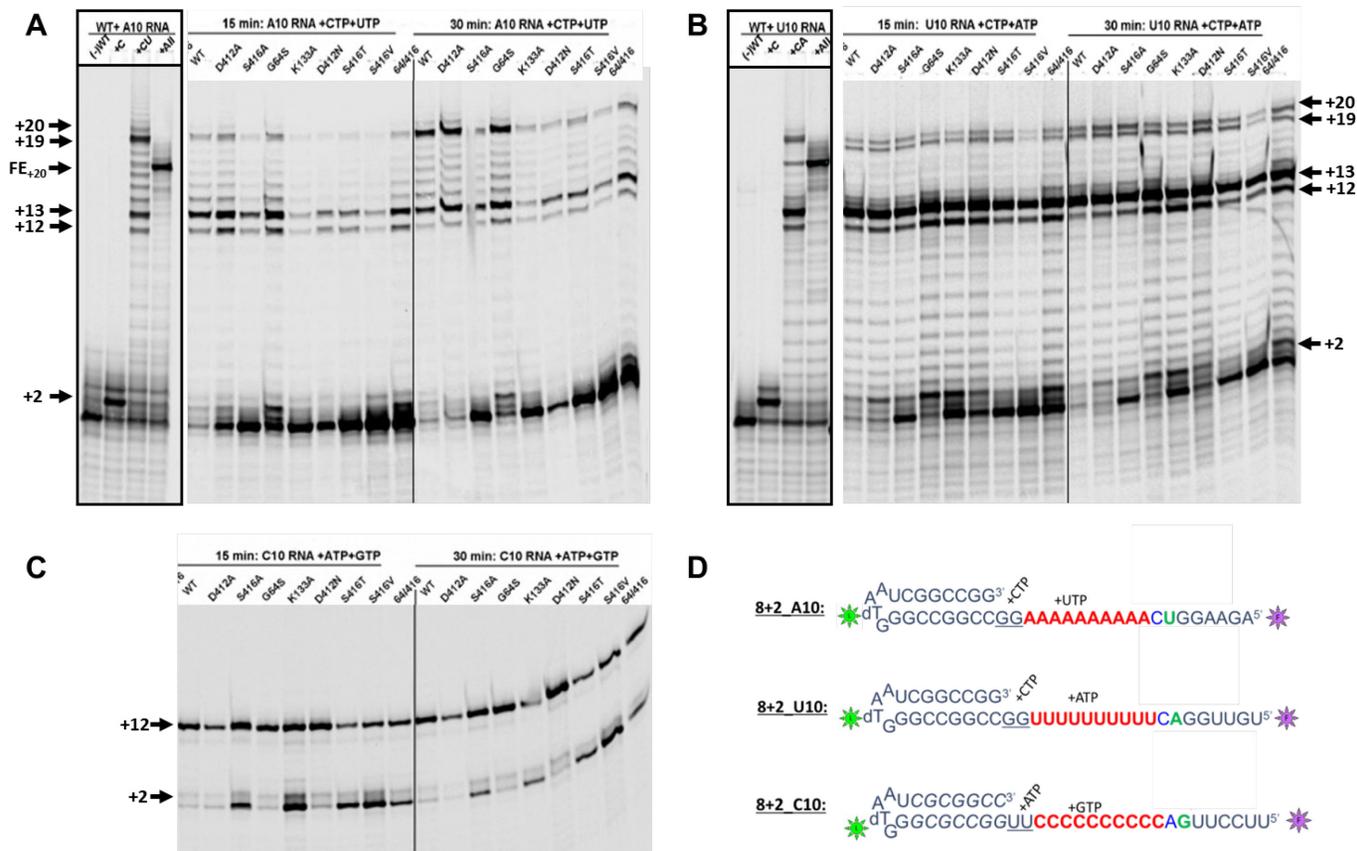


Fig. 4.7. Testing 8+2_A10, U10, and C10 RNA with 3D^{pol} mutants.

Elongation complexes were pre-formed by mixing mutant polymerase with either (A) A10, (B) U10, or (C) C10 RNA, and the appropriate pre-locking nucleotide, as indicated above the two underlined bases in the image of the RNAs in (D). The ECs were then mixed with the poly(N)₁₀ homopolymer cognate nucleotide, as indicated above the red bases in (D), and incubated at room temperature. Samples were taken 15 minutes and 30 minutes later and run on a denaturing PAGE gel. The resulting +12 to +20 products are indicated by arrows. The outlined panels to the left of the gels in (A) and (B) are from Fig. 4.6 for comparison of the various products formed by wild type 3D^{pol} in the presence of different nucleotides. The reactions contained a final concentration of 15 μM wild type 3D^{pol}, 1 μM RNA, and 65 μM of the indicated NTP.

Table 4.2. Quantitation of poly(A) and poly(U) transcripts.⁸

Mutant	Time: Band	8+2_A10 RNA					8+2_U10 RNA				
		15 min		30 min		30-15 min	15 min		30 min		30-15 min
		Intensity	% of Lane ⁹	Intensity	% of Lane		Intensity	% of Lane	Intensity	% of Lane	
WT	+0	445000	4.63%	348000	3.93%	-0.70%	280000	2.87%	242000	2.64%	-0.23%
	+2	397000	4.13%	333000	3.76%	-0.37%	297000	3.05%	210000	2.29%	-0.75%
	+12	405000	4.21%	371000	4.19%	-0.02%	362000	3.71%	309000	3.37%	-0.34%
	+13	788000	8.20%	626000	7.07%	-1.13%	3460000	35.49%	3320000	36.24%	0.76%
	+19	379000	3.94%	702000	7.93%	3.99%	220000	2.26%	266000	2.90%	0.65%
	+20	226000	2.35%	251000	2.84%	0.48%	197000	2.02%	249000	2.72%	0.70%
	Lane ¹⁰	9610000		8850000			9750000		9160000		
D412A	+0	853000	8.81%	581000	5.59%	-3.23%	414000	4.10%	287000	2.73%	-1.37%
	+2	416000	4.30%	341000	3.28%	-1.02%	426000	4.22%	316000	3.01%	-1.21%
	+12	539000	5.57%	434000	4.17%	-1.40%	473000	4.68%	427000	4.07%	-0.62%
	+13	1010000	10.43%	984000	9.46%	-0.97%	3280000	32.48%	4070000	38.76%	6.29%
	+19	393000	4.06%	811000	7.80%	3.74%	254000	2.51%	339000	3.23%	0.71%
	+20	210000	2.17%	256000	2.46%	0.29%	241000	2.39%	377000	3.59%	1.20%
	Lane	9680000		10400000			10100000		10500000		
S416A	+0	2380000	22.04%	1990000	20.12%	-1.92%	1370000	13.30%	904000	8.69%	-4.61%
	+2	424000	3.93%	412000	4.17%	0.24%	357000	3.47%	324000	3.12%	-0.35%
	+12	385000	3.56%	365000	3.69%	0.13%	447000	4.34%	445000	4.28%	-0.06%
	+13	535000	4.95%	502000	5.08%	0.12%	2610000	25.34%	3360000	32.31%	6.97%
	+19	305000	2.82%	379000	3.83%	1.01%	248000	2.41%	304000	2.92%	0.52%
	+20	229000	2.12%	218000	2.20%	0.08%	234000	2.27%	368000	3.54%	1.27%
	Lane	10800000		9890000			10300000		10400000		
G64S	+0	959000	8.96%	597000	5.63%	-3.33%	417000	3.18%	373000	3.03%	-0.15%
	+2	684000	6.39%	605000	5.71%	-0.68%	823000	6.28%	508000	4.13%	-2.15%
	+12	582000	5.44%	474000	4.47%	-0.97%	959000	7.32%	783000	6.37%	-0.95%
	+13	1120000	10.47%	933000	8.80%	-1.67%	4720000	36.03%	4900000	39.84%	3.81%
	+19	393000	3.67%	867000	8.18%	4.51%	258000	1.97%	299000	2.43%	0.46%
	+20	211000	1.97%	269000	2.54%	0.57%	201000	1.53%	244000	1.98%	0.45%
	Lane	10700000		10600000			13100000		12300000		
K133A	+0	2280000	21.31%	1660000	16.87%	-4.44%	2100000	17.80%	1550000	13.25%	-4.55%
	+2	472000	4.41%	376000	3.82%	-0.59%	841000	7.13%	693000	5.92%	-1.20%
	+12	381000	3.56%	340000	3.46%	-0.11%	904000	7.66%	906000	7.74%	0.08%
	+13	441000	4.12%	381000	3.87%	-0.25%	2280000	19.32%	2750000	23.50%	4.18%
	+19	303000	2.83%	386000	3.92%	1.09%	216000	1.83%	296000	2.53%	0.70%
	+20	218000	2.04%	221000	2.25%	0.21%	182000	1.54%	274000	2.34%	0.80%
	Lane	10700000		9840000			11800000		11700000		
D412N	+0	1160000	14.52%	624000	7.33%	-7.19%	790000	6.42%	482000	3.68%	-2.74%
	+2	343000	4.29%	267000	3.14%	-1.16%	443000	3.60%	392000	2.99%	-0.61%
	+12	380000	4.76%	325000	3.82%	-0.94%	772000	6.28%	703000	5.37%	-0.91%
	+13	393000	4.92%	497000	5.84%	0.92%	4280000	34.80%	5010000	38.24%	3.45%
	+19	243000	3.04%	371000	4.36%	1.32%	312000	2.54%	389000	2.97%	0.43%
	+20	175000	2.19%	214000	2.51%	0.32%	232000	1.89%	368000	2.81%	0.92%
	Lane	7990000		8510000			12300000		13100000		
S416T	+0	3130000	26.53%	2610000	23.51%	-3.01%	1820000	18.20%	1200000	12.05%	-6.15%
	+2	542000	4.59%	440000	3.96%	-0.63%	310000	3.10%	297000	2.98%	-0.12%
	+12	452000	3.83%	430000	3.87%	0.04%	438000	4.38%	461000	4.63%	0.25%
	+13	520000	4.41%	602000	5.42%	1.02%	2270000	22.70%	2540000	25.50%	2.80%
	+19	294000	2.49%	451000	4.06%	1.57%	266000	2.66%	334000	3.35%	0.69%
	+20	209000	1.77%	223000	2.01%	0.24%	211000	2.11%	290000	2.91%	0.80%
	Lane	11800000		11100000			10000000		9960000		
S416V	+0	4590000	33.50%	3410000	29.15%	-4.36%	3410000	31.87%	2160000	23.43%	-8.44%
	+2	702000	5.12%	633000	5.41%	0.29%	405000	3.79%	319000	3.46%	-0.33%
	+12	313000	2.28%	308000	2.63%	0.35%	435000	4.07%	336000	3.64%	-0.42%
	+13	416000	3.04%	409000	3.50%	0.46%	1290000	12.06%	1290000	13.99%	1.94%
	+19	296000	2.16%	325000	2.78%	0.62%	160000	1.50%	183000	1.98%	0.49%
	+20	217000	1.58%	206000	1.76%	0.18%	177000	1.65%	154000	1.67%	0.02%
	Lane	13700000		11700000			10700000		9220000		
G64S/S416A	+0	3720000	27.76%	2490000	21.65%	-6.11%	2130000	15.66%	1230000	10.25%	-5.41%
	+2	800000	5.97%	565000	4.91%	-1.06%	743000	5.46%	640000	5.33%	-0.13%
	+12	483000	3.60%	448000	3.90%	0.29%	763000	5.61%	786000	6.55%	0.94%
	+13	793000	5.92%	758000	6.59%	0.67%	3600000	26.47%	3350000	27.92%	1.45%
	+19	370000	2.76%	511000	4.44%	1.68%	225000	1.65%	346000	2.88%	1.23%
	+20	214000	1.60%	209000	1.82%	0.22%	266000	1.96%	342000	2.85%	0.89%
	Lane	13400000		11500000			13600000		12000000		

Footnotes shown on next page

⁸ Quantitation of the +0, +2, +12, +19 and +20 RNA bands from the 8+2_A10 and 8+2_U10 gels shown in Fig. 4.11A&B. The rows containing the +13 product values, corresponding to the first apparent poly(N) slippage event, for each mutant have been shaded in red to.

⁹ This value represents the intensity of the indicated RNA band divided by the intensity of the lane.

¹⁰ The lane intensity represents the intensity measured for the portion of the lane ranging from, and including, the +0 to the +20 bands. The portions of the lane above and below this region were not included in this measurement.

The gels in Fig. 4.7A,B were further analyzed to determine the intensities of the unelongated (+0), +2, +12, +13, +19, and +20 RNA bands as a fraction of the total intensity of this portion of the lane (i.e. the total intensity of the lane ranging from +0 to +20) for each mutant at 15 and 30 minutes. As shown in Table 4.2, these values were then used to determine relative amount of product by calculating the percent of each band in relation to the lane (% of Lane). The relative difference in the amount of each product between 15 and 30 minutes is also provided in Table 4.2 (30-15 min). *Analysis of these data suggest that the incorporation of an extra nucleotide on the U10 template occurs fairly rapidly to generate the +13 poly(A)₁₁ product, at which point an energy barrier must be overcome before there is a rapid incorporation of an additional six nucleotides to form the +19 poly(A)₁₇ product.* In all the reactions there is a decrease in the amount of unelongated RNA (+0) between 15 to 30 minutes, suggesting that initiation is still occurring and/or the unelongated material is being degraded. Not surprisingly, larger amounts of unelongated material are observed in the presence of mutants with more severe initiation and stability defects, with S416V having the largest percentage unelongated (+0) material. Although the other difference between the mutants for each RNA are minor, there are variable trends that emerge in the presence of the 8+2_U10 and 8+2_A10. In the presence of both the A10 and U10 RNA, the +13 product is the most abundant elongated (+12 to +20) product, although there is a much larger amount of +13 material in the U10 reactions than in the A10 reactions, as demonstrated in Fig. 4.8. Between 15 and 30 minutes, the amount of +13 material increases in all the U10 reactions and in the A10 reactions carried out with the S416A, D412N, S416T, S416V and G64S/S416A mutants, whereas the amount of +13 material

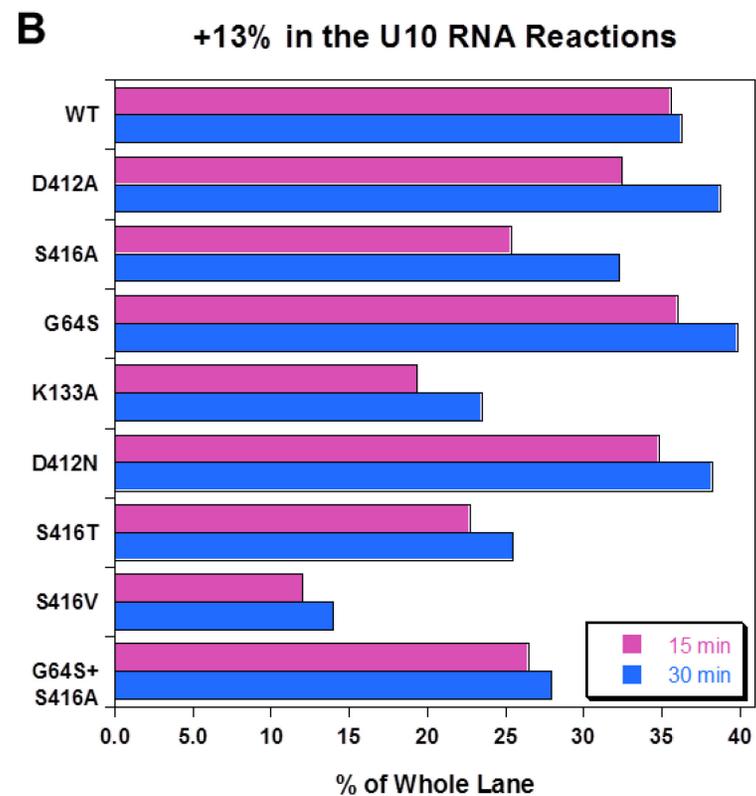
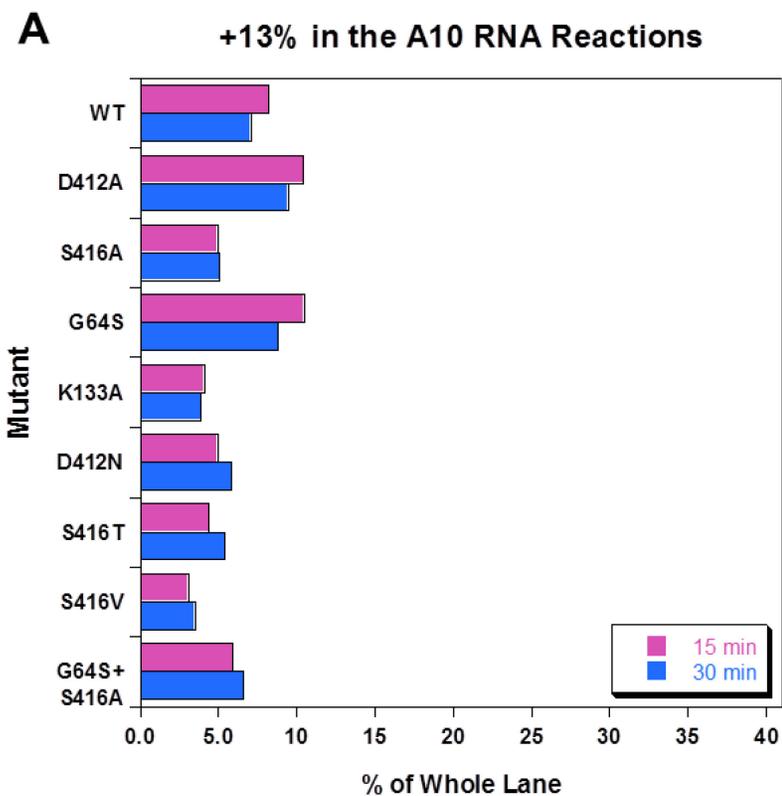


Fig. 4.8. Histogram of +13 band in A10 and U10 RNA reactions.

Histogram plots of the percent of +13 product values listed in Table 4.2 for the reactions containing wild type or mutant polymerase and either the (A) 8+2_A10 RNA or (B) 8+2_U10 RNA at 15 minutes (pink) and 30 minutes (blue).

decreases in the A10 reactions containing wild type 3D^{pol} and the D412A and G64S mutants (Fig. 4.8). In the reactions containing the U10 RNA, there is a significant buildup of +13 poly(A)₁₁ product that ranges from ≈3-fold to 20-fold more abundant than the other elongated products. The +13 product also displays the largest increase between 15 to 30 minutes in all the U10 reactions and only minor increases are observed for the +19 and +20 products during this timeframe. In the reactions containing A10 RNA, the amount of +13 product is only slightly greater than the elongated (+12 to +20) products, and is actually equivalent to or lower than the initiated (+2) material in the presence of the S416T, S416V, and G64S/S416A 3D^{pol} mutants. Interestingly, the second most abundant elongated (+12 to +20) product in the A10 RNA reactions is the +12 poly(U)₁₀ product at 15 minutes and the +19 poly(U)₁₇ product at 30 minutes, with the overall largest relative increase observed for the +19 product for all mutants between 15 and 30 minutes. Finally, this appears to be followed by another, fairly rapid addition of one A, resulting in the +20 poly(A)₁₈ product. A similar process seems to occur on the poly(A)₁₀ template, although the main energy barrier appears to be related to the incorporation of the single additional nucleotide to produce the +13 and +19 products, as demonstrated by a reduction in accumulated +13 poly(U)₁₁ product at 15 minutes followed by a more significant increase in the +19 poly(U)₁₇ product at 30 minutes, as compared to the U10 RNA. Similarly, the relative increase of +20 poly(U)₁₈ product generated during the 15 minute span is at least 3-fold less than the relative +19 poly(U)₁₇ product, whereas the relative increase of the +20 poly(A)₁₈ product during this time is similar to or greater than the amount of +19 poly(A)₁₇ product, with the exception of the S416V mutant which could be due to the very low stability of this mutant.

In summary, these data show that 3D^{pol} is able to rapidly generate poly(A) and poly(U) transcripts that are longer than the 10 nucleotide homopolymer sequence from the U10 and A10 RNAs, but longer poly(G) transcripts are not transcribed from the C10 RNA. The products that were generated from the A10 and U10 RNAs are longer than the 12 base template (i.e. the 2 nt pre-lock sequence followed by the poly(A/U)₁₀ sequence) and represent a +13 poly(N)₁₁ product, a +19 poly(N)₁₇ product, and a +20 poly(N)₁₈ product. The relative amounts of the +19 and +20 products typically increased between 15 to 30 minutes in both the U10 and A10 RNA reactions; however the rates/proportions at which these products accumulate differ between the two RNAs, suggesting there may be differences in the energetics/kinetics when 3D^{pol} is reiteratively transcribing a poly(U) versus a poly(A) template.

Poly(A) slippage is time and ATP concentration dependent

To determine whether production of these longer poly(A) products were dependent on ATP concentrations, we carried out a poly(A) slippage assay using the 8+2_U10 RNA and D412N 3D^{pol} mutant in the presence of 15, 30, and 60 μM ATP (Fig. 4.9). We initially decided to use the D412N mutant as it appeared in to be one of the more efficient of the α-helix mutants to produce poly(U) products longer than the template from the experiment shown in Fig. 4.7B were curious as to whether it would generate transcripts longer than the +20 product over a longer timecourse. These reactions were pre-locked with CTP at room temperature for 5 minutes and a sample was taken just before and after the 5 min incubation (-5 and 0, respectively). The reactions were then added to the appropriate NTP solutions at room temperature and

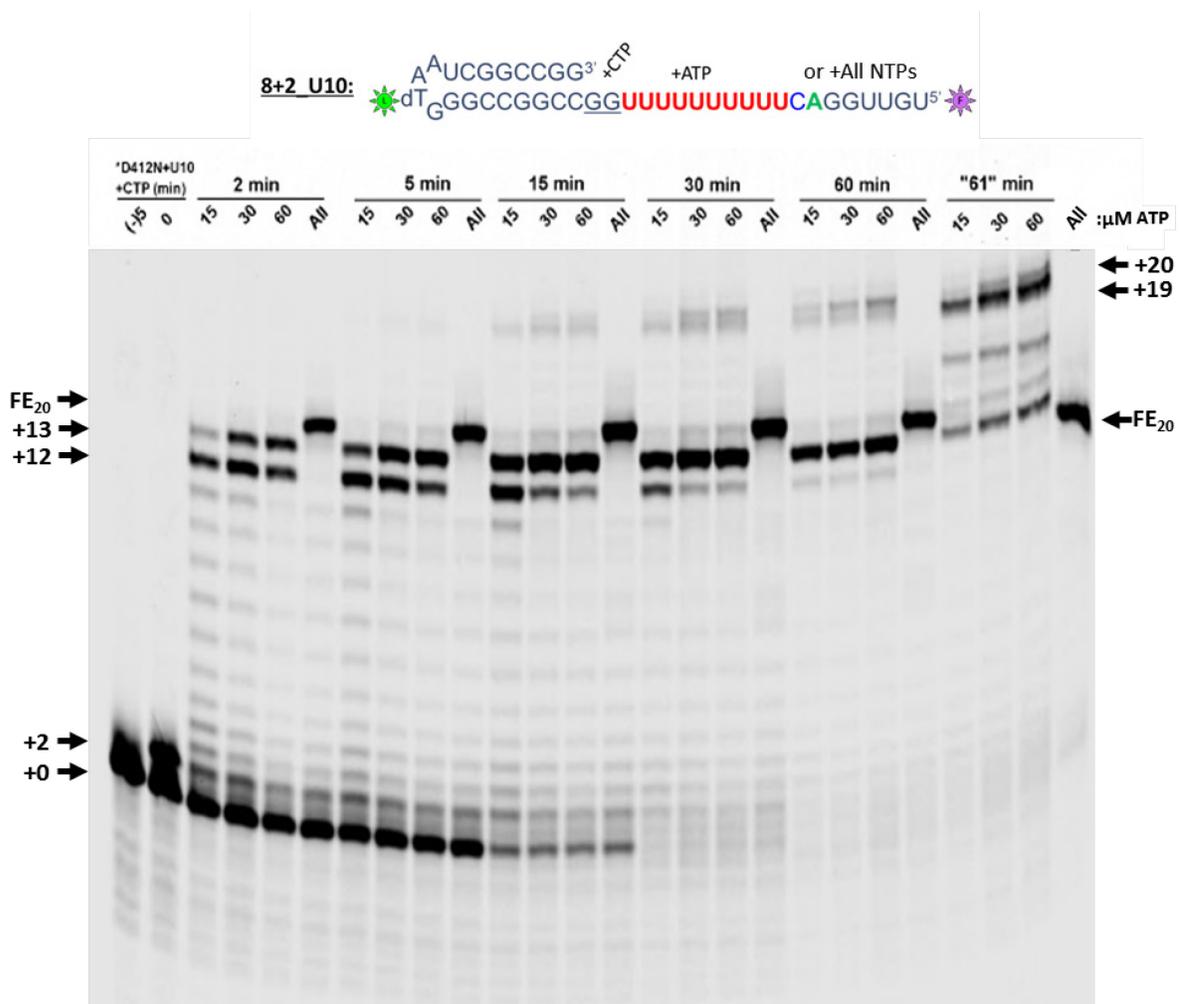


Fig. 4.9. Poly(A) slippage of D412N 3D^{pol} dependence on ATP concentration.

Time course series of the D412N 3D^{pol} mutant poly(A) slippage in the presence of the 8+2_U10 RNA template and increasing amounts of ATP or all NTPs. The elongation complex was formed by mixing RNA, protein and CTP at room temperature for 5 minutes, from which a sample was taken immediately before putting at room temperature (-5) and after the 5 minute incubation (0). The reactions were then mixed with various concentrations of ATP or all NTPs so that the reactions contained a final concentration of 15 μ M 3D^{pol}, 1 μ M U10 RNA, and either 15, 30, or 60 μ M ATP. A reaction was also run using the same protein and RNA concentrations in the presence of all NTPs at 31.25 μ M each (All). Samples were taken from each reaction and quenched at 2 min, 5 min, 15 min, 30 min, and 60 min. An additional sample was taken at 60 min from the reactions containing ATP only and were chased for one minute with all NTPs at a final concentration of 30 μ M ("61" min).

monitored over the course of 60 minutes. As a control, the pre-locked mixture was also used to carry out a reaction in the presence of 31 μM (each) of all four NTPs. At the end of the 60 minute time course, an extra 4 μL sample was taken from the reactions containing ATP, added to 1 μL of buffer containing 150 μM (each) of all the NTPs (resulting in a final concentration of 30 μM), and these chase reactions were allowed to proceed for 1 minute before being quenched ("61" min samples). As can be seen from the resulting gel in Fig. 4.9, there is a clear dependence of the poly(A) transcript length on both the concentration of ATP in the reaction and the length of time the reaction is allowed to proceed. Two minutes after adding the NTPs, there are nearly equivalent amounts of +12 and +13 product in the reactions containing 60 μM ATP, whereas there is very little +13 product compared to +12 product in the reactions containing 15 μM ATP and slightly less +13 than +12 product in the 30 μM ATP reactions. At 5 minutes, there is more +13 than +12 product in the 60 μM ATP reactions, $\approx 1:1$ ratio of +12 to +13 product in the 30 μM ATP reactions, and there is still more +12 than +13 product in the 15 μM ATP reactions. At 15 minutes, the 15 μM ATP reactions contain nearly equal amounts of +12 and +13 products, the +19 product becomes evident, and both the +19 and, faintly, the +20 products become visible in the 30 and 60 μM reactions. There is also a significant reduction in the amount of unelongated (+0) product at this time, again suggesting that initiation is still occurring in the reactions. At 30 minutes, the +20 product begins to appear in the 15 μM reactions and is nearly equivalent to the amount of +20 product in the 30 and 60 μM ATP reactions. By 60 minutes, there is very little +12 product remaining in all the reactions and slightly more +20 than +19 products in the reactions containing 30 and 60 μM ATP and another band above the +20 product is

faintly visible in all the reactions. In the “61” min chase reactions, the +12 product is no longer visible and the intensity of the +13 product is significantly reduced while the intensity of the +19 product increases. There is also an appearance of a new band between the +13 and +19 product. It is not fully clear what RNA species these bands represent. From the reactions containing all the NTPs, we know that the fully elongated product (FE₂₀), in which 20 nucleotides have been incorporated, migrates just above the +13 poly(A)₁₁ product as opposed to running at the same location as the +20 poly(A)₁₈ product. Therefore, it would stand to reason that if the 8 terminal nucleotides were added to the +13 product, the resulting FE₂₁ product would also electrophorese more quickly on the gel and could be responsible for the new band that appears between the +13 and +19 bands or perhaps is the cause of the increased intensity at the +19 band. It should be noted that the FE₂₁ product tends to “frown” in polyacrylamide gels due to the temperature gradient across the gel, as can be seen by the far right lane in which all NTPs were present during the reaction and the FE₂₀ band in Fig. 4.10. Thus, the disappearance of the +12 products in the “61” min chase reaction is most likely due to elongation to the FE₂₁ product but the small amount of material present is masked by the overlapping and more intense +13 product.

Effects of salt concentration on poly(A) slippage

Since it appeared that initiation was still occurring during the previous reactions, we wanted to investigate whether 3D^{pol} was dissociating and then re-binding the RNA in the process of generating the poly(A) transcripts. In other words, we wanted to find out whether the polymerase was actually “slipping” on the poly(U) template to generate

longer poly(A) sequences, or was complete release and re-binding of the RNA taking place. To test this, we decided to compare the poly(A) transcript lengths generated from the 8+2_U10 RNA by wild type 3D^{pol}, the highly stable D412N mutant, and the unstable S419T mutant, in high versus low salt as high salt would prevent the polymerase from re-binding the RNA (84, 130). Stable elongation complexes were formed at room temperature for 5 minutes in the presence of CTP under low salt conditions then added to either low or high salt buffer containing ATP, with the final reaction containing 45 μ M ATP in either 87.5 mM NaCl or 300 mM NaCl. We decided to monitor these reactions over the course of 2 hours to see if longer poly(A) transcripts would be generated than the +20 poly(A)₁₈ products that we had observed during the previous 1 hour reactions. Similar to the previous experiment, we took an additional sample from each reaction at the end of the time course and did a one minute chase (“121” min), this time in the presence of 60 μ M GTP so as to only allow incorporation of one additional nucleotide after the poly(U) sequence in the template and hopefully allow us to better identify the resulting products.

The resulting gel is shown in Fig. 4.10 and the RNA bands from the 2, 20 and 120 minute time point samples were quantitated (as described for Table 4.2) and are given in Table 4.3. Visual inspection of the gel demonstrates that the high salt buffer condition is inhibiting re-initiation from occurring, as can be seen by the significant loss of the +0 and +2 RNA in low salt between 2 to 120 minutes, whereas the relative amounts of the +0 to +2 material in high salt changes very little. There are also significantly less +19 and +20 products and more +12 material in the high salt reactions

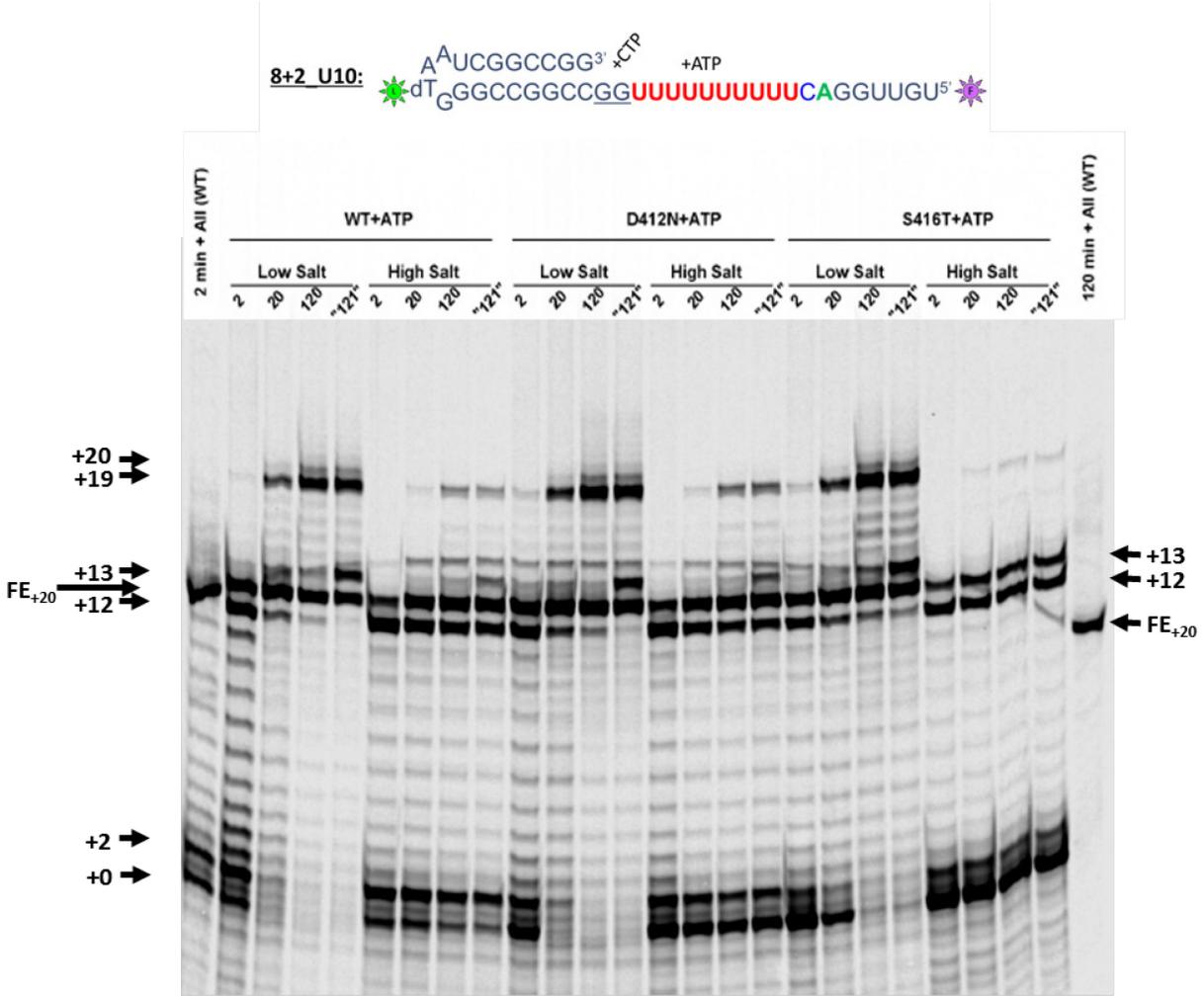


Fig. 4.10. WT , D412N, and S416T poly(A) slippage in high vs. low salt.

Poly(A) slippage assay using 8+2_U10 RNA with of wild type, the D412N mutant, or S416T mutant 3D^{pol} in the presence of 45 μM ATP and either low salt (87.5 mM NaCl) or high salt (300 mM NaCl) over the course of 2 hours. At the end of the 2 hour period, an extra sample was taken from each reaction and chased with 45 μM GTP ("121" min). 2 minute and 120 minute samples from a control reaction containing wild type polymerase, 8+2_U10 RNA, and 45 μM each of all NTPs in low salt buffer are also included on the outside lanes of the gel.

Table 4.3. Quantitation results of poly(A) smearing in high versus low salt.¹¹

U10 RNA + ATP in Low Salt										
Mutant	RNA Band	2 min			20 min			120 min		
		Intensity	% of Lane	% of +12 to +20 Products ¹²	Intensity	% of Lane	% of +12 to +20 Products	Intensity	% of Lane	% of +12 to +20 Products
WT	+0	840000	5.12%		301000	2.45%		195000	2.02%	
	+2	1640000	10.00%		376000	3.06%		240000	2.49%	
	+12	2780000	16.95%	52.04%	529000	4.30%	9.68%	317000	3.29%	8.17%
	+13	2330000	14.21%	43.62%	4160000	33.82%	76.15%	2080000	21.58%	53.59%
	+19	232000	1.41%	4.34%	539000	4.38%	9.87%	1110000	11.51%	28.60%
	+20				235000	1.91%	4.30%	374000	3.88%	9.64%
	Lane	16400000			12300000			9640000		
D412N	+0	1900000	11.95%		390000	2.79%		309000	2.47%	
	+2	1180000	7.42%		365000	2.61%		235000	1.88%	
	+12	2230000	14.03%	39.84%	629000	4.49%	9.43%	412000	3.30%	6.67%
	+13	3110000	19.56%	55.57%	5050000	36.07%	75.70%	3870000	30.96%	62.63%
	+19	257000	1.62%	4.59%	754000	5.39%	11.30%	1530000	12.24%	24.76%
	+20				238000	1.70%	3.57%	367000	2.94%	5.94%
	Lane	15900000			14000000			12500000		
S416T	+0	4560000	30.00%		1460000	10.81%		330000	2.41%	
	+2	768000	5.05%		417000	3.09%		246000	1.80%	
	+12	1100000	7.24%	33.10%	649000	4.81%	12.07%	451000	3.29%	6.56%
	+13	1960000	12.89%	58.98%	3760000	27.85%	69.95%	4130000	30.15%	60.11%
	+19	263000	1.73%	7.91%	737000	5.46%	13.71%	1890000	13.80%	27.51%
	+20				229000	1.70%	4.26%	400000	2.92%	5.82%
	Lane	15200000			13500000			13700000		

U10 RNA + ATP in High Salt										
Mutant	RNA Band	2 min			20 min			120 min		
		Intensity	% of Lane	% of +12 to +20 Products	Intensity	% of Lane	% of +12 to +20 Products	Intensity	% of Lane	% of +12 to +20 Products
WT	+0	877000	6.36%		769000	5.83%		498000	4.08%	
	+2	1800000	13.04%		1630000	12.35%		1070000	8.77%	
	+12	3610000	26.16%	77.73%	2100000	15.91%	43.80%	1400000	11.48%	29.96%
	+13	808000	5.86%	17.40%	2280000	17.27%	47.55%	2760000	22.62%	59.06%
	+19	226000	1.64%	4.87%	230000	1.74%	4.80%	326000	2.67%	6.98%
	+20				185000	1.40%	3.86%	187000	1.53%	4.00%
	Lane	13800000			13200000			12200000		
D412N	+0	2720000	17.89%		1860000	15.12%		1460000	12.27%	
	+2	1330000	8.75%		966000	7.85%		685000	5.76%	
	+12	3340000	21.97%	72.91%	1300000	10.57%	34.19%	829000	6.97%	20.46%
	+13	1020000	6.71%	22.27%	2040000	16.59%	53.66%	2580000	21.68%	63.67%
	+19	221000	1.45%	4.82%	243000	1.98%	6.39%	420000	3.53%	10.37%
	+20			0.00%	219000	1.78%	5.76%	223000	1.87%	5.50%
	Lane	15200000			12300000			11900000		
S416T	+0	5120000	34.36%		4870000	32.68%		3560000	26.97%	
	+2	920000	6.17%		815000	5.47%		701000	5.31%	
	+12	1670000	11.21%	63.23%	1480000	9.93%	50.19%	1440000	10.91%	51.14%
	+13	743000	4.99%	28.13%	1020000	6.85%	34.59%	903000	6.84%	32.07%
	+19	228000	1.53%	8.63%	256000	1.72%	8.68%	287000	2.17%	10.19%
	+20				193000	1.30%	6.54%	186000	1.41%	6.61%
	Lane	14900000			14900000			13200000		

Footnotes shown on next page

¹¹ Quantitation results of the 2, 20, and 120 minute timepoint samples in both low and high salt from Fig. 4.10. Quantitation was done the same as described for Table 4.4 wherein the % of lane is the intensity of the indicated RNA band divided by the intensity of the lane, and the lane intensity was determined for the region ranging from the +0 to the +20 bands.

¹² % of +12 to +20 products was calculated by dividing the intensity of the indicated band by the sum of the +12, +13, +19, and +20 intensities

over the course of the experiment as compared to the low salt reactions. Interestingly, faint bands above the +20 band do become apparent in the low salt reactions after 2 hours.

These bands are more distinct in the D412N 120 and “121” minute samples and there appear to be \approx 6-7 nucleotides longer than the +20 band in these lanes. In the low salt “121” min reactions in which samples of the 120 minute reactions were chased with GTP for one minute, the amount of +12 product is visibly reduced, there is little change of the +13 band, and the +14 band becomes much more prominent, suggesting the majority of the +12 poly(A)₁₀ RNA is able to incorporate the next nucleotide and only a portion of the +13 poly(A)₁₁ product is able to do so. In the reactions carried out in high salt, this chase from +12 to +13 and +13 to +14 products is significantly attenuated in the presence of wild type and the D412N mutant 3D^{pol}, and the chase does not appear to occur at all in the presence of the S416T mutant. The differing migration and frowning effect of the FE₊₂₀ product can be seen in the 2 min and 120 min reactions containing wild type 3D^{pol} and all NTPs, and a similarly migrating species can be seen throughout the samples. We attribute this to alternate denaturation states of the RNA products, similar to that observed in Fig. 4.6 in which the FE₊₂₀ products from the U10 and A10 RNA migrate much differently than the C10 FE₊₂₀ product that is the same length but does not denature as readily in the gel.

Quantitation of the +0 to +20 bands from Fig. 4.14, shown in Table 4.3, reveals that the reactions in the low salt buffer conditions follow a similar pattern as seen for the U10 RNA reactions in Fig. 4.7B wherein there is a significant accumulation of the +13 material, representing \approx 40% or more of the elongated (+12 to +20) products after only

two minutes. From 20 minutes to 2 hours in low salt conditions, the relative amount of +13 product diminishes in the reactions containing wild type and the D412N mutant 3D^{pol} and stays relatively the same in the reaction containing the S416T mutant while the relative amount of +19 product has increased \approx 2-fold, and there is only a small amount of +12 product remaining. Alternatively, in the high salt reactions the +13 product accumulates more slowly, with the +12 product representing \approx 60% or more of the elongated products at two minutes into the reaction. It isn't until 20 minutes after the reactions were started that the +13 product in the reactions containing wild type and the D412N mutant 3D^{pol} in high salt buffer reaches similar levels to that seen at 2 minutes in the low salt reactions, and the amounts of +13 product do not surpass that of the +12 product in the reaction containing the S416T mutant in high salt even after 2 hours. In addition, there is much more +12 product remaining and much less +19 and +20 products produced in the high salt reactions after 2 hours compared to the low salt reactions. The D412N 3D^{pol} mutant, which is \approx 2-fold more stable than wild type, generates the most +19 poly(A)₁₇ and +20 poly(A)₁₈ products and the S416T mutant, that is \approx 3-fold less stable than wild type, generates the least of these products under high salt conditions over the course of 2 hours. In contrast, all three polymerases produce similar relative amounts of the elongated products under low salt conditions during this time. Taken together, these data suggest that that both dissociation and slippage may take place during reiterative poly(A) transcription, appearing that dissociation is not required but rather aids the process.

4.3.4 Mutations in the Thumb have Variable Effects on Initiation and Stability

Elongation Complex Initiation

To determine how these mutations affected the ability of 3D^{pol} to initiate transcription and form a stable elongation complex, we used the 9+2-24 self-priming PETE RNA (Fig. 4.2) and only included the first two nucleotides, ATP and GTP, in the reaction. For each mutant, samples were taken from the reaction over the course of 30 minutes, resolved on a denaturing polyacrylamide gel (as shown for WT and S416V in Fig. 4.3), and the bands were quantified to measure the amount of +2 product formation. These values were then plotted as a function of time and analyzed by fitting to a single exponential curve.

The resulting initiation rates for the wild type and mutant 3D^{pol} constructs are provided in Fig. 4.11, with the majority of the constructs being within ≈ 2 -fold range of wild type. Of the helix mutants, mutations at residue 416 consistently had the most significant effect on elongation complex formation, ranging from ≈ 2 - to 6-fold slower initiation times compared to wild type for S416A and S416V, respectively. This feature appears to carry over in the double G64S/S416A mutant, resulting in an initiation time that is ≈ 3 -fold slower than wild type as opposed to the ≈ 2 -fold difference demonstrated by the individual mutants. Of all the mutants, K133A had the most profound influence on initiation with a time nearly 10-fold lower than wild type, which is consistent with it affecting RNA binding (130).

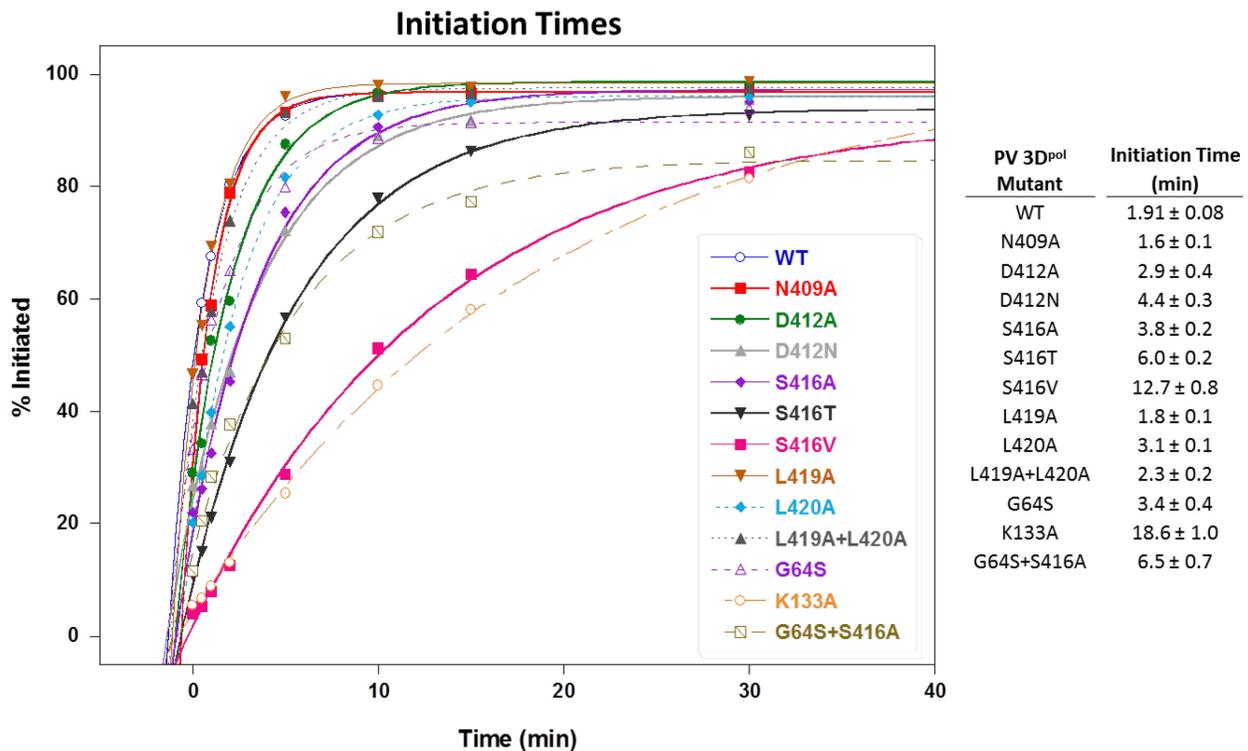


Fig. 4.11. 3D^{pol} initiation times.

The percent of initiated (+2) product formed in the presence of the 9+2-24 PETE RNA, wild type or mutant 3D^{pol}, and the first two nucleotides (ATP and GTP) was monitored over the course of 30 minutes. These values were plotted and curve fit to a single exponential in order to obtain the initiation time for each mutant, which are listed in the table to the right of the graph along.

Elongation Complex Stability

After the 30 minute initiation timecourse was complete, the stability of each resulting elongation complex was assessed. This was done by diluting the initiated (+2) 3D^{pol} and 9+2-24 PETE RNA complexes 10-fold in buffer containing 300 mM NaCl, thereby preventing any protein that may dissociate from rebinding and reinitiating. A control PETE RNA was also added in the dilution buffer to verify that only the previously intact +2 elongation complexes were being elongated. The diluted complexes were monitored over the course of 5 hours during which samples were taken, mixed with ATP, UTP, and GTP for 3 minutes, and analyzed to determine the amount of +2 material that was fully elongated to form the +20 product. These values were then plotted and fit to an exponential curve to obtain a dissociation rate for each mutant, as shown in Fig. 4.12.

These data show a range of effects on 3D^{pol} stability, with L419A and G64S being the only two mutants that are similar to the \approx 110 min wild type dissociation rate at \approx 90 and \approx 130 minutes, respectively, and the K133A and D412N mutations both demonstrating a nearly 2-fold increase in stability compared to wild type. It should be noted, however, that our laboratory has previously found the G64S and K133A mutants to be less stable than wild type (90, 130). It is unclear why the dissociation rates reported in this work for the G64S and K133A mutants are higher than those previously observed, but it would be worthwhile to retest these two mutants directly alongside wild type to determine whether these differences may be related to variations between reaction conditions. Of the mutations in the thumb α -helix, the majority show stability defects compared to wild type, with G64S/S416A, D412A, N409A, L420A,

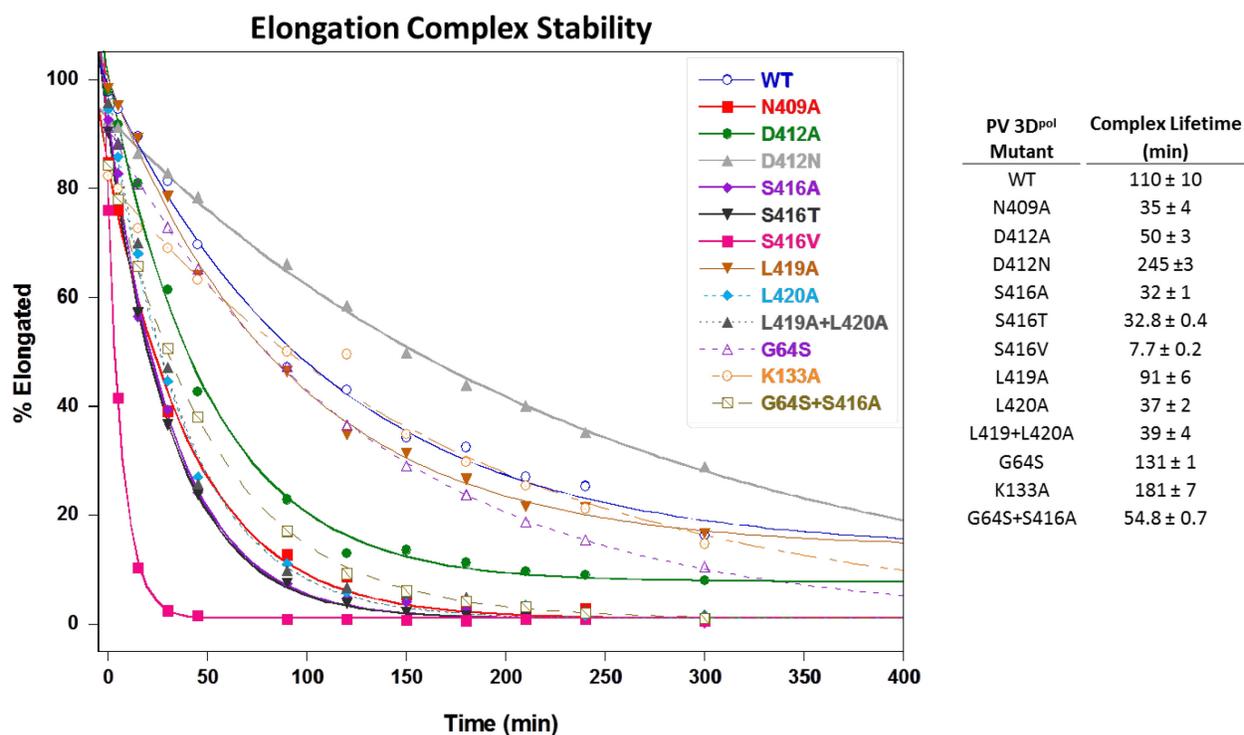


Fig. 4.12. 3D^{pol} elongation complex stability.

To determine the stability of the elongation complex formed by each mutant, initiation reactions were diluted into high salt buffer. Samples were periodically taken and mixed with NTPs in high salt buffer for 3 minutes to allow only elongation complexes that remained intact to incorporate the next 20 nucleotides on the 9+2.24 PETE RNA template. The samples were analyzed to determine the percent of material that was elongated and the resulting data was plotted and curve fit in order to obtain dissociation rates for each mutant, which are listed in the table to the right.

L419A/L420A, S416A, and S416T displaying ≈ 2 - to 3-fold reduction in stability. S416V is the least stable of all the mutants, having ≈ 14 -fold lower stability than wild type, suggesting that size as well as polarity is important at this position since neither the Ala nor the Thr mutation at this residue had such a dramatic effect. Taken together, mutations in the α -helix of the thumb that are associated with generating altered poly(A) tail lengths tend to demonstrate more significant effects on the stability of the elongation complex and relatively minor effects on the ability of the complex to initiate.

4.3.5 Thumb Mutants have Minor Effects on Elongation and Fidelity

Elongation Rates

To evaluate whether the thumb mutants affect the speed at which 3D^{pol} elongates RNA, we determined the Michaelis-Menten parameters for each mutant using the fluorescently labeled 10+4_22 RNA (Fig. 4.2B) via a stopped-flow fluorescence assay. The 10+4_22 PETE RNA has a self-priming hairpin, a 4 nt lock sequence, followed by an additional 24 nt single stranded sequence. Stalled +4 elongation complexes were pre-formed in the presence of ATP and GTP and then used to analyze the speed of elongation in the presence of 2-2500 μ M NTPs. In these assays, there is an initial lag phase in which the fluorescence intensity drops corresponding to elongation through the initial 18 nucleotides, followed by an increase in intensity as the polymerase reaches the 5th nucleotide from the 5' end (129). Thus, the nt/sec elongation rate can be determined by dividing the 18 nt by the lag phase and the resulting values for each NTP concentration can then be plotted to determine the K_m and V_{max} (Fig. 4.13 and Table 4.4). As seen in Fig. 4.13C, we typically observe a

decrease in elongation speeds at NTP concentrations greater than $\approx 750 \mu\text{M}$, for which a similar observation has been previously reported (143). Although we are unclear as to why this inhibition occurs, we were able to determine the Michaelis-Menten kinetics as well as the apparent $k_{\text{inhibition}}$ for each construct using a self-inhibition curve fit and the resulting values are listed in Table 4.4.

Overall, the resulting K_m and V_{max} values for each of the mutants are similar to wild type. The exceptions to this is G64S that is ≈ 2 -fold slower than wild type, a value similar to the previously reported 3-fold reduction in speed observed by the Cameron laboratory (90). The only other mutant that appears to have a faster elongation rate is S416V, but it is important to note that the extremely poor stability of this protein made it very difficult to obtain data above background levels, and thus the low quality of data for this mutant is reflected in the high error.

Fidelity

To examine the fidelity of each mutant, we tested the ability of each construct to misincorporate a 2' deoxyribose CTP across from a guanine nucleotide, a technique that has been previously utilized by our laboratory to test the fidelity of coxsackievirus polymerase (128). This fidelity assay is based off the crystal structure of the poliovirus elongation complex, in which active site closure and catalysis is, in part, is triggered by recognition of the 2' and 3'-ribose hydroxyl groups (82, 128). As the 10+4_22 RNA was designed with a unique guanine at the 8th nucleotide from the 5' end (Fig. 4.2B), we used the same pre-formed 10+4_22 elongation complex for these experiments. The incorporation rate of varying amounts of CTP versus 2'-dCTP was measured while

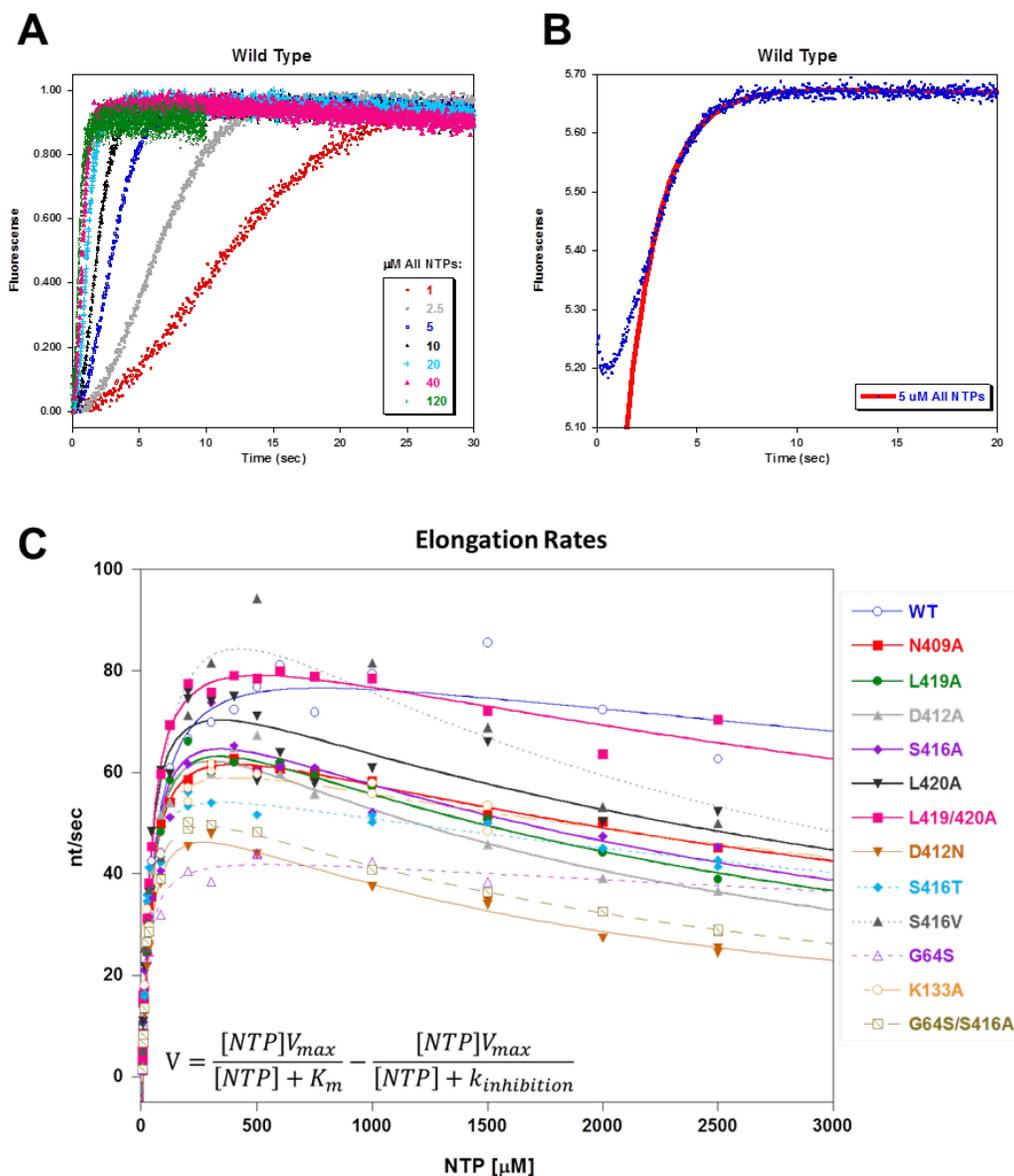


Fig. 4.13. Determination of 3D^{pol} elongation rates.

A) Normalized fluorescence signals obtained for wild type 3D^{pol} at several different NTP concentrations demonstrating the drop in the fluorescence signal that corresponds to elongation of first 18 nucleotides in the 10+4₂₂ that is followed by a signal increase as the polymerase reaches the 5th nucleotide from the 5' end and the length of the lag phase decreases in the presence of higher concentrations. B) Raw data points obtained from wild type 3D^{pol} in the presence of 5 μM NTPs showing how the data was fit in order to determine the length of the lag phase. C) This value was then divided by 18 to obtain a nucleotide per second rate that were then plotted as a function of NTP concentration to generate a Michaelis-Menten plot. Shown in the bottom left corner of the graph is the equation used to determine the elongation rate (V_{max}), K_m , and the $k_{inhibition}$ term (to account for the observed decrease in rate at NTP concentrations greater than $\approx 750 \mu\text{M}$), and the values obtained for each mutant are listed in Table 4.4.

Table 4.4. 3D^{pol} elongation rates.

PV 3D^{pol} Mutant	K_m (μM)	V_{max} (nt/sec)	k_{Inhibition} (mM)
WT	52 \pm 8	88 \pm 5	12 \pm 5
N409A	38 \pm 2	75 \pm 1	4.2 \pm 0.3
D412A	42 \pm 4	83 \pm 4	2.1 \pm 0.3
D412N	38 \pm 2	62 \pm 2	1.9 \pm 0.1
S416A	36 \pm 5	86 \pm 5	2.4 \pm 0.5
S416T	20 \pm 3	61 \pm 3	6.1 \pm 1.6
S416V	86 \pm 32	128 \pm 26	2.0 \pm 0.8
L419A	45 \pm 4	82 \pm 3	2.6 \pm 0.3
L420A	33 \pm 6	86 \pm 5	3.4 \pm 0.8
L419+L420A	38 \pm 3	92 \pm 3	7 \pm 1
G64S	24 \pm 3	46 \pm 2	13 \pm 8
K133A	30 \pm 4	69 \pm 3	5.3 \pm 1.5
G64S+S416A	34 \pm 3	63 \pm 2	2.2 \pm 0.2

keeping the other NTPs at a constant concentration. Similar to that described in the elongation rate assays, the data was curve fit to obtain a nt/sec rate at each CTP and 2'-dCTP concentration, which was then plotted as a function of CTP or 2'-dCTP concentration to determine the Michaelis-Menten kinetics for each (Fig. 4.14 and Table 4.5). By dividing the V_{max} by the K_m we can obtain the incorporation efficiency for CTP and 2'-dCTP and from this we can determine a discrimination factor by dividing the CTP efficiency by the dCTP efficiency, thus the discrimination factor $= \frac{CTP\ Efficiency}{2'-dCTP\ Efficiency} = \frac{(V_{max}/K_m)_{CTP}}{(V_{max}/K_m)_{2'-dCTP}}$, and these values are also listed in Table 4.5. A higher discrimination factor reflects an improved ability to preferentially incorporate the correct CTP nucleotide over the 2'-dCTP nucleotide, thus indicating a polymerase with higher fidelity (128).

Using this assay, we investigated the fidelity of wild type polymerase, the five original thumb mutants (N409A, D412A, S416A, L419A, L420A, and L419/420A), G64S, K133A, and D412N and the results are shown in Fig. 4.14. As expected, the G64S mutant had a ≈ 3 -fold higher discrimination factor than wild type 3D^{pol}, which is consistent with results from previous studies (89, 90, 144). Unexpectedly, the S416A and K133A mutations, that had been identified *in vivo* as having a lower mutation frequency than G64S, had discrimination factors very similar to that of wild type. Indeed, the discrimination factor for S416A at 121 ± 9 was within error of the wild type factor of 117 ± 11 , while K133A had a slightly lower value of 85 ± 12 . The remaining mutants had very similar discrimination factors to wild type, ranging from ≈ 82 -88, with

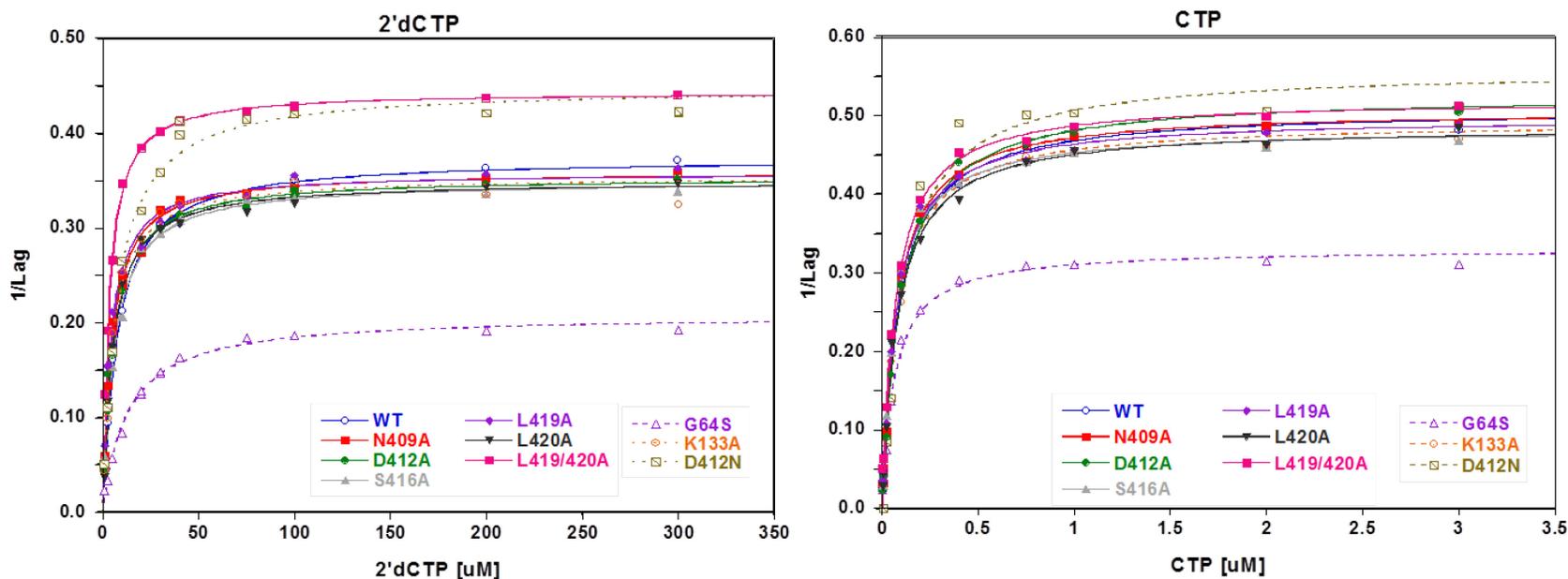


Fig. 4.14. Determining the fidelity of poliovirus 3D^{pol}.

The fidelity of wild type and mutant PV 3D^{pol} was obtained by determining the Michaelis-Menten type kinetics of the polymerase when either 2'dCTP (left panel) or CTP (right panel) was titrated into reactions containing elongation complexes that were pre-formed with the 10+4_{22v4} PETE RNA. The resulting K_m and V_{max} values were used to calculate the discrimination factor for each polymerase, which are listed in Table 4.5. Please note that there is a data point for D412N at 3 μM CTP that lies approximately at the same value as N409A and D412A.

Table 4.5. Fidelity of poliovirus 3Dpol.

PV 3D^{pol} Mutant	2'dCTP		CTP		Discrimination Factor
	K_m (μM)	V_{max} (sec^{-1})	K_m (μM)	V_{max} (sec^{-1})	
WT	7.3 \pm 0.3	0.374 \pm 0.003	0.085 \pm 0.007	0.508 \pm 0.009	117 \pm 11
N409A	4.5 \pm 0.3	0.360 \pm 0.004	0.074 \pm 0.005	0.507 \pm 0.006	86 \pm 7
D412A	5.2 \pm 0.4	0.353 \pm 0.005	0.095 \pm 0.005	0.527 \pm 0.007	82 \pm 8
D412N	7.4 \pm 0.7	0.448 \pm 0.008	0.10 \pm 0.02	0.56 \pm 0.02	88 \pm 16
S416A	5.8 \pm 0.3	0.350 \pm 0.004	0.066 \pm 0.004	0.483 \pm 0.005	121 \pm 9
L419A	4.0 \pm 0.3	0.358 \pm 0.004	0.073 \pm 0.005	0.498 \pm 0.007	77 \pm 8
L420A	5.1 \pm 0.1	0.349 \pm 0.004	0.080 \pm 0.005	0.487 \pm 0.007	88 \pm 9
L419/420A	3.1 \pm 0.1	0.444 \pm 0.003	0.068 \pm 0.003	0.521 \pm 0.004	53 \pm 3
G64S	12.7 \pm 0.8	0.208 \pm 0.003	0.066 \pm 0.005	0.330 \pm 0.005	307 \pm 32
K133A	4.8 \pm 0.6	0.354 \pm 0.008	0.079 \pm 0.007	0.419 \pm 0.010	85 \pm 12

the exception of the double L419/L420 mutant which was down about 2-fold. Overall, these results from this fidelity assay do not indicate that any of the mutants tested have higher fidelity than wild type 3D^{pol} based on a 2' hydroxyl discrimination.

4.4 Discussion

The poliovirus poly(A) tail at the 3' end of the viral genome is important for efficient negative-strand RNA synthesis, viral infectivity, and protein translation (135-139). The poly(A) tail ranges from \approx 20-150 nucleotides in length, and a minimum length of at least poly(A)₂₀ is necessary for efficient negative strand RNA synthesis (132, 133, 139). To determine which regions of the poliovirus polymerase are important for mediating poly(A) tail length, the Barton laboratory examined the effects of mutations in 3D^{pol} on the polyadenylation of virion RNA produced by transfected HeLa cells and found that mutations at several residues the α -helix in the thumb domain altered the length of the poly(A) tail (142). These findings were supported by our initial *in vitro* poly(A) slippage assays in which we found that these alanine mutations in the thumb generated poly(A) transcripts with altered lengths that demonstrated very similar phenotypes to that observed by the Barton laboratory (Fig. 4.4 and Fig. 4.5). Thus, our goal in the studies presented here was to further probe the mechanism by which mutations in the thumb α -helix were causing altered poly(A) phenotypes using purified polymerase *in vitro*.

4.4.1 Multiple Determinants of Altered Poly(A) Tail Lengths

In order to assess whether these mutations were altering a particular aspect of the polymerase function, we investigated how these mutations affected the ability of 3D^{pol} to initiate and form an active elongation complex, the stability of the elongation complex, and polymerase elongation rates. A total of twelve 3D^{pol} mutants were characterized, including the original alanine helix mutants identified by the Barton laboratory (N409A, D412A, S416A, L419A, and L420A), alongside wild type 3D^{pol}. In addition, we assayed the fidelity of several of these constructs by measuring their propensity to incorporate 2'-dCTP in comparison to CTP across from a guanine residue to obtain a discrimination factor. Table 4.6 summarizes these *in vitro* results as well as the *in vivo* poly(A) tail length and fidelity results determined by the Barton laboratory.

The Barton laboratory identified the K133A and S416A mutants as possible high fidelity mutants since they both resulted in fewer mutations than the known high fidelity G64S mutant as determined by sequencing the 3D^{pol} region of thirty TOPO-TA clones from virion RNA for each mutant. The discrimination factors obtained using our fidelity assay do reflect the high fidelity of the G64S mutation, but do not indicate that the K133A and S416A mutations have higher fidelity. The discrimination factor for K133A is lower than that of wild type and similar to that of the other thumb helix mutants. The discrimination factor for the S416A mutant is slightly higher, but within error of the discrimination factor for wild type. Taken together, either the sequencing method used by the Barton laboratory or our fidelity assay is not accurately reflecting the fidelity of these mutants. Our fidelity assay was designed to evaluate how well the polymerase was able to discriminate between the correct 2'-hydroxyl of the incoming CTP and an

Table 4.6. Summary of *in vitro* and *in vivo* characterization results.

PV 3D ^{pol} Mutant	<i>In Vitro</i> Results ¹³					<i>In Vivo</i> Results ¹⁴		
	Initiation/Stability		Elongation ¹⁵		Fidelity ¹⁶	Length of poly(A) tails		Replication Fidelity
	Initiation Time (min)	Complex Lifetime (min)	K _m (μ M)	V _{max} (nt/sec)	Discrimination Factor	Phenotype	Range ¹⁷	No. of mutations
WT	1.91 ± 0.08	110 ± 10	52 ± 8	88 ± 5	117 ± 11	WT	25–93	8
G64S	3.4 ± 0.4	131 ± 1	24 ± 3	46 ± 2	307 ± 32	WT	<i>30–101</i>	2
K133A	18.6 ± 1.0	181 ± 7	30 ± 4	69 ± 3	85 ± 12	Longer	29–96	1
N409A	1.6 ± 0.1	35 ± 4	38 ± 2	75 ± 1	86 ± 7	Longer	25–110	3
D412A	2.9 ± 0.4	50 ± 3	42 ± 4	83 ± 4	82 ± 8	Shorter	21–80	4
D412N	4.4 ± 0.3	245 ± 3	38 ± 2	62 ± 2	88 ± 16			
S416A	3.8 ± 0.2	32 ± 1	36 ± 5	86 ± 5	121 ± 9	Broader	<i>28–97</i>	0
S416T	6.0 ± 0.2	32.8 ± 0.4	20 ± 3	61 ± 3				
S416V	12.7 ± 0.8	7.7 ± 0.2	86 ± 32	128 ± 26				
L419A	1.8 ± 0.1	91 ± 6	45 ± 4	82 ± 3	77 ± 8	Longer	39–113	3
L420A	3.1 ± 0.1	37 ± 2	33 ± 6	86 ± 5	88 ± 9	Longer	28–121	9
L419+L420A	2.3 ± 0.2	39 ± 4	38 ± 3	92 ± 3	53 ± 3			
G64S+S416A	6.5 ± 0.7	54.8 ± 0.7	34 ± 3	63 ± 2				

¹³ These are results from experiments done by me.

¹⁴ These are results from Table 4.1 that were obtained by the Barton laboratory (142).

¹⁵ Michaelis-Menten parameters determined in the presence of increasing amounts of all NTPs as reported in Table 4.2.

¹⁶ This value represents incorporation efficiency CTP versus 2'deoxy-CTP as reported in Table 4.3.

¹⁷ Values in red italics indicate poly(A) tail length ranges that are not statistically significantly different than wild type.

incorrect 2'-deoxy-CTP (128), but perhaps these mutations are affecting another aspect of elongation, such as the translocation step, that could alter the apparent fidelity of the polymerase. This may make some sense since these mutations are far from the catalytic center and do not appear to affect active site closure.

Characterization of the thumb α -helix mutants indicates that mutations in this helix tend to have the most significant effect on elongation complex stability, although there does not appear to be a singular characteristic that correlates with alteration of poly(A) tail length. Indeed, the L419A mutant that is most similar to wild type 3D^{pol} in its initiation rate and stability generates longer poly(A) tails *in vivo*. Alternatively, the S416A mutation that resulted in some of the most drastic defects in initiation and stability compared to wild type *in vitro* also generated poly(A) tails with statistically similar lengths to wild type *in vivo*. Similarly, there is no clear link between the Michaelis-Menten kinetics of the mutants and altered poly(A) tail lengths. Interestingly, the mutants with a discrimination factors that are similar to (S416A) or higher than (G64S) that of the wild type polymerase also generate poly(A) tails with statistically similar lengths to wild type, whereas mutants with lower discrimination factors generate altered poly(A) tail lengths. However, among the mutants with altered poly(A) tail lengths, the discrimination factor does not reflect a longer or shorter poly(A) phenotype. In fact, altered poly(A) tail lengths do not appear to be associated with just a defect in initiation, stability, elongation rate, or discrimination factor, but perhaps involve a combination of these traits.

4.4.2 A Model for Poly(N) Slippage by Poliovirus 3D^{pol}

Our initial poly(A) slippage assay was carried out using the T4 PETE RNA template and the thumb helix alanine mutants (Fig. 4.4, Fig. 4.5). Reiterative transcription of this RNA resulted in products of heterogeneous length, ranging from poly(A)₁₀ to greater than poly(A)₃₉, that appeared as smears on the gel, but still reflected the phenotypes observed by the Barton laboratory for these mutants. We next utilized the 8+2_A10, 8+2_U10, and 8+2_C10 PETE RNAs, as shown in Fig. 4.2. These were designed to have a longer, more stable priming hairpin that we hoped would prevent the heterogeneous smearing effect observed for the T4 RNA, and a two nucleotide lock sequence before the poly(N)₁₀ homopolymer region that would allow us to pre-form stable elongation complexes. Initial testing of these RNAs with wild type or mutant polymerase in the presence of the cognate poly(N)₁₀ NTP confirmed that 3D^{pol} is able to generate poly(A) and poly(U) transcripts that are longer than the poly(U/A)₁₀ sequence in the U10 and A10 template (Fig. 4.6, Fig. 4.7A&B). This was not the case with the C10 RNA as all of the constructs tested only elongated this RNA by 10 nucleotides in the presence of GTP (Fig. 4.7C).

As opposed to the smear of transcripts that had been generated from the T4 RNA (Fig. 4.4 and Fig. 4.5), slippage assays done in the presence of the A10 and U10 RNAs generated several distinct products that were longer than the 12 base template, including a +13 poly(N)₁₁ product, a +19 poly(N)₁₇ product, and a +20 poly(N)₁₈ product. Initially, we were surprised that all the poly(A) transcripts from the 8+2_A10 RNA appeared to stop at the +20 poly(A)₁₈ product, as we had previously observed transcripts from the T4 RNA that became longer than poly(A)₃₄ (Fig. 4.4). However,

upon closer inspection and extending the time of the reactions containing the U10 RNA, faint bands above the +20 product become apparent, as can best be seen in Fig. 4.7A and Fig. 4.10, indicating that longer poly(A) transcripts are being generated, just much more slowly than from the T4 RNA. Interestingly, we can discern $\approx 6-7$ bands above the +20 band, suggesting the presence of a +26/27 product. It is unclear why transcription from the A10 RNA occurs more slowly, but could be related to the greater stability of the 8 bp stem loop, as opposed to the $\approx 2-3$ bp hairpin in the T4 RNA, acting to reduce the flexibility of the complex.

From the single nucleotide resolution afforded by the A10 and U10 RNAs, we were able to monitor and compare the poly(U) and poly(A) transcripts produced under various conditions. In reactions containing the 8+2_U10 RNA in low salt buffer, we observed an ATP concentration dependent accumulation of a +13 poly(A)₁₁ product that occurred fairly rapidly (within 2 minutes), followed by a slower (within 15-20 minutes) formation of a prominent +19 poly(A)₁₇ product. Accompanying the +19 product was a faint +20 poly(A)₁₈ product that appeared to accumulate over the course of the 2 hour reactions, as can be seen in Fig. 4.9. Similarly, the +13, +19, and +20 products were also the major transcripts to appear in the reactions containing the 8+2_A10 RNA. However, the accumulation of the +13 product was not nearly as substantial in these reactions, comprising a much smaller fraction of the elongated RNA as compared to the +13 product in the 8+2_U10 RNA reactions, which typically represented the majority of all the RNA in the lane (Table 4.2). In addition, the +13 poly(A)₁₁ product experienced the largest relative increase between 15 and 30 minutes, whereas the +19 poly(U)₁₇ increased proportionally more than the other poly(U) products during this time,

suggesting there are slight mechanistic differences between poliovirus 3D^{pol} reiteratively transcribing a poly(U) versus poly(A) template.

As discussed in section 1.4.3 , the crystal structure of the 3D^{pol} elongation complex shows ≈10 nucleotides making key interactions with the polymerase, of which ≈6 of these nucleotides are encircled in the core of the polymerase. This includes the +1 and +2 positions of the single stranded template, that will serve as the next two sites of nucleotide incorporation and are located directly across and just downstream from the active site, respectively (82). Also enclosed by the polymerase are the four most recently incorporated nucleotides duplexed with the template strand at positions -1 to -4, with the -1 base being the newest and immediately upstream of the active site and the -4 position corresponding to the fourth base pair upstream of the active site. At approximately the -4/-5 position, the product-template duplex begins to contact the 3D^{pol} thumb α-helix, which packs into the minor groove of the A-form helix, and begins threading out of the polymerase between the thumb and the pinky domains until it loses contacts with the thumb at approximately the -8 position (82, 85).

Based on this observation and in conjunction with our data, I propose a model by which poliovirus 3D^{pol} may be reiteratively transcribing poly(A) sequences, as shown in Fig. 4.15. In the presence of the U10 RNA and poliovirus 3D^{pol}, the +2 elongation complex is formed by the addition of CTP. As demonstrated in the image of the +2 complex, 3D^{pol} encompasses six nucleotides, including the two G-C base pairs from the stem loop, the two newly formed C-G base pairs, and the next two single stranded uridine residues in the template. Upon the addition of ATP, 10 adenosines are incorporated to form the +12 poly(A)₁₀ product that stalls upon reaching the end of the

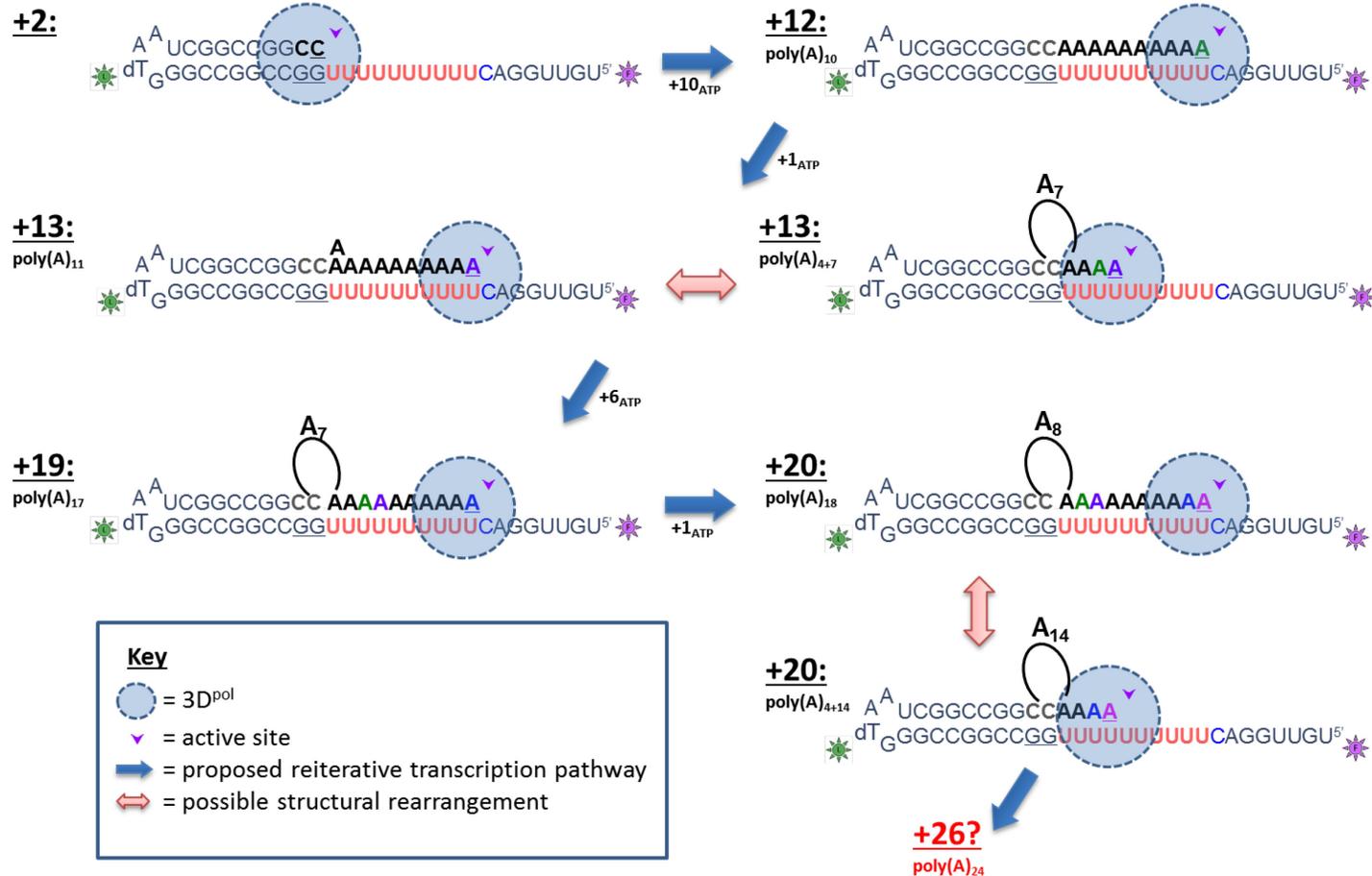


Fig. 4.15. Proposed poly(A) slippage model.

Proposed poly(A) slippage mechanism in which the elongation complex stalls at the end of the homopolymer stretch of the +12 poly(A)₁₀ product in the absence of additional nucleotides, promoting a single nucleotide slippage to allow for incorporation of an additional nucleotide. This destabilizes the complex and allows 3D^{pol} to slip back to the 3' end of the poly(U) template. Here it encompasses four A-U base pairs in its -1 to -4 positions and the active site is poised to incorporate an A across from the fifth poly(U) in the template. Please note that the location of the poly(A) bulges/loops are for illustrative purposes only and are not based on any of our experimental data.

poly(U)₁₀ portion of the template. This is followed by a single nucleotide slippage to allow the incorporation of a single adenosine, resulting in the +13 poly(A)₁₁ product. Based on our data showing the accumulation of this +13 product, we postulate that a structural rearrangement occurs within the +13 complex to form a poly(A)₄₊₇ complex, wherein the polymerase slips back to the 3' end of the poly(U)₁₀ template leaving four A-U base pairs downstream of the active site and a non-paired poly(A)₇ loop. Once this occurs, the polymerase is able to transcribe the remaining six residues of the poly(U)₁₀ template, resulting in the +19 poly(A)₁₇ product. After this another single nucleotide slippage occurs to generate the +20 poly(A)₁₈ product, similar to the +12 to +13 transition. At this point, we hypothesize that the +20 complex then undergoes a slow rearrangement, similar to the +13 complex, back to the 3' end of the poly(U)₁₀ stretch to form the +20 poly(A)₄₊₁₄ complex, containing four A-U base pairs downstream of the active site and a unpaired poly(A)₁₄ loop, before it can go on to incorporate six more adenosines and generate a +26 poly(A)₂₄ product. From our data, it appears that poly(U) transcription from a poly(A) template follows the same basic pathway, although our results from Fig. 4.7A and Table 4.2 suggest that the transition from the +12 to +13 product and +19 to +20 product may be the rate limiting step, rather than the +13 rearrangement step.

Transcript slippage is a common mechanism employed by eukaryotic, bacterial, and viral RNA polymerases that is important for a variety of functions, including regulation of gene expression (145), transcription termination (146), and maintaining genomic sequence integrity (147). Telomerase reverse transcriptase (TERT) is one such enzyme that maintains chromosomal telomeres by reiterative transcription of an

RNA template and shares the canonical 'right hand' structural features observed for polymerases (148). It has been proposed that picornaviral poly(A) tails may serve a similar purpose as telomeres, with 3D^{pol} carrying out polyadenylation as a means by which to maintain the tail length and prevent erosion of the genome (142).

Additional studies will be necessary in order to test the slippage model proposed in this report and to further investigate the mechanistic differences between poly(U) and poly(A) reiterative transcription and how these mutations affect this process, which I further address in Chapter 5. Supplementary studies are also being carried out by the Barton laboratory to ascertain the effects of the new thumb helix mutants (D412N, S416T, S416V, L419A+L420A) on poly(A) tail length *in vivo*. In conclusion, these studies provide important insight into the mechanism governing reiterative transcription of poly(A) and poly(U) sequences by poliovirus 3D^{pol} and lays the groundwork for further characterizing this poorly understood phenomenon.

CHAPTER 5

CONCLUSIONS AND FUTURE DIRECTIONS

5.1 Introduction

Picornaviruses are a family of viruses that are able to cause disease and can potentially be fatal in humans, animals, plants, and insects alike. Recent efforts in combating these viruses have been geared towards designing anti-viral compounds that target viral non-structural proteins that are important for the viral replication cycle (18-20). The goal of this work has been to investigate the structure-function relationship of two such proteins, the poliovirus 2C^{ATPase} and 3D^{pol}. The focus of Part I of this dissertation is the 2C protein, an oligomeric ATPase that is able to bind RNA and is involved in membrane rearrangement, both as a fully the processed protein and as the precursor 2BC protein (37, 38, 51, 61, 75, 76). Here, we demonstrate how proteolytic processing regulates 2C ATPase activity and also explore its putative helicase activity in the context of membrane bilayers. The subject of Part II is the poliovirus 3D^{pol} protein, the RNA-dependent RNA polymerase, in which we investigate the role of the thumb α -helix in reiterative transcription of poly(A) and poly(U) sequences and the mechanism by which this occurs. These studies provide important insight into the functions of these enzymes and lay the foundation for future exploration of these processes. In this chapter, I provide a brief summary of these results and propose future experiments for these projects.

5.2 2C^{ATPase}: Summary and Future Directions

In Chapter 2, we demonstrate a novel approach for expressing poliovirus 2C on lipid bilayers using an *in vitro* transcription/translation system in the presence of nanodiscs. To determine whether the nanodisc-bound 2C is functional, we assessed its ATPase activity and found that it possessed much higher activity than had been previously reported, suggesting that association with membranes is important for enzymatic activity. We confirmed that 2C is responsible for this activity by introducing mutations at the canonical AAA+ ATPase motifs, Walker A, Walker B and motif C, which significantly reduced or, in the case of the Walker A mutation, virtually abolished ATPase activity. Using the same nanodisc expression system, we also generated various combinations of 2C in the context of its membrane-associated polyprotein neighbors to generate 2BC, 2C3A, and 2BC3AB. Assessment of the ATPase activity of these various constructs revealed that 2C has the highest activity, followed by 2BC, then 2C3A and 2BC3AB, suggesting that 2C may always have innate ATPase activity that is allosterically modulated by the presence of polyprotein neighbors. In Chapter 3, we investigated the potential helicase activity of nanodisc bound 2C. Although we did not see clear evidence of RNA strand separation, we did observe 2C-dependent protection of RNAs that have a 3' and 5' single stranded overhang, in the RNase T1 helicase assays.

In my mind, one of the most intriguing questions regarding 2C that remains to be answered is the purpose of 2C ATPase activity. Although membrane association is important for enhancing 2C ATPase activity, as observed from our nanodisc-bound protein, it has been previously reported that mutations in the Walker A and Walker B

motifs of 2C do not impair membrane rearrangement (37). Based on the presence of motif C, which is typically found in the SF3 helicases, it has been proposed that 2C may act as a helicase during viral replication. However, our laboratory has observed that 3D^{pol} alone is able to separate dsRNA (82, 149). In addition, attempts by me and other laboratories to demonstrate 2C helicase activity have been unsuccessful (66, 70). Thus, there is little support for the theory that 2C is responsible for strand separation during RNA replication. Interestingly, mutations in 2C confer resistance to the antiviral compound hydantoin, which inhibits RNA encapsidation, (64). In addition, several recent studies have shown evidence for interactions between 2C and viral capsid proteins VP3 and VP1 via identification of suppressor mutations that rescue encapsidation defects (65, 66). Taken together, it may be that 2C is acting as a helicase/translocase during encapsidation wherein interactions between 2C and the capsid proteins induce a conformational change that allows 2C to facilitate the translocation of newly synthesized (+) strand RNA into the viral capsid. The most straightforward way to test this theory would be to assess how mutations at the Walker A, Walker B, and motif C residues in 2C affect packaging *in vivo*. This could be done using the “reporter virus” assay developed by the Wimmer laboratory (65, 66), perhaps in collaboration with Dave Barton’s laboratory as they have the expertise to carry out these experiments in HeLa cells. It may also be possible to test this model in our *in vitro* helicase assays by adding purified VP3 or VP1 to the reactions to see whether these proteins are able to stimulate 2C helicase activity. If there is a link between 2C ATPase activity and encapsidation, mutations of the 2C residues implicated in interactions with the VP3 and VP1 proteins, such as N252 and C323 (65, 66), could

also be tested to further explore how these interactions may be mediating 2C helicase/translocase activity.

Alternatively, it is possible that 2C does not act as a helicase at all but could potentially act as an RNA chaperone, as was recently demonstrated for the 2C protein from the *Ectropis oblique* picorna-like virus (EoV) (81). In this study, the authors demonstrated that EoV 2C was able to destabilize nucleic acid helices containing both a 3' and 5'-overhang, with a preference for 5' overhangs, and that this activity did not require ATP and was actually inhibited in the presence of greater than 2 mM ATP (81). As shown in Chapter 3, several of my experiments suggested that 2C was interacting with RNA-A and RNA-C, having a 3' and 5' overhang respectively, in both the presence and absence of ATP, and it would be interesting to further explore whether this could be indicative of PV 2C functioning as an RNA chaperone.

5.3 3D^{pol}: Summary and Future Directions

In Chapter 5, I investigated the role of several residues within 3D^{pol} in determining the length of reiteratively transcribed poly(A) and poly(U) sequences. These studies focused on residues located in an α -helix of the thumb domain that forms part of a protein clamp with the pinky domain around the upstream duplex RNA as it exits the polymerase (82). In collaboration with the Barton laboratory, we demonstrated that alanine point mutations at these residues (D412, S416, L419, and L420) altered the length of poly(A) transcripts compared to wild type, both *in vivo* and *in vitro*. I further analyzed mutations at these residues for the effects on initiation and stability of the elongation complex, elongation rates, and fidelity and found that mutations in the in the

thumb α -helix generally had a greater impact on elongation complex stability compared to the other traits that were tested. The fact that these mutations in the thumb α -helix affect the stability of the elongation complex is not too surprising, as these residues make contacts with the upstream product-template duplex as it exits the polymerase. It may be that these mutations have altered initiation and stability phenotypes on poly(A) and poly(U) sequences that allow them to more easily slip on these homopolymeric sequences and generate longer transcripts. Thus, it would be useful in the future to further characterize the initiation rates and elongation complex dissociation rates of these mutants using the A10, U10 and C10 RNAs. These experiments may also shed some light on potential mechanistic differences between reiterative transcription of poly(A) and poly(U) templates.

I also found that poliovirus 3D^{pol} is able to generate longer transcripts from homopolymeric poly(A) and poly(U) templates, but not poly(C) templates. Interestingly, there were several distinct poly(A) and poly(U) products that were generated from the A10 and U10 RNA as opposed to the wide range of transcript lengths that had been produced from the T4 RNA. These products corresponded to a poly(N) transcript that was 1 nucleotide longer than the template (+13 poly(N)₁₁), followed by a product that was 6 nucleotides longer than the +13 product (+19 poly(N)₁₇), and finally another product that was 1 nucleotide longer than the +19 product (+20 poly(N)₁₈). This prompted us to propose a slippage mechanism that appears to be similar for both the A10 and U10 RNA. In our model, the elongation complex stalls at the end of the homopolymer stretch after transcribing the +12 poly(A)₁₀ product in the absence of additional nucleotides. This promotes a single nucleotide slippage event to form the

+13 poly(A)₁₁ product, that may be followed by a slow structural rearrangement in which polymerase slips back to the 3' end of the poly(U)₁₀ template where it contains four A-U base pairs in the -1 to -4 positions and the 5th poly(U) nucleotide of the template is located in the +1 position. The polymerase is then able to readily incorporate the next six nucleotides across from the poly(U) template to generate the +19 poly(A)₁₇. At this point we think the process may repeat, wherein the stalled complex experiences a single nucleotide slippage and incorporates another nucleotide, generating a +20 poly(A)₁₈ complex, which may help destabilize the duplex enough to allow 3D^{pol} to slip back to the first six bases at the 3'-end of the poly(U) template with two of the bases in the +1 and +2 positions within the polymerase and could subsequently incorporate and additional six nucleotides to generate a +26 poly(A)₂₄ product. This model could be tested fairly easily using slightly modified PETE RNAs that contain longer and shorter poly(A/U) stretches. If this model is correct, then I would expect that longer poly(N) products would not be generated from homopolymeric stretches that are six nucleotides or shorter. Similarly, if this model is correct then I would expect that poly(A/U)_n sequences that are longer than six nucleotides would tend to result in the formation of products that, not including the lock sequence, correspond to a product that is the length of the homopolymer stretch (n), n+1, and 2n-3. For example, from a poly(U)₁₄ template you would observe products corresponding to poly(A)₁₄, poly(A)₁₅, and poly(A)₂₅ transcript lengths. Alternatively, if this is not an accurate model then it may be that the hairpin in the A10 and U10 RNAs are restricting the ability of 3D^{pol} to slip on the template and may explain why are not observing more products from the A10 and U10 that are longer than the +20 poly(N)₁₈. If this is the case, then perhaps modifying the

hairpin in the A10 and U10 RNAs would allow us to better investigate how mutations in the thumb α -helix affect the length of poly(A) transcripts.

5.4 Conclusion

The studies presented in this dissertation demonstrate a diverse approach towards a common goal of better understanding the structure and function of poliovirus and other picornaviral proteins and their roles in the viral replication cycle. Indeed, this work provided important functional and structural insight into the enzymatic activity of the enigmatic poliovirus 2C^{ATPase} as well as the poliovirus RNA dependent RNA polymerase, 3D^{pol}, and left a foundation from which to expand these studies. Both of these proteins play an important role in viral replication and are prime targets for development of antiviral compounds that could help to treat poliovirus as well as other picornaviral species. It is my hope that this research can contribute in some small way to combating these diseases.

REFERENCES

1. **Eggers HJ.** 2002. History of Poliomyelitis and Poliomyelitis Related Research, p. 3-14, Molecular biology of picornaviruses. ASM Press, Washington D.C.
2. **Underwood M.** 1799. A Treatise on the Diseases of Children with Directions for the Management of Infants from the Birth. J. Mathews.
3. **Heine J.** 1840. Beobachtungen über Lähmungszustände der untern Extremitäten und deren Behandlung, von J. Heine. F. H. Köhler, Stuttgart.
4. **Racaniello VR.** 2006. One hundred years of poliovirus pathogenesis. *Virology* **344**:9-16.
5. **Landsteiner K, Popper E.** 1908. Mikroskopische präparate von einem menschlichen und zwei affentückermarker. *Wein klin Wschr* **21**:1930.
6. **Enders JF, Weller TH, Robbins FC.** 1949. Cultivation of the Lansing Strain of Poliomyelitis Virus in Cultures of Various Human Embryonic Tissues. *Science* **109**:85-87.
7. **Mueller S, Wimmer E, Cello J.** 2005. Poliovirus and poliomyelitis: A tale of guts, brains, and an accidental event. *Virus Res.* **111**:175-193.
8. **Nomoto A, Lee YF, Wimmer E.** 1976. The 5' end of poliovirus mRNA is not capped with m7G(5')ppp(5')Np. *Proc. Natl. Acad. Sci. U. S. A.* **73**:375-380.
9. **Jang SK, Kräusslich HG, Nicklin MJ, Duke GM, Palmenberg AC, Wimmer E.** 1988. A segment of the 5' nontranslated region of encephalomyocarditis virus RNA directs internal entry of ribosomes during in vitro translation. *J. Virol.* **62**:2636-2643.
10. **Nomoto A, Detjen B, Pozzatti R, Wimmer E.** 1977. The location of the polio genome protein in viral RNAs and its implication for RNA synthesis. *Nature* **268**:208-213.

11. **Hogle JM, Chow M, Filman DJ.** 1985. Three-dimensional structure of poliovirus at 2.9 Å resolution. *Science* **229**:1358-1365.
12. **Khan S, Peng X, Yin J, Zhang P, Wimmer E.** 2008. Characterization of the New World Monkey Homologues of Human Poliovirus Receptor CD155. *J. Virol.* **82**:7167-7179.
13. **Centers for Disease Control and Prevention.** 2012. Poliomyelitis. In Atkinson W, Wolfe S, Hamborsky J (ed.), *Epidemiology and Prevention of Vaccine-Preventable Diseases*, 12 ed. Public Health Foundation, Washington DC.
14. **De Jesus N.** 2007. Epidemics to eradication: the modern history of poliomyelitis. *Virology journal* **4**:70-70.
15. **Minor PD.** 2004. Polio eradication, cessation of vaccination and re-emergence of disease. *Nat. Rev. Microbiol.* **2**:473-482.
16. **World Health Organization,** posting date. [Online.]
17. **Savolainen-Kopra C, Blomqvist S.** 2010. Mechanisms of genetic variation in polioviruses. *Rev. Med. Virol.* **20**:358-371.
18. **Norder H, De Palma AM, Selisko B, Costenaro L, Papageorgiou N, Arnan C, Coutard B, Lantéz V, De Lamballerie X, Baronti C, Solà M, Tan J, Neyts J, Canard B, Coll M, Gorbalenya AE, Hilgenfeld R.** 2011. Picornavirus non-structural proteins as targets for new anti-virals with broad activity. *Antivir. Res.* **89**:204-218.
19. **Thibaut HJ, De Palma AM, Neyts J.** 2012. Combating enterovirus replication: State-of-the-art on antiviral research. *Biochem. Pharmacol.* **83**:185-192.
20. **Chen T-C, Weng K-F, Chang S-C, Lin J-Y, Huang P-N, Shih S-R.** 2008. Development of antiviral agents for enteroviruses. *J. Antimicrob. Chemother.* **62**:1169-1173.
21. **Brandenburg B, Lee LY, Lakadamyali M, Rust MJ, Zhuang X, Hogle JM.** 2007. Imaging Poliovirus Entry in Live Cells. *PLoS Biol.* **5**:e183.

22. **Coyne CB, Kim KS, Bergelson JM.** 2007. Poliovirus entry into human brain microvascular cells requires receptor-induced activation of SHP-2. *The EMBO journal* **26**:4016-4028.
23. **Murray KE, Barton DJ.** 2003. Poliovirus CRE-dependent VPg uridylylation is required for positive-strand RNA synthesis but not for negative-strand RNA synthesis. *J. Virol.* **77**:4739-4750.
24. **Bergelson JM.** 2008. New (fluorescent) light on poliovirus entry. *Trends Microbiol.* **16**:44-47.
25. **Virgen-Slane R, Rozovics JM, Fitzgerald KD, Ngo T, Chou W, van der Heden van Noort GJ, Filippov DV, Gershon PD, Semler BL.** 2012. An RNA virus hijacks an incognito function of a DNA repair enzyme. *Proceedings of the National Academy of Sciences* **109**:14634-14639.
26. **Rozovics JM, Virgen-Slane R, Semler BL.** 2011. Engineered Picornavirus VPg-RNA Substrates: Analysis of a Tyrosyl-RNA Phosphodiesterase Activity. *PLoS ONE* **6**:e16559.
27. **Kok CC, McMinn PC.** Picornavirus RNA-dependent RNA polymerase. *The International Journal of Biochemistry & Cell Biology* **In Press, Corrected Proof.**
28. **Agol VI, Belov GA, Bienz K, Egger D, Kolesnikova MS, Romanova LI, Sladkova LV, Tolskaya EA.** 2000. Competing Death Programs in Poliovirus-Infected Cells: Commitment Switch in the Middle of the Infectious Cycle. *J. Virol.* **74**:5534-5541.
29. **Jackson WT, Giddings TH, Jr., Taylor MP, Mulinyawe S, Rabinovitch M, Kopito RR, Kirkegaard K.** 2005. Subversion of Cellular Autophagosomal Machinery by RNA Viruses. *PLoS Biol.* **3**:e156.
30. **Kindberg E, Ax C, Fiore L, Svensson L.** 2009. Ala67Thr mutation in the poliovirus receptor CD155 is a potential risk factor for vaccine and wild-type paralytic poliomyelitis. *J. Med. Virol.* **81**:933-936.
31. **Taylor MP, Burgon TB, Kirkegaard K, Jackson WT.** 2009. Role of Microtubules in Extracellular Release of Poliovirus. *J. Virol.* **83**:6599-6609.

32. **Castello A, Alvarez E, Carrasco L.** 2011. The multifaceted poliovirus 2A protease: regulation of gene expression by picornavirus proteases. *J. Biomed. Biotechnol.* **2011**:369648.
33. **Paul AV.** 2002. Possible unifying mechanism of picornavirus genome replication, p. 227-246, *Molecular biology of picornaviruses.* ASM Press.
34. **Bienz K, Egger D, Troxler M, Pasamontes L.** 1990. Structural organization of poliovirus RNA replication is mediated by viral proteins of the P2 genomic region. *J. Virol.* **64**:1156-1163.
35. **Bienz K, Egger D, Pfister T, Troxler M.** 1992. Structural and functional characterization of the poliovirus replication complex. *J. Virol.* **66**:2740-2747.
36. **Belov GA, Nair V, Hansen BT, Hoyt FH, Fischer ER, Ehrenfeld E.** 2012. Complex Dynamic Development of Poliovirus Membranous Replication Complexes. *J. Virol.* **86**:302-312.
37. **Cho MW, Teterina N, Egger D, Bienz K, Ehrenfeld E.** 1994. Membrane rearrangement and vesicle induction by recombinant poliovirus 2C and 2BC in human cells. *Virology* **202**:129-145.
38. **Teterina NL, Gorbalenya AE, Egger D, Bienz K, Ehrenfeld E.** 1997. Poliovirus 2C protein determinants of membrane binding and rearrangements in mammalian cells. *J. Virol.* **71**:8962-8972.
39. **Yin J, Liu Y, Wimmer E, Paul AV.** 2007. Complete protein linkage map between the P2 and P3 non-structural proteins of poliovirus. *J. Gen. Virol.* **88**:2259-2267.
40. **Lin J-Y, Chen T-C, Weng K-F, Chang S-C, Chen L-L, Shih S-R.** 2009. Viral and host proteins involved in picornavirus life cycle. *J. Biomed. Sci.* **16**:103.
41. **Thompson AA, Peersen OB.** 2004. Structural basis for proteolysis-dependent activation of the poliovirus RNA-dependent RNA polymerase. *The EMBO journal* **23**:3462-3471.
42. **Campagnola G, Weygandt M, Scoggin K, Peersen O.** 2008. Crystal structure of coxsackievirus B3 3Dpol highlights the functional importance of residue 5 in picornavirus polymerases. *J. Virol.* **82**:9458-9464.

43. **Salonen A, Ahola T, Kääriäinen L.** 2005. Viral RNA Replication in Association with Cellular Membranes, p. 139-173. *In* Marsh M (ed.), Membrane Trafficking in Viral Replication, vol. 285. Springer Berlin Heidelberg.
44. **Dales S, Eggers HJ, Tamm I, Palade GE.** 1965. Electron microscopic study of the formation of poliovirus. *Virology* **26**:379-389.
45. **Bienz K, Egger D, Rasser Y, Bossart W.** 1980. Kinetics and location of poliovirus macromolecular synthesis in correlation to virus-induced cytopathology. *Virology* **100**:390-399.
46. **Caliguiri LA, Tamm I.** 1970. Characterization of poliovirus-specific structures associated with cytoplasmic membranes. *Virology* **42**:112-122.
47. **Pfister T, Egger D, Bienz K.** 1995. Poliovirus subviral particles associated with progeny RNA in the replication complex. *J. Gen. Virol.* **76**:63-71.
48. **Hsu N-Y, Ilnytska O, Belov G, Santiana M, Chen Y-H, Takvorian PM, Pau C, van der Schaar H, Kaushik-Basu N, Balla T, Cameron CE, Ehrenfeld E, van Kuppeveld FJM, Altan-Bonnet N.** 2010. Viral Reorganization of the Secretory Pathway Generates Distinct Organelles for RNA Replication. *Cell* **141**:799-811.
49. **Belov GA, van Kuppeveld FJM.** 2012. (+)RNA viruses rewire cellular pathways to build replication organelles. *Current Opinion in Virology* **2**:740-747.
50. **Trahey M, Oh HS, Cameron CE, Hay JC.** 2012. Poliovirus Infection Transiently Increases COPII Vesicle Budding. *J. Virol.* **86**:9675-9682.
51. **Aldabe R, Carrasco L.** 1995. Induction of membrane proliferation by poliovirus proteins 2C and 2BC. *Biochem. Biophys. Res. Commun.* **206**:64-76.
52. **Suhy DA, Giddings TH, Jr., Kirkegaard K.** 2000. Remodeling the endoplasmic reticulum by poliovirus infection and by individual viral proteins: an autophagy-like origin for virus-induced vesicles. *J. Virol.* **74**:8953-8965.
53. **Parsley TB, Towner JS, Blyn LB, Ehrenfeld E, Semler BL.** 1997. Poly (rC) binding protein 2 forms a ternary complex with the 5'-terminal sequences of poliovirus RNA and the viral 3CD proteinase. *RNA* **3**:1124-1134.

54. **Gamarnik AV, Andino R.** 1997. Two functional complexes formed by KH domain containing proteins with the 5' noncoding region of poliovirus RNA. *RNA* **3**:882-892.
55. **Herold J, Andino R.** 2001. Poliovirus RNA Replication Requires Genome Circularization through a Protein–Protein Bridge. *Mol. Cell* **7**:581-591.
56. **Nagy PD, Pogany J.** 2012. The dependence of viral RNA replication on co-opted host factors. *Nat Rev Micro* **10**:137-149.
57. **Brunner JE, Ertel KJ, Rozovics JM, Semler BL.** 2010. Delayed kinetics of poliovirus RNA synthesis in a human cell line with reduced levels of hnRNP C proteins. *Virology* **400**:240-247.
58. **Banerjee R, Echeverri A, Dasgupta A.** 1997. Poliovirus-encoded 2C polypeptide specifically binds to the 3'-terminal sequences of viral negative-strand RNA. *J. Virol.* **71**:9570-9578.
59. **Banerjee R, Tsai W, Kim W, Dasgupta A.** 2001. Interaction of poliovirus-encoded 2C/2BC polypeptides with the 3' terminus negative-strand cloverleaf requires an intact stem-loop b. *Virology* **280**:41-51.
60. **Brunner JE, Nguyen JHC, Roehl HH, Ho TV, Swiderek KM, Semler BL.** 2005. Functional Interaction of Heterogeneous Nuclear Ribonucleoprotein C with Poliovirus RNA Synthesis Initiation Complexes. *J. Virol.* **79**:3254-3266.
61. **Rodriguez PL, Carrasco L.** 1995. Poliovirus protein 2C contains two regions involved in RNA binding activity. *J. Biol. Chem.* **270**:10105-10112.
62. **Barton DJ, Flanagan JB.** 1997. Synchronous replication of poliovirus RNA: initiation of negative-strand RNA synthesis requires the guanidine-inhibited activity of protein 2C. *J. Virol.* **71**:8482-8489.
63. **Li JP, Baltimore D.** 1990. An Intragenic Revertant of a Poliovirus 2c Mutant has an Uncoating Defect. *J. Virol.* **64**:1102-1107.
64. **Vance LM, Moscufo N, Chow M, Heinz BA.** 1997. Poliovirus 2C region functions during encapsidation of viral RNA. *J. Virol.* **71**:8759-8765.

65. **Liu Y, Wang C, Mueller S, Paul AV, Wimmer E, Jiang P.** 2010. Direct interaction between two viral proteins, the nonstructural protein 2C and the capsid protein VP3, is required for enterovirus morphogenesis. *PLoS pathogens* **6**:e1001066.
66. **Wang C, Jiang P, Sand C, Paul AV, Wimmer E.** 2012. Alanine Scanning of Poliovirus 2CATPase Reveals New Genetic Evidence that Capsid Protein/2CATPase Interactions Are Essential for Morphogenesis. *J. Virol.* **86**:9964-9975.
67. **Echeverri A, Banerjee R, Dasgupta A.** 1998. Amino-terminal region of poliovirus 2C protein is sufficient for membrane binding. *Virus Res.* **54**:217-223.
68. **Mirzayan C, Wimmer E.** 1994. Biochemical studies on poliovirus polypeptide 2C: evidence for ATPase activity. *Virology* **199**:176-187.
69. **Pfister T, Wimmer E.** 1999. Characterization of the nucleoside triphosphatase activity of poliovirus protein 2C reveals a mechanism by which guanidine inhibits poliovirus replication. *J. Biol. Chem.* **274**:6992-7001.
70. **Rodriguez PL, Carrasco L.** 1993. Poliovirus Protein 2c Has ATPase and GTPase Activities. *J. Biol. Chem.* **268**:8105-8110.
71. **Pfister T, Jones KW, Wimmer E.** 2000. A Cysteine-Rich Motif in Poliovirus Protein 2CATPase Is Involved in RNA Replication and Binds Zinc In Vitro. *J. Virol.* **74**:334-343.
72. **Hanson PI, Whiteheart SW.** 2005. AAA+ proteins: have engine, will work. *Nat. Rev. Mol. Cell Biol.* **6**:519-529.
73. **Bennett-Lovsey RM, Herbert AD, Sternberg MJ, Kelley LA.** 2008. Exploring the extremes of sequence/structure space with ensemble fold recognition in the program Phyre. *Proteins* **70**:611-625.
74. **Erzberger JP, Berger JM.** 2006. Evolutionary relationships and structural mechanisms of AAA+ proteins. *Annu. Rev. Biophys. Biomol. Struct.* **35**:93-114.

75. **Adams P, Kandiah E, Effantin G, Steven AC, Ehrenfeld E.** 2009. Poliovirus 2C protein forms homo-oligomeric structures required for ATPase activity. *J. Biol. Chem.* **284**:22012-22021.
76. **Sweeney TR, Cisnetto V, Bose D, Bailey M, Wilson JR, Zhang X, Belsham GJ, Curry S.** 2010. Foot-and-Mouth Disease Virus 2C Is a Hexameric AAA+ Protein with a Coordinated ATP Hydrolysis Mechanism. *J. Biol. Chem.* **285**:24347-24359.
77. **Baltimore D, Eggers HJ, Franklin RM, Tamm I.** 1963. Poliovirus-induced RNA polymerase and the effects of virus-specific inhibitors on its production. *Proc. Natl. Acad. Sci. U. S. A.* **49**:843-849.
78. **Rightsel WA, Dice JR, McAlpine RJ, Timm EA, McLean IW, Jr., Dixon GJ, Schabel FM, Jr.** 1961. Antiviral Effect of Guanidine. *Science* **134**:558-559.
79. **Pincus SE, Diamond DC, Emini EA, Wimmer E.** 1986. Guanidine-selected mutants of poliovirus: mapping of point mutations to polypeptide 2C. *J. Virol.* **57**:638-646.
80. **Tolskaya EA, Romanova LI, Kolesnikova MS, Gmyl AP, Gorbalenya AE, Agol VI.** 1994. Genetic studies on the poliovirus 2C protein, an NTPase A plausible mechanism of guanidine effect on the 2C function and evidence for the importance of 2C oligomerization. *J. Mol. Biol.* **236**:1310-1323.
81. **Cheng Z, Yang J, Xia H, Qiu Y, Wang Z, Han Y, Xia X, Qin C-F, Hu Y, Zhou X.** 2013. The Nonstructural Protein 2C of a Picorna-Like Virus Displays Nucleic Acid Helix Destabilizing Activity That Can Be Functionally Separated from Its ATPase Activity. *J. Virol.* **87**:5205-5218.
82. **Gong P, Peersen OB.** 2010. Structural basis for active site closure by the poliovirus RNA-dependent RNA polymerase. *Proceedings of the National Academy of Sciences* **107**:22505-22510.
83. **Arnold JJ, Cameron CE.** 2000. Poliovirus RNA-dependent RNA Polymerase (3Dpol). Assembly of stable, elongation-competent complexes by using a symmetrical primer-template substrate (sym/sub). *J. Biol. Chem.* **275**:5329-5336.

84. **Hobdey SE, Kempf BJ, Steil BP, Barton DJ, Peersen OB.** 2010. Poliovirus Polymerase Residue 5 Plays a Critical Role in Elongation Complex Stability. *J. Virol.* **84**:8072-8084.
85. **Gong P, Kortus MG, Nix JC, Davis RE, Peersen OB.** 2013. Structures of Coxsackievirus, Rhinovirus, and Poliovirus Polymerase Elongation Complexes Solved by Engineering RNA Mediated Crystal Contacts. *PLoS ONE* **8**:e60272.
86. **Domingo E, Menéndez-Arias L, Holland JJ.** 1997. RNA virus fitness. *Rev. Med. Virol.* **7**:87-96.
87. **Domingo E, Baranowski E, Escarmis C, Sobrino F, Holland JJ.** 2002. Error Frequencies of Picornavirus RNA Polymerases: Evolutionary Implications for Virus Populations, p. 285-298, *Molecular biology of picornaviruses*. ASM Press, Washington D.C.
88. **Holland JJ, Torre JC, Steinhauer DA.** 1992. RNA Virus Populations as Quasispecies, p. 1-20. *In* Holland J (ed.), *Genetic Diversity of RNA Viruses*, vol. 176. Springer Berlin Heidelberg.
89. **Pfeiffer JK, Kirkegaard K.** 2003. A single mutation in poliovirus RNA-dependent RNA polymerase confers resistance to mutagenic nucleotide analogs via increased fidelity. *Proc. Natl. Acad. Sci. U. S. A.* **100**:7289-7294.
90. **Arnold JJ, Vignuzzi M, Stone JK, Andino R, Cameron CE.** 2005. Remote Site Control of an Active Site Fidelity Checkpoint in a Viral RNA-dependent RNA Polymerase. *J. Biol. Chem.* **280**:25706-25716.
91. **Pfeiffer JK, Kirkegaard K.** 2005. Increased Fidelity Reduces Poliovirus Fitness and Virulence under Selective Pressure in Mice. *PLoS pathogens* **1**:e11.
92. **Vignuzzi M, Wendt E, Andino R.** 2008. Engineering attenuated virus vaccines by controlling replication fidelity. *Nat. Med.* **14**:154-161.
93. **Roberts L.** 2012. The Polio Emergency. *Science* **337**:514-516.
94. **Boldog T, Grimme S, Li M, Sligar SG, Hazelbauer GL.** 2006. Nanodiscs separate chemoreceptor oligomeric states and reveal their signaling properties. *Proceedings of the National Academy of Sciences* **103**:11509-11514.

95. **Leitz AJ, Bayburt TH, Barnakov AN, Springer BA, Sligar SG.** 2006. Functional reconstitution of Beta2-adrenergic receptors utilizing self-assembling Nanodisc technology. *BioTechniques* **40**:601-602.
96. **Flanegan JB, Van Dyke TA.** 1979. Isolation of a soluble and template-dependent poliovirus RNA polymerase that copies virion RNA in vitro. *J. Virol.* **32**:155-161.
97. **Harris KS, Reddigari SR, Nicklin MJ, Hammerle T, Wimmer E.** 1992. Purification and characterization of poliovirus polypeptide 3CD, a proteinase and a precursor for RNA polymerase. *J. Virol.* **66**:7481-7489.
98. **Limpens RW, van der Schaar HM, Kumar D, Koster AJ, Snijder EJ, van Kuppeveld FJ, Barcena M.** 2011. The transformation of enterovirus replication structures: a three-dimensional study of single- and double-membrane compartments. *mBio* **2**.
99. **Schlegel A, Giddings TH, Jr., Ladinsky MS, Kirkegaard K.** 1996. Cellular origin and ultrastructure of membranes induced during poliovirus infection. *J. Virol.* **70**:6576-6588.
100. **Agirre A, Barco A, Carrasco L, Nieva JL.** 2002. Viroporin-mediated membrane permeabilization. Pore formation by nonstructural poliovirus 2B protein. *J. Biol. Chem.* **277**:40434-40441.
101. **Agirre A, Lorizate M, Nir S, Nieva JL.** 2008. Poliovirus 2b insertion into lipid monolayers and pore formation in vesicles modulated by anionic phospholipids. *Biochim. Biophys. Acta* **1778**:2621-2626.
102. **Nieva JL, Agirre A, Nir S, Carrasco L.** 2003. Mechanisms of membrane permeabilization by picornavirus 2B viroporin. *FEBS Lett.* **552**:68-73.
103. **Paul AV, Molla A, Wimmer E.** 1994. Studies of a putative amphipathic helix in the N-terminus of poliovirus protein 2C. *Virology* **199**:188-199.
104. **Echeverri AC, Dasgupta A.** 1995. Amino terminal regions of poliovirus 2C protein mediate membrane binding. *Virology* **208**:540-553.

105. **Strauss DM, Glustrom LW, Wuttke DS.** 2003. Towards an understanding of the poliovirus replication complex: the solution structure of the soluble domain of the poliovirus 3A protein. *J. Mol. Biol.* **330**:225-234.
106. **Towner JS, Ho TV, Semler BL.** 1996. Determinants of membrane association for poliovirus protein 3AB. *J. Biol. Chem.* **271**:26810-26818.
107. **Teterina NL, Levenson E, Rinaudo MS, Egger D, Bienz K, Gorbalenya AE, Ehrenfeld E.** 2006. Evidence for functional protein interactions required for poliovirus RNA replication. *J. Virol.* **80**:5327-5337.
108. **Bayburt TH, Sligar SG.** 2003. Self-assembly of single integral membrane proteins into soluble nanoscale phospholipid bilayers. *Protein Sci.* **12**:2476-2481.
109. **Civjan NR, Bayburt TH, Schuler MA, Sligar SG.** 2003. Direct solubilization of heterologously expressed membrane proteins by incorporation into nanoscale lipid bilayers. *BioTechniques* **35**:556-563.
110. **Katzen F, Fletcher JE, Yang J-P, Kang D, Peterson TC, Cappuccio JA, Blanchette CD, Sulchek T, Chromy BA, Hoepflich PD, Coleman MA, Kudlicki W.** 2008. Insertion of Membrane Proteins into Discoidal Membranes Using a Cell-Free Protein Expression Approach. *J. Proteome Res.* **7**:3535-3542.
111. **Gohara DW, Ha CS, Kumar S, Ghosh B, Arnold JJ, Wisniewski TJ, Cameron CE.** 1999. Production of "Authentic" Poliovirus RNA-Dependent RNA Polymerase (3Dpol) by Ubiquitin-Protease-Mediated Cleavage in *Escherichia coli*. *Protein Expr. Purif.* **17**:128-138.
112. **Denisov IG, Baas BJ, Grinkova YV, Sligar SG.** 2007. Cooperativity in Cytochrome P450 3A4. *J. Biol. Chem.* **282**:7066-7076.
113. **Denisov IG, Grinkova YV, Lazarides AA, Sligar SG.** 2004. Directed Self-Assembly of Monodisperse Phospholipid Bilayer Nanodiscs with Controlled Size. *J. Am. Chem. Soc.* **126**:3477-3487.
114. **Barnett RE.** 1970. Effect of monovalent cations on the ouabain inhibition of the sodium and potassium ion activated adenosine triphosphatase. *Biochemistry* **9**:4644-4648.

115. **Russell R, Jordan R, McMacken R.** 1998. Kinetic Characterization of the ATPase Cycle of the DnaK Molecular Chaperone†. *Biochemistry* **37**:596-607.
116. **Denisov IG, McLean MA, Shaw AW, Grinkova YV, Sligar SG.** 2005. Thermotropic phase transition in soluble nanoscale lipid bilayers. *The Journal of Physical Chemistry. B* **109**:15580-15588.
117. **Nyborg JK, Peersen OB.** 2004. That zincing feeling: the effects of EDTA on the behaviour of zinc-binding transcriptional regulators. *Biochem. J.* **381**:e3-4.
118. **Falke JJ, Drake SK, Hazard AL, Peersen OB.** 1994. Molecular tuning of ion binding to calcium signaling proteins. *Q. Rev. Biophys.* **27**:219-290.
119. **Wendler P, Ciniawsky S, Kock M, Kube S.** 2012. Structure and function of the AAA+ nucleotide binding pocket. *Biochim. Biophys. Acta* **1823**:2-14.
120. **Palmenberg AC.** 1990. PROTEOLYTIC PROCESSING OF PICORNAVIRAL POLYPROTEIN. *Annu. Rev. Microbiol.* **44**:603-623.
121. **Singleton MR, Dillingham MS, Wigley DB.** 2007. Structure and mechanism of helicases and nucleic acid translocases. *Annu. Rev. Biochem.* **76**:23-50.
122. **Gorbalenya AE, Koonin EV, Wolf YI.** 1990. A new superfamily of putative NTP-binding domains encoded by genomes of small DNA and RNA viruses. *FEBS Lett.* **262**:145-148.
123. **Gorbalenya AE, Koonin EV.** 1993. Helicases: amino acid sequence comparisons and structure-function relationships. *Curr. Opin. Struct. Biol.* **3**:419-429.
124. **James JA, Escalante CR, Yoon-Robarts M, Edwards TA, Linden RM, Aggarwal AK.** 2003. Crystal Structure of the SF3 Helicase from Adeno-Associated Virus Type 2. *Structure* **11**:1025-1035.
125. **Li D, Zhao R, Lilyestrom W, Gai D, Zhang R, DeCaprio JA, Fanning E, Jochimiak A, Szakonyi G, Chen XS.** 2003. Structure of the replicative helicase of the oncoprotein SV40 large tumour antigen. *Nature* **423**:512-518.

126. **Castella S, Burgin D, Sanders CM.** 2006. Role of ATP hydrolysis in the DNA translocase activity of the bovine papillomavirus (BPV-1) E1 helicase. *Nucleic Acids Res.* **34**:3731-3741.
127. **Mestas SP, Sholders AJ, Peersen OB.** 2007. A fluorescence polarization-based screening assay for nucleic acid polymerase elongation activity. *Anal. Biochem.* **365**:194-200.
128. **Gnädig NF, Beaucourt S, Campagnola G, Bordería AV, Sanz-Ramos M, Gong P, Blanc H, Peersen OB, Vignuzzi M.** 2012. Coxsackievirus B3 mutator strains are attenuated in vivo. *Proceedings of the National Academy of Sciences* **109**:E2294–E2303.
129. **Gong P, Campagnola G, Peersen OB.** 2009. A quantitative stopped-flow fluorescence assay for measuring polymerase elongation rates. *Anal. Biochem.* **391**:45-55.
130. **Kortus MG, Kempf BJ, Haworth KG, Barton DJ, Peersen OB.** 2012. A Template RNA Entry Channel in the Fingers Domain of the Poliovirus Polymerase. *J. Mol. Biol.* **417**:263-278.
131. **Pagano JM, Clingman CC, Ryder SP.** 2011. Quantitative approaches to monitor protein-nucleic acid interactions using fluorescent probes. *RNA* **17**:14-20.
132. **Ahlquist P, Kaesberg P.** 1979. Determination of the length distribution of poly(A) at the 3' terminus of the virion RNAs of EMC virus, poliovirus, rhinovirus, RAV-61 and CPMV and of mouse globin mRNA. *Nucleic Acids Res.* **7**:1195-1204.
133. **Spector DH, Baltimore D.** 1975. Polyadenylic acid on poliovirus RNA. II. poly(A) on intracellular RNAs. *J. Virol.* **15**:1418-1431.
134. **Paul AV, van Boom JH, Filippov D, Wimmer E.** 1998. Protein-primed RNA synthesis by purified poliovirus RNA polymerase. *Nature* **393**:280-284.
135. **Bergamini G, Preiss T, Hentze MW.** 2000. Picornavirus IRESes and the poly(A) tail jointly promote cap-independent translation in a mammalian cell-free system. *RNA* **6**:1781-1790.

136. **Ogram SA, Spear A, Sharma N, Flanegan JB.** 2010. The 5' CL-PCBP RNP complex, 3' poly(A) tail and 2Apro are required for optimal translation of poliovirus RNA. *Virology* **397**:14-22.
137. **Sarnow P.** 1989. Role of 3'-end sequences in infectivity of poliovirus transcripts made in vitro. *J. Virol.* **63**:467-470.
138. **Spector DH, Baltimore D.** 1974. Requirement of 3' -Terminal Poly(adenylic acid) for the Infectivity of Poliovirus RNA. *Proceedings of the National Academy of Sciences* **71**:2983-2987.
139. **Silvestri LS, Parilla JM, Morasco BJ, Ogram SA, Flanegan JB.** 2006. Relationship between poliovirus negative-strand RNA synthesis and the length of the 3' poly(A) tail. *Virology* **345**:509-519.
140. **Svitkin YV, Costa-Mattioli M, Herdy B, Perreault S, Sonenberg N.** 2007. Stimulation of picornavirus replication by the poly(A) tail in a cell-free extract is largely independent of the poly(A) binding protein (PABP). *RNA* **13**:2330-2340.
141. **Steil BP, Kempf BJ, Barton DJ.** 2010. Poly(A) at the 3' End of Positive-Strand RNA and VPg-Linked Poly(U) at the 5' End of Negative-Strand RNA Are Reciprocal Templates during Replication of Poliovirus RNA. *J. Virol.* **84**:2843-2858.
142. **Kempf BJ, Kelly MM, Springer CL, Peersen OB, Barton DJ.** 2013. Structural features of a picornavirus polymerase involved in the polyadenylation of viral RNA. *J. Virol.* **87**:5629-5644.
143. **Freistadt MS, Vaccaro JA, Eberle KE.** 2007. Biochemical characterization of the fidelity of poliovirus RNA-dependent RNA polymerase. *Virology journal* **4**:44.
144. **Arias A, Arnold JJ, Sierra M, Smidansky ED, Domingo E, Cameron CE.** 2008. Determinants of RNA-dependent RNA polymerase (in)fidelity revealed by kinetic analysis of the polymerase encoded by a foot-and-mouth disease virus mutant with reduced sensitivity to ribavirin. *J. Virol.* **82**:12346-12355.
145. **Turnbough Jr CL.** 2011. Regulation of gene expression by reiterative transcription. *Curr. Opin. Microbiol.* **14**:142-147.

146. **Barr JN, Wertz GW.** 2001. Polymerase slippage at vesicular stomatitis virus gene junctions to generate poly(A) is regulated by the upstream 3'-AUAC-5' tetranucleotide: implications for the mechanism of transcription termination. *J. Virol.* **75**:6901-6913.
147. **Barr JN, Fearn R.** 2010. How RNA viruses maintain their genome integrity. *J. Gen. Virol.* **91**:1373-1387.
148. **Autexier C, Lue NF.** 2006. The Structure and Function of Telomerase Reverse Transcriptase. *Annu. Rev. Biochem.* **75**:493-517.
149. **Cho MW, Richards OC, Dmitrieva TM, Agol V, Ehrenfeld E.** 1993. RNA duplex unwinding activity of poliovirus RNA-dependent RNA polymerase 3Dpol. *J. Virol.* **67**:3010-3018.
150. **Thompson AA, Albertini RA, Peersen OB.** 2007. Stabilization of Poliovirus Polymerase by NTP Binding and Fingers-Thumb Interactions. *J. Mol. Biol.* **366**:1459-1474.
151. **Van Der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJC.** 2005. GROMACS: Fast, flexible, and free. *J Comput Chem* **26**:1701-1718.
152. **Jorgensen WL, Maxwell DS, Tirado-Rives J.** 1996. Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *J. Am. Chem. Soc.* **118**:11225-11236.

APPENDIX I

MOLECULAR DYNAMICS SIMULATIONS OF POLIOVIRUS 3D^{pol}

The material in this appendix summarizes the in silico molecular dynamics simulations I carried out on poliovirus 3D^{pol} during my first year in the Peersen laboratory to investigate structural changes associated with polymerase activity as there was no structure of an RNA bound elongation complex available at that time. However, these simulations were put on hold when the crystal structure of the 3D^{pol} elongation complex was solved by Peng Gong in our laboratory (82). At that point I began working on the 2C protein and these studies were not resumed.

AI.1 Introduction

PV 3D^{pol} is an RNA-dependent RNA polymerase comprised of 461 amino acids and is solely responsible for synthesizing both positive- and negative- sense viral RNA. Our laboratory published the first complete structure of picornaviral 3D^{pol} demonstrating its structure is analogous to a "right hand", having thumb, palm and fingers domains (as discussed in section 1.4.1 and shown in Fig. 1.10 (41)). The index finger forms the template RNA entry channel, playing a key role in maintaining 3D^{pol} stability and has been implicated in incorporating incoming NTPs and (41, 42, 150). Contact between the index finger and thumb is maintained by hydrophobic interactions involving Phe30 and Phe34, and thermal denaturation studies suggest that 3D^{pol} unfolding is linked to the disruption of these interactions (Fig. AI.1A) (150). A key feature within the index

finger of 3D^{pol} is the burial of the N-terminus, Gly1, into a pocket at the back of the palm, which is held in place via a total of four hydrogen bonds: two with residues 239 and 241, and two with Gly64 (Fig. AI.1B) (41), and deletion or addition one residue to the N-terminus or mutating Gly1 abolishes 3D^{pol} activity. The index finger then continues with residues 2-9 being hydrogen bonded to the middle finger in a canonical β -sheet conformation, with the exception of Trp5, which deviates from this secondary structure and is flipped out into solution. This seeming anomaly of an exposed large hydrophobic residue is actually highly conserved among picornavirus polymerases, such as Phe5 in coxsackie virus 3D^{pol} (Fig. AI.1C) (42). Initial biochemical and biophysical analysis revealed that mutations at these sites have significant effects on 3D^{pol} activity. Our laboratory has found that residues Trp5 and Gly64 are important for formation and stability of the processive 3D^{pol} elongation complex ((84) and unpublished results). For example, mutations at Gly64 to Ala, Ser, Asp and Asn have a reduced elongation rate as compared to wild type and have a $K_{m,app}$ similar to that of wild type (K_m and V_{max} values for G64S is presented in Chapter 3 of this dissertation, data for other G64 mutants not shown). Mutational analysis of Phe5 in coxsackievirus 3D^{pol} demonstrates that 3D^{pol} activity decreases in relation to the decreasing size of the hydrophobic residue (42). Burial of Gly1 is stabilized by hydrogen bonding interactions with Gly64 and, prior to solving the crystal structure of the elongation complex, we thought it may help anchor movements within the index finger domain, such as a possible flipping motion of Trp5 into a hydrophobic pocket in the ring finger (Fig. AI.1D). We postulated that Trp5 and Gly64 function were linked via Gly1, and that burial of Gly1 in the pocket in the thumb was needed to provide a sturdy anchor that may allow Trp5 to flip into the

hydrophobic pocket during EC formation and that larger hydrophobics at residue 5 may help lock in this conformation (42). In order to test these theories and supplement our current biochemical data, I carried out molecular dynamics (MD) simulations of 3D^{pol} containing mutations at residue 5 to and residue 64. Additionally, I modeled the thermal denaturation of wild type 3D^{pol} in order to validate my MD simulation. I did not observe residue 5 flipping into the hydrophobic pocket in the ring finger during my simulations of both wild type and mutant polymerases, thus these simulations are not discussed in this appendix. Simulations were done with the GROMACS software program (151) using the OPLSAA force field (152) and the calculations were carried out on the CSU Center for Bioinformatics computer cluster. The protocols for running each MD experiment includes adding waters and ions to our protein system and performing various rounds of energy minimization and position restrained equilibration before carrying out the full MD production run.

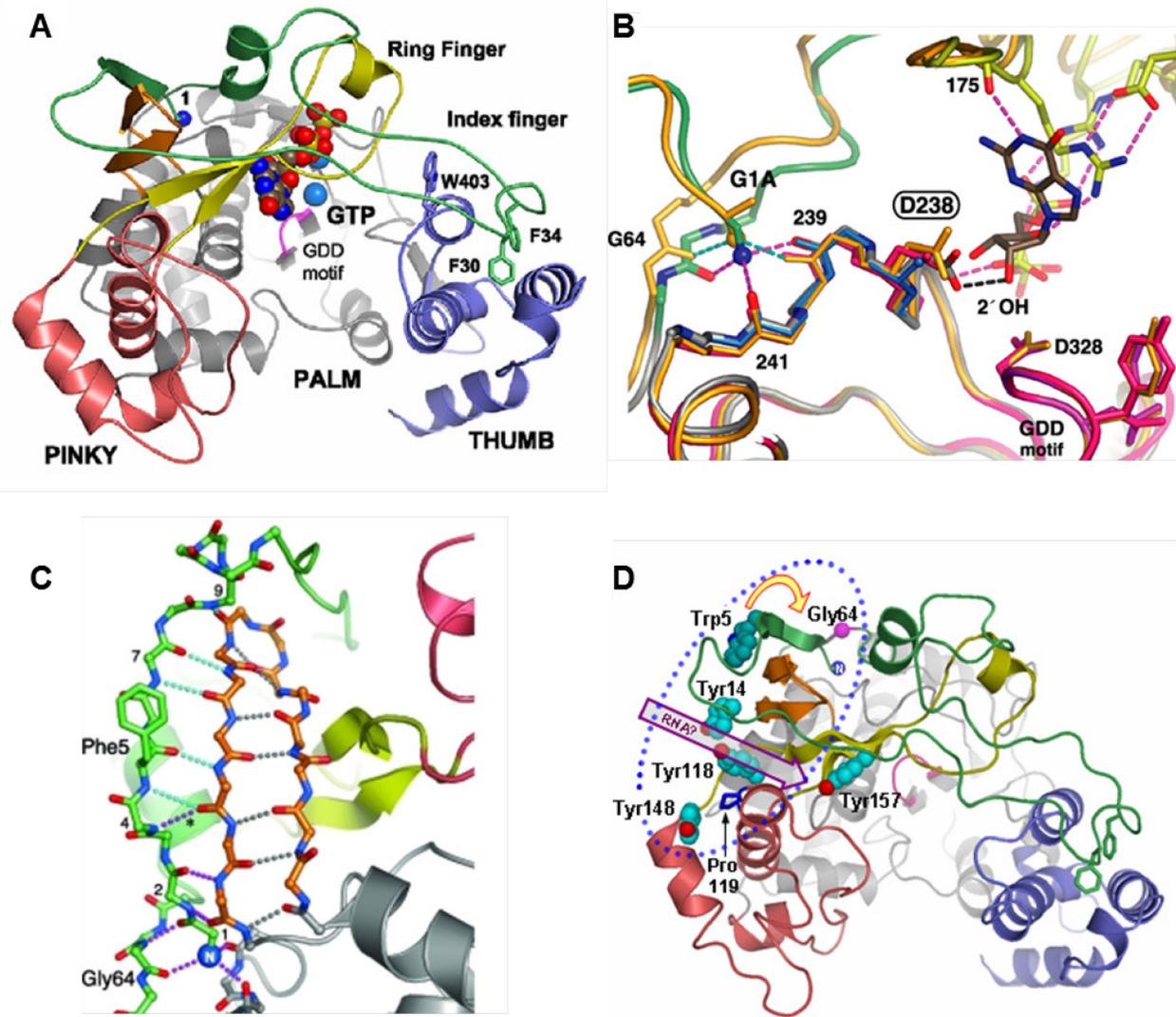


Fig. A1.1. Crystal structures of 3D^{pol} showing important structural features.

A) PV 3D^{pol} with GTP bound indicating residues important for maintaining structural integrity. Gly1 is shown as a blue sphere on the upper left of the polymerase, the active site GDD motif is shown in magenta, and the aromatic residues that mediate interactions between the thumb domain (blue) and index finger (green) are shown in licorice (150). B) Close up of the PV 3D^{pol} buried N-terminus, Gly1, forming three hydrogen bonds with residues 64, 239, and 241 that help position Asp238 in the active site to interact with the 2'OH of the incoming NTP.(41). C) Crystal structure of coxsackie polymerase showing the deviation of Phe5 from the canonical β -sheet and the hydrogen bonding interactions between Gly64 and Gly1 that may help anchor Phe5 (or Trp5 in PV) (42). D) PV 3D^{pol} crystal structure in which the orange arrow indicates the hypothesized motion of Trp5 into a pocket between the middle and index fingers. The magenta arrow indicates the path of template RNA.

AI.2 Results and Discussion

AI.2.1 Validation of MD Simulations by Modeling 3D^{pol} Thermal Denaturation

In order to determine whether our MD simulations were valid, we carried out simulations of wild type 3D^{pol} at various temperatures to model protein thermal denaturation. Our laboratory has shown that 3D^{pol} unfolds at ~40°C in a process which appears to occur two steps (150). First, the pinky finger becomes very dynamic, fluctuating rapidly as the temperature increases and exhibiting molten globule behavior. Second, as the temperature reaches approximately 40°C, there is an abrupt unfolding attributed to the disruption of hydrophobic interactions between the index finger and thumb. Phe30 and Phe34 in the index finger appear to be important in maintaining these interactions as mutating these residues to alanine or aspartic acid decreases the melting temperature ~4°C (150). To validate our MD simulations, I carried out MD simulations of wild type 3D^{pol} at temperatures ranging from 298K to 498K. As shown in Fig. AI.2, coloring by root mean square fluctuation (RMSF) of 3D^{pol} simulated at 498K reveals that the pinky finger and the thumb display the largest amount of fluctuation, and fluctuations in the thumb disrupt the hydrophobic interactions with Phe30 and Phe34 in the index finger. Preliminary analysis of averaged root mean square fluctuations (RMSF) per 3D^{pol} domain at 298K, 323K and 498K showed that the pinky finger does indeed exhibit more fluctuation as the temperature increases, followed by the thumb and then the index finger (Fig. AI.3).

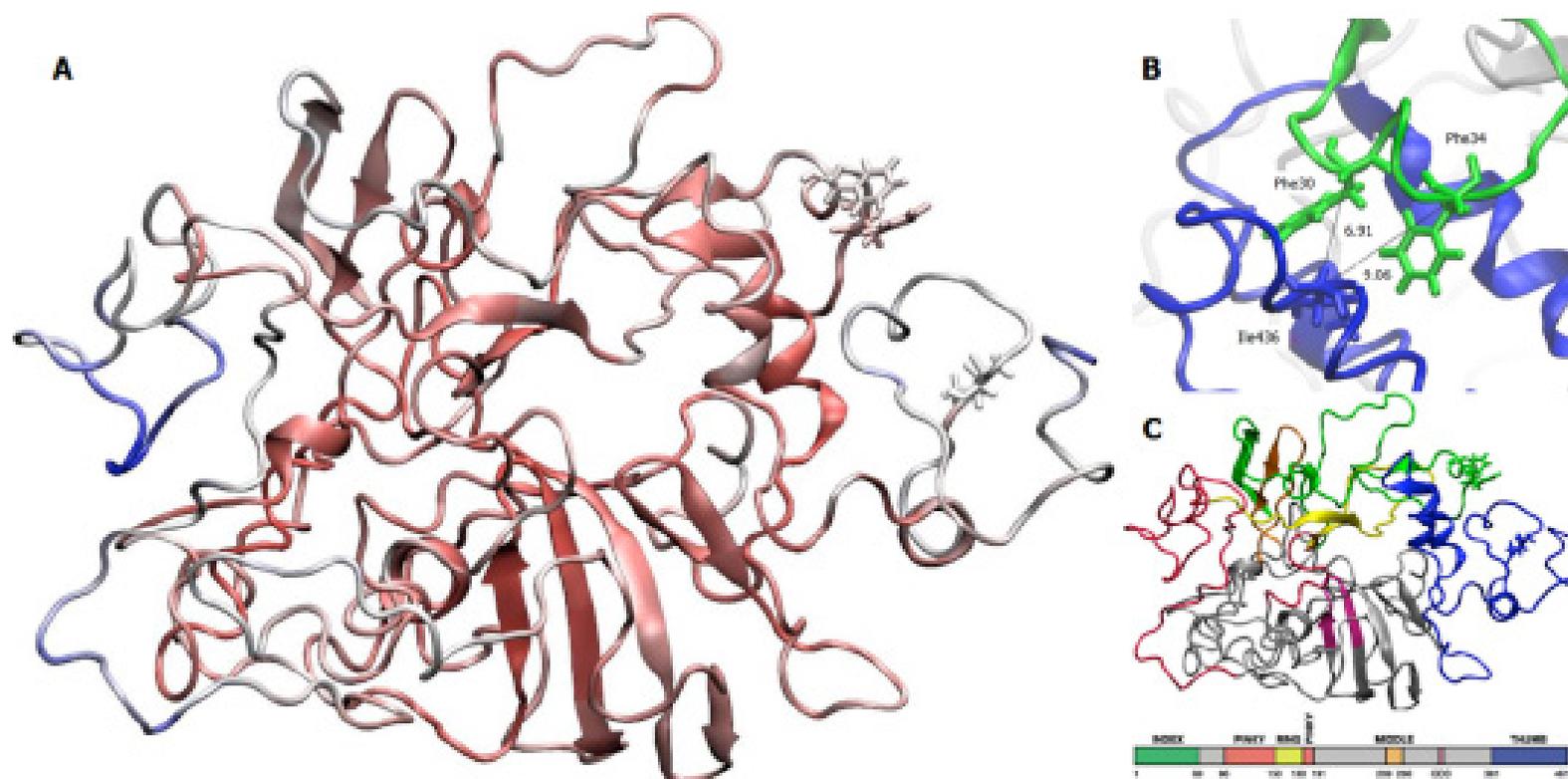


Fig. AI.2. 3D^{pol} MD simulation at 498K.

Phe30 and Phe34 of the index finger and Ile436 of the thumb are drawn in licorice to demonstrate disruption between the index finger and thumb. A) Last frame of ~15 ns simulation colored by RMSF. Blue indicates regions that are most dynamic and red regions are least dynamic. B) Close up of hydrophobic interactions between the index finger (green) and thumb (blue) at the beginning of the simulation (distances in nm). C) Fig. AI.2A in standard 3D^{pol} coloring.

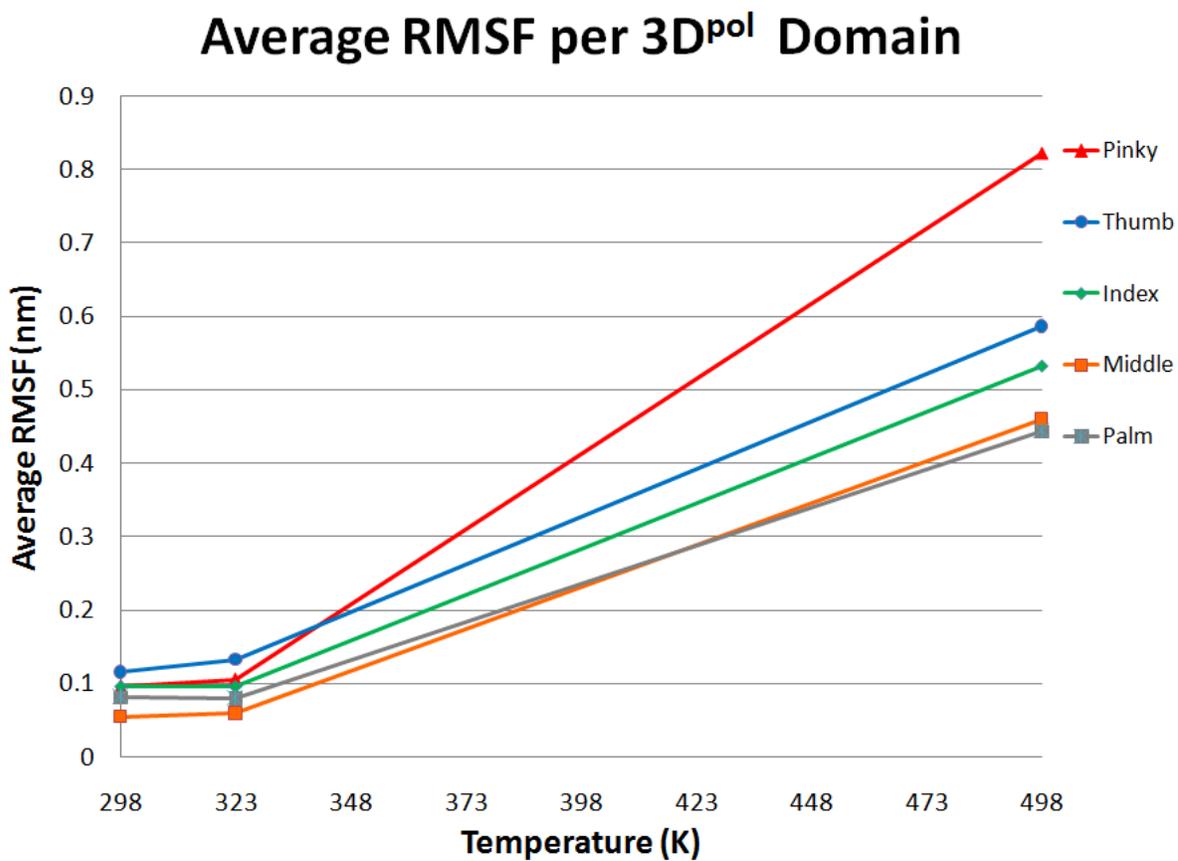


Fig. AI.3. Average RMSF of 3D^{pol} domains at 298K, 323K, and 498K.

Root mean square fluctuations (RMSF) of residues in the pinky, thumb, index finger, middle finger, or palm domains were averaged at each temperature.

AI.2.2 MD Simulations of G64 Mutants

Preliminary analysis of 20 ns simulations carried out with the Gly64 mutants (listed above) and wild type 3D^{pol} showed structural deviations between the simulations and the crystal structure and compared to one another (Fig. AI.4). Within the first 3 ns, all the simulations show a flip in Asp238 in which it forms a salt bridge with Lys61, an interaction that had not been previously observed in the apo crystal structure whereas a salt bridge between Lys61 and Glu177 had been observed (41) (Fig. AI.4, bottom panels). Interestingly, the interaction between Asp238 and Lys61 appears to be quite stable only in the G64S and G64N mutants, (Fig. AI.4, maroon and green, respectively) compared to wild type and the other mutants. The salt bridge between Lys61 and Glu177 that had been originally observed in the crystal structure appears to be a fairly dynamic interaction in the wild type, G64A, and G64D 3D^{pol} simulations. Alternatively, this Lys61-Glu177 interaction is lost \approx 8 ns into the G64S simulation and is not reformed during the remainder of the 20 ns simulation and this interaction is actually maintained throughout the entire G64N simulation. This data suggests that there may be a correlation between the mobility of Asp238 and fidelity, wherein introduction of a serine at residue 64 shifts the position of residue Lys61, bringing it in closer proximity to Asp238 and promoting the formation of a salt bridge between these two residues.

It has been previously suggested by our laboratory that the high fidelity imparted by the G64S mutation may be related to altered positioning of Asp238 (41, 82). Indeed, the recent structure of the elongation complex at various stages of catalysis demonstrates that active site closure requires movement of Asp238 out of the active

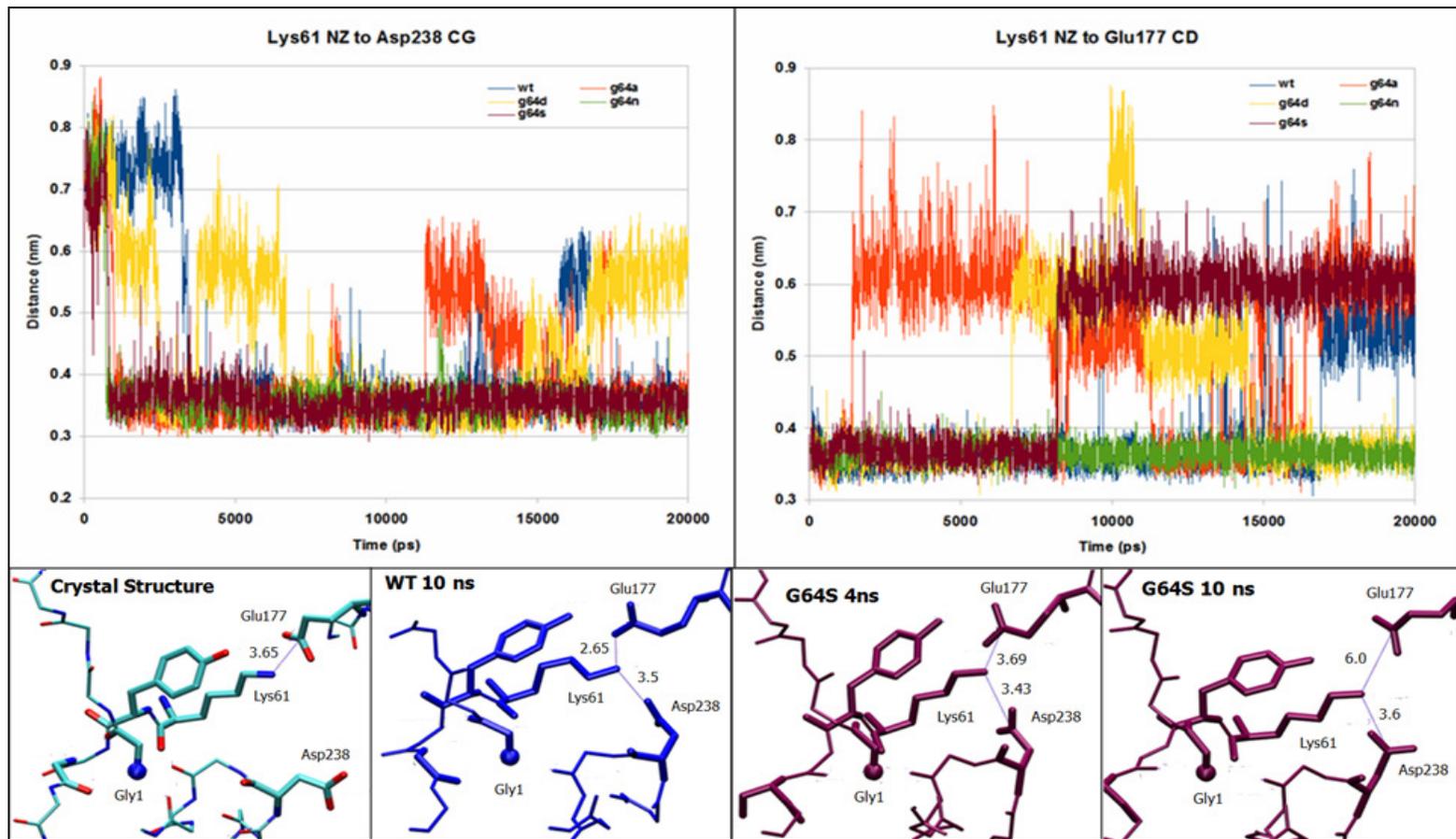


Fig. A1.4. Inter-residue distances among 3D^{pol} Gly64 mutants.

Top Panels: Distances between Lys61-Asp238 and Lys61-Glu177 for WT 3D^{pol} and the four G64 mutants. (WT in blue, G64A in orange, G64D in yellow, G64N in green, G64S in maroon)

Bottom Panels (from left to right): Wild type crystal structure; wild type simulation at 10 ns reflecting altered position of Asp238; G64S simulation at 4 ns and 10 ns reflecting the two different interactions observed between Lys61 and Glu177. Note that the distances shown here are to the carboxylic acid carbon.

site upon NTP binding, a conformation that is stabilized via interactions with Lys61 and Ser288, which leads to the repositioning of Asn297 (82). These movements result in the formation of extensive a hydrogen bonding network with the ribose hydroxyls that helps position the NTP for catalysis and trigger active site closure. Thus, it is possible that the tight interaction observed between Asp238 and Lys61 *in silico* may be reflecting enhanced affinity between these two residues *in vivo* that may contribute to altering the energetics associated with NTP positioning and active site closure to increase polymerase fidelity. In conclusion, these simulations reveal some of the more subtle factors that may contribute to 3D^{pol} structure and fidelity.