INVESTIGATION OF THE NUCLEOTIDE SELECTION MECHANISM OF
POLIOVIRUS RNA-DEPENDENT RNA POLYMERASE

A Dissertation in
Chemistry
by
Xinran Liu

© 2015 Xinran Liu

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

December 2015
The dissertation of Xinran Liu was reviewed and approved* by the following:

David D. Boehr  
Associate Professor of Chemistry  
Dissertation Advisor  
Chair of Committee

Craig E. Cameron  
Professor of Biochemistry and Molecular Biology  
Eberly Chair in Biochemistry and Molecular Biology

Scott A. Showalter  
Associate Professor of Chemistry  
Associate Professor of Biochemistry and Molecular Biology

John H. Golbeck  
Professor of Biochemistry and Biophysics  
Professor of Chemistry

Kenneth S. Feldman  
Professor of Chemistry  
Head of the Department of Chemistry

*Signatures are on file in the Graduate School
ABSTRACT

RNA viruses cause many diseases including severe acute respiratory syndrome (SARS), the common cold, hepatitis C, poliomyelitis, and so on. However, antiviral strategies against RNA virus infections are very limited. For example, there are no FDA approved (Food and Drug Administration) antiviral compounds for the treatment of picornavirus infection. Only three vaccines are available to prevent the transmission of picornaviruses. The severity of public health issues associated with these viruses and the scarcity of treatment options make the development of antiviral drugs and vaccines high priorities. One very promising antiviral target is the virally encoded RNA-dependent RNA polymerase (RdRp). The RdRp is the most conserved protein among RNA viruses. One antiviral strategy is to modulate the RdRp’s error rate of nucleotide incorporation. The working principle of this antiviral strategy derives from the ‘quasispecies’ nature of RNA viruses. RNA viruses utilize the error-prone RdRp to replicate a genetically diverse population where viral genomes do not contain a unique sequence but a pool of genetically variable sequences. It has been shown that mere two-fold changes in the RdRp error rate (either higher or lower error) lead to viral attenuation. Increasing the viral mutation rate through the action of nucleoside analogs leads to lethal mutagenesis due to the accumulation of an excess amount of mutations and loss of viable genetic information. Decreasing the viral mutation rate is also detrimental to the virus due to a constrained viral ability to adapt to the host environment. Understanding the nucleotide selection mechanisms of RdRps would therefore opens up new treatment strategies including the development of live, attenuated vaccines and/or antiviral drugs.

Poliovirus (PV) RdRp is a great model system to study the nucleotide selection mechanism. This dissertation mainly focuses on investigations into the fidelity mechanism of PV RdRp via a combination of kinetic and NMR experiments. Among all the DNA and RNA
polymerases, the prechemistry conformational change is a critical fidelity checkpoint. In RdRps and other polymerases, it has been proposed that this step involves the rearrangement of the triphosphate group and nucleobase of the incoming nucleotide into productive conformation, and the repositioning of a general acid to help catalyze phosphodiester bond formation. In PV RdRp, conserved structural motif D is responsible for the repositioning of the general acid via an “open” to a “closed” state transition during the prechemistry conformational change. In this dissertation, I show that the T362I substitution, which originates from the Sabin 1 vaccine strain, lowers enzyme fidelity by shifting the motif D equilibrium more to the “closed” state. PV encoding this low fidelity variant is more sensitive to ribavirin and is moderately attenuated in a mouse model. Other Sabin substitutions in the RdRp also change catalysis and fidelity. I show that the Sabin RdRp, which has all four substitutions (i.e. D53N, Y73H, K250E and T362I), discriminates against nucleotides with noncognate nucleobase to the same extent as wild-type (WT) enzyme, but more efficiently catalyzes incorporation of 2’-deoxy nucleotides. My studies suggest that there is more selective pressure to maintain nucleobase discrimination than sugar discrimination in the Sabin vaccine strain.

Besides motif D, motif F is another structure that we propose is involved in nucleotide selection. I propose that substitutions on motif F lead to changes in RdRp fidelity by perturbing triphosphate conformations necessary for efficient nucleotide addition. My studies on the nucleotide selection mechanism of PV RdRp likely extend to RdRps of other viruses due to the strong structural and functional similarities among the RdRps. As such, my studies open up new possibilities for engineering attenuated vaccine candidates through modifications of motifs D and F in these RNA viruses. Such developments will be critical in the treatment of current and future virus outbreaks.
TABLE OF CONTENTS

List of Figures ................................................................................................................. ix

List of Tables ..................................................................................................................... xvi

List of Abbreviations ......................................................................................................... xviii

Acknowledgements ........................................................................................................... xxi

Chapter 1 Introduction ....................................................................................................... 1

  1.1 Viruses ....................................................................................................................... 1

  1.2 Picornaviruses .......................................................................................................... 2

  1.3 Historical perspective on antiviral strategies ............................................................ 7

  1.4 New antiviral strategies targeting RdRp fidelity ....................................................... 9

  1.5 PV as a model system for understanding RdRp function ......................................... 10

  1.6 Catalysis of PV RdRp ............................................................................................... 11

  1.7 Structure and function relationship in the fidelity of PV RdRp ............................... 15

  1.8 Summary .................................................................................................................. 18

  1.9 References ............................................................................................................... 19

Chapter 2 Vaccine-Derived Mutation in Motif D of Poliovirus RNA-dependent RNA
Polymerase Lowers NTP Incorporation Fidelity ................................................................. 24

  2.1 Introduction ................................................................................................................ 24

  2.2 Methods .................................................................................................................... 27

      2.2.1 Materials ............................................................................................................. 27

      2.2.2 Plasmid construction ......................................................................................... 28

      2.2.3 Expression and purification of RdRp ................................................................. 28

      2.2.4 Purification, 5'-32P end labeling and annealing of sym/sub ........................... 30

      2.2.5 PV RdRp assay46 ............................................................................................. 31

      2.2.6 Stopped flow experiments36 ............................................................................. 32

      2.2.7 Benchtop assay for slow nucleotide incorporation47 ...................................... 33

      2.2.8 Product analysis and data analysis ................................................................. 33

      2.2.9 Sample preparation for NMR spectroscopy43 ................................................ 33

      2.2.10 1H - 13C HSQC NMR experiment43 ............................................................. 34

  2.3 Results ....................................................................................................................... 34

      2.3.1 The T362I substitution lowers RdRp fidelity by making the pre-chemistry
conformational change for misincorporation more efficient ...................................... 34

      2.3.2 Solution-state NMR indicates that the T362I substitution induces a
structural rearrangement within motif D to lower RdRp fidelity .............................. 44

      2.3.3 The T362I substitution results in long-range structural changes that may
allow for more efficient misincorporation ................................................................. 47

      2.3.4 T362I PV expresses a mutator phenotype in vivo and has reduced
virulence in the mouse model .................................................................................... 50

  2.4 Discussion ................................................................................................................... 52

  2.5 References ............................................................................................................... 54
Chapter 3 Selective pressure to re-establish RNA-dependent RNA polymerase fidelity in the sabin vaccine strain ................................................................. 57

3.1 Introduction .................................................................................. 57
3.2 Methods ....................................................................................... 60
  3.2.1 Materials ................................................................................ 60
  3.2.3 Plasmid construction .............................................................. 61
  3.2.3 Overexpression and protein purification .................................. 61
  3.2.4 Kinetic assays ......................................................................... 61
  3.2.5 Determination of kinetic constants (K_{d,app} and k_{pol}) for nucleotide incorporation catalyzed by RdRp ....................... 62
  3.2.6 NMR sample preparation and spectroscopy .............................. 62
3.3 Results .......................................................................................... 63
  3.3.1 The three single variants D53N, Y73H, and K250E RdRps, and Sabin RdRp do not change assembly and stability of enzyme complexes .......... 63
  3.3.2 D53N, Y73H, and K250E RdRps have different sugar and nucleobase fidelities.............................................. 66
  3.3.3 Sabin 1 RdRp has lower sugar selectivity but similar nucleobase selectivity as WT enzyme ...................................................... 69
  3.3.4 Cooperative interactions among the D53N, Y73H, K250E, and T3621 substitutions are small ............................................. 72
  3.3.5 Other conformational changes compensate for the changes of motif D to regulate the fidelity of Sabin RdRp ......................... 75
  3.3.6 The Sabin RdRp is more susceptible to 2’ modified nucleotides ........ 77
3.4 Discussion ..................................................................................... 77
3.5 References ..................................................................................... 81

Chapter 4 Triphosphate Re-orientation of the incoming nucleotide as a fidelity checkpoint in the viral RNA-dependent RNA polymerase ........................................ 84

4.1 Introduction .................................................................................. 84
4.2 Methods ....................................................................................... 88
  4.2.1 Materials ................................................................................. 88
  4.2.2 Plasmid construction .............................................................. 88
  4.2.3 Expression and purification of RdRp ........................................ 89
  4.2.4 Purification, 5’ - 32P end labeling and annealing of sym/sub ..... 89
  4.2.5 PV RdRp Assay ...................................................................... 89
  4.2.6 Stopped flow experiment ......................................................... 89
  4.2.7 Benchtop assay for slow nucleotide incorporation ............... 90
  4.2.8 Fluorescence polarization experiment .................................... 90
  4.2.9 Product analysis and data analysis .......................................... 91
  4.2.10 Sample preparation for NMR spectroscopy ......................... 91
  4.2.11 1H - 13C HSQC NMR experiment ....................................... 92
  4.2.12 31P NMR experiment ........................................................... 92
4.3 Results .......................................................................................... 92
  4.3.1 The RNA binding affinity and stability of the RdRp-RNA binary complexes are slightly different between K167R and R174K RdRps ........ 92
  4.3.2 The single Arg substitution at Lys167 results in low fidelity, but the single Lys substitution at Arg174 results in high fidelity ......................... 97
4.3.3 The pre-chemistry conformational change in both K167R and R174K variants significantly contributes to changes in fidelity. ..........................100
4.3.4 The pre-chemistry conformational change in the R174K variant is significantly damaged due to an inability to form a stable “closed” state. 103
4.3.5 Motif F and motif D are independent of each other during the pre-chemistry conformational change. ..............................................106
4.3.6 31P NMR monitors conformations of the triphosphate of the incoming nucleotide in RdRp-RNA-nucleotide ternary complexes ..........................109
4.3.7 Triphosphate orientation in the ternary complexes is a fidelity checkpoint. 112
4.3.8 Substitutions interacting with α, β, or γ phosphate contribute to fidelity in different ways.................................................................114
4.3.9 The function of motif F is not necessarily coordinated with that of motif D. 115
4.4 Discussion ..................................................................................117
4.5 References ..................................................................................122

Chapter 5 A Novel antiviral inhibitor targeting the rna template channel of coxsackievirus

b3 124

5.1 Introduction .................................................................................124
5.2 Methods ....................................................................................128
  5.2.1 Materials ..............................................................................128
  5.2.2 Purification, 5′-32P end labeling, and annealing of sym/subU RNA 128
  5.2.3 CVB3 RdRp assay .................................................................128
  5.2.4 Homopolymeric template elongation assay ..........................129
  5.2.5 Template switching assay .....................................................130
  5.2.6 Fluorescence polarization experiments .................................130
5.3 Results .......................................................................................130
  5.3.1 6-FAM-sym/subU RNA allows nucleotide incorporation by CVB3 WT RdRp. ........................................................................130
  5.3.2 GPC-N114 contributes to fluorescence and prevents quantitative inhibition studies. .................................................................133
  5.3.3 GPC-N114 competes against template binding in the homopolymeric template assay. .................................................................138
  5.3.4 GPC-N114 has no inhibitory effects on the binding of sym/subU RNA to CVB3 WT RdRp. ..................................................................143
  5.3.5 GPC-N114 has no effects on template switching. .......................145
5.4 Discussion ..................................................................................147
5.5 References ..................................................................................149

Chapter 6 Conclusions and future directions........................................153

6.1 Conclusions ...............................................................................153
6.2 Future directions ........................................................................156
  6.2.1 Fidelity opens up a new way to design live attenuated vaccine strains 156
  6.2.2 Engineering motif F to generate RdRp variants suitable for live attenuated vaccine .......................................................................161
  6.2.3 Elucidating the roles of motifs E and G on RdRp function ............................164
  6.2.4 Other strategies to change the genetic composition of diverse viral populations ........................................................................168
6.3 References ................................................................................................................................................................168
LIST OF FIGURES

Figure 1-1. The structure of viral genome in enteroviruses. The RNA contains one open reading frame that consists of genes encoding P1, P2, and P3 proteins. VPg and poly(A) tail are at the 5’ and 3’ end of the genome respectively. The 5’ untranslated region (5’ UTR) contains a cloverleaf structure and an internal ribosomal entry site (IRES). B. The polyprotein is then proteolytically processed into multiple precursors and fully mature viral proteins. ..........................6

Figure 1-2. The crystal structure of PV RdRp (PDB 1RA6). A. A ribbon representation of PV RdRp with “Fingers” (cyan), “Palm” (magenta), and “Thumb” (green) subdomains. B. Highly conserved functional motifs in PV RdRp include motifs A (green), B (hot pink), C (orange), D (cyan), E (blue), F (yellow) and G (red).................................................................13

Figure 1-3. Chemical mechanism of PV RdRp. Nucleotide incorporation is the consequence of formation of phosphodiester bond between the 3’ OH of primer nucleotide (blue) and the α phosphate of the incoming nucleotide (green). Aspartates from motifs A and C coordinate with metal ion B and A respectively. Metal A lowers the pKa of the 3’ OH of primer strand, and an unknown base (B) deprotonates the same hydroxyl group. The Lys359 from motif D acts as a general acid to protonate the pyrophosphate group (green) and facilitates the formation of the phosphodiester bond. This figure was adapted with permission from reference 88. Copyright 2012 by the Elsevier Ltd. .................................................................14

Figure 1-4. The pre-steady-state kinetic mechanism of single nucleotide incorporation by PV RdRp. When nucleotide binds to the enzyme-RNA binary complex (step 1), the incoming nucleotide rearranges into a catalytically competent conformation (step 2). Then, the phosphodiester bond forms between the primer nucleotide and the incoming nucleotide. Pyrophosphate (PPi) also forms (step 3). The competent state relaxes to a nonproductive state accompanied by RNA translocation (step 4). The PPi releases (step 5). The RdRp-RNA complex is then ready for the next incoming nucleotide. ........................................................................14

Figure 1-5. The conformational change of motif D can be monitored by NMR. A. The conformational change of motif D in the X ray crystal structure of norovirus RdRp complexes with nucleotide (green) and without nucleotide (orange). B. The conformational change of motif D (as sensed by the changes in chemical shift of Met354) of PV RdRp ternary complexes with cognate nucleotide (black) and noncognate nucleotide (red) in NMR. When cognate nucleotide binds, motif D comes to the “closed” state. When noncognate nucleotide binds, motif D stays at the “open” state. The picture was adapted with permission from reference 95. Copyright 2014 by the Elsevier B.V.................................................................14

Figure 2-1. The T362I substitution may impact the structural dynamics of motif D and other regions to affect RdRp catalysis and fidelity. A, B. Crystal structure of PV RdRp (PDB 1RA6). Colored motifs include motif A (green), B (hot pink), C (orange), D (cyan), E (blue), F (yellow) and G (red). Thr362, Lys359, and probe Met354 are shown in purple, cyan, and purple respectively. The Met probes that are sensitive to RNA or nucleotide binding are shown in gold. C. Kinetic schemes of single nucleotide incorporation for both Mg2+ and Mn2+ include 5 steps. From steps 1 to 5, those are nucleotide binding, pre-chemistry conformational change, chemical reaction, post-chemistry conformational change, and release of the pyrophosphate group. For cognate nucleotide insertion, step 2 (red) is rate limiting only in Mg2+, and step 3 (light blue) is rate limiting in both Mg2+ and Mn2+. .................................................................36
Figure 2-2. Formation and stability of RdRp-RNA complexes are not affected by the T362I substitution in motif D. Reactions contained 50 mM HEPES 7.5, 10 mM BME, 5 mM MgCl$_2$ and 60 µM ZnCl$_2$. A. Sequence of both sym/subU and sym/subUA RNAs. B, C. In the assembly assay, ATP and RNA were firstly incubated for 5 min followed by addition of RdRp to initiate the reaction. Reactions were quenched by addition of 25 mM EDTA. Final concentrations of RdRp, RNA, and ATP were 1, 1, and 500 µM respectively. The data of formed WT RdRp-RNA complexes (empty circle) and T362I RdRp-RNA complexes (solid circle) is plotted as a function of time. C, D. In dissociation assay, RdRp-RNA binary complexes were formed in 90 s followed by addition of trap (100 µM). At each time point, ATP was allowed to react for 30 s followed by addition of 25 mM EDTA to quench the reaction. Final concentrations of RdRp, RNA, and ATP were 1, 0.1, and 500 µM respectively. Percentage of the remaining binary complex is plotted as a function of time for both WT (solid circle) and T362I (empty circle) RdRps. The solid lines represent the fit of the data into a single exponential function. The $k_{\text{dis}}$ of WT enzyme for sym/subU and sym/subUA RNAs are $4.00 \pm 0.09 \times 10^{-4}$ s$^{-1}$ and $1.70 \pm 0.07 \times 10^{-5}$ s$^{-1}$ respectively. The $k_{\text{dis}}$ of T362I RdRp for sym/subU and sym/subUA RNAs are $3.90 \pm 0.02 \times 10^{-4}$ s$^{-1}$ and $2.70 \pm 0.03 \times 10^{-5}$ s$^{-1}$ respectively. The figure was reprinted from reference 1. © The American Society for Biochemistry and Molecular Biology.

Figure 2-3. The T362I substitution lowers fidelity in the presence of Mg$^{2+}$. Reactions contained 50 mM HEPES 7.5, 10 mM BME, 5 mM MgCl$_2$, and 60 µM ZnCl$_2$. RdRp and RNA were firstly incubated for 2 min to form RdRp-RNA binary complexes. Reactions were initiated by the addition of an equal volume of nucleotide at various concentrations. Changes in fluorescence were recalculated as product formed in time. Final concentrations of RdRp and RNA were 0.5 and 0.5 µM respectively. A. Reaction scheme of single nucleotide incorporation in the presence of Mg$^{2+}$. B. For cognate ATP incorporation, T362I RdRp (empty circle) is as efficient as WT RdRp (solid circle). C, D. For noncognate 2’dATP and GTP incorporations, T362I RdRp (empty circle) is more efficient than WT RdRp (solid circle). Solid lines represent the fit of the data into hyperbola function. The obtained $k_{\text{pol}}$ and $K_{\text{d,app}}$ for each nucleotide incorporation are shown in table 2-1. The figure was reprinted from reference 1. © The American Society for Biochemistry and Molecular Biology.

Figure 2-4. The T362I substitution lowers sugar and nucleobase selections by making the pre-chemistry conformational change more efficient. Reactions contained 50 mM HEPES 7.5, 10 mM BME, 5 mM MnCl$_2$, and 60 µM ZnCl$_2$. RdRp and RNA were firstly incubated for 2 min to form RdRp-RNA binary complexes. Reactions were initiated by the addition of an equal volume of nucleotide at various concentrations. Changes in fluorescence were recalculated as product formed instantly. Final concentrations of RdRp and RNA were 0.5 and 0.5 µM respectively. A, B. For cognate ATP and noncognate 2’dATP incorporations, T362I RdRp (empty circle) is as efficient as WT RdRp (solid circle). C. For noncognate GTP incorporation, T362I RdRp (empty circle) is more efficient than WT RdRp (solid circle). Solid lines represent the fit of the data into hyperbola function. The obtained $k_{\text{pol}}$ and $K_{\text{d,app}}$ for each nucleotide incorporation are shown in table 2-1. The figure was reprinted from reference 1. © The American Society for Biochemistry and Molecular Biology.

Figure 2-5. The T362I substitution alters the motif D equilibrium between the open and closed states as revealed by NMR spectra of [Methyl$^{13}$C]Met-labeled RdRp. A, B, and C. WT RdRp clearly distinguishes cognate UTP (black) from noncognate 2’dUTP (red) and ATP (blue) by binding only UTP into the “closed” state and remaining the “open” state of 2’dUTP or ATP. D, E, and F. Motif D in T362I RdRp completely or partially shifts to the “closed” state when UTP...
(black) or 2’dUTP (red) is bound respectively. However, T362I RdRp is unable to significantly reverse the motif D equilibrium when ATP binds (blue). The figure was reprinted from reference 1. © The American Society for Biochemistry and Molecular Biology.

Figure 2-6. The T362I substitution in motif D leads to structural and dynamics changes consistent with changes to RdRp function as revealed by MD simulations. A. The average structure of WT RdRp from the previous study (grey) and T362I RdRp (blue). The structures are overlapped at the Cα atoms. Four substitution sites (Asp53, Tyr73, Lys250, and Thr362) in RdRp of Sabin I vaccine strain are shown as sticks. B. Zoom-in view of the surrounding regions at Thr362 in A. Conformational changes of selected residues (shown as sticks) near Thr362 or Ile362 are obvious especially Glu364. C. Zoom-in view of the active site distant from the substitution site T362I in A. Conformations of many conserved, functional residues (shown as sticks) in motif A (Asp238), B (Thr293), C (Asp328 and Asp329), and D (Lys359) are changed by the single substitution of T362I. D. The per residue root mean square deviation (per RMSD) calculated from the last 10 ns of WT (grey) and T362I (black) trajectories. The per RMSD corresponding to the four substitution sites in the Sabin I vaccine strain are shown in dark red diamond in WT RdRp and dark red triangle in T362I RdRp. The T362I substitution constrains the overall dynamics of RdRp and is also able to communicate with other Sabin sites via changing their dynamics. The figure was reprinted from reference 1. © The American Society for Biochemistry and Molecular Biology.

Figure 2-7. PV encoding the T362I substitution in motif D expresses a mutator phenotype in cell culture and is attenuated in vivo. A. In the ribavirin sensitivity assay, HeLa cells were pretreated with ribavirin at various concentrations and were infected with WT, G64S, and T362I PVs at 50 plaque forming unit. Plaque numbers relative to control (without ribavirin) are plotted as a function of ribavirin concentration. The solid lines represent the fit of the data to a sigmoidal dose response equation (four parameter logistic model). The IC_{50} values of WT, G64S, and T362I PVs were 0.67 ± 0.06, 2.15 ± 0.75, and 0.3 ± 0.02 mM respectively. B, C. T362I PV was attenuated in the cPVR (poliovirus receptor) mouse model. cPVR mice (4 to 6 weeks old) were infected via intraperitoneal route with serum-free media containing WT and T362I PVs at the indicated genome copy and observed for 10 days. Mice that developed symptoms of double limb paralysis or paralysis that compromised the ability to get food and water were euthanized. The figure was reprinted from reference 1. © The American Society for Biochemistry and Molecular Biology.

Figure 3-1. The structure of PV RdRp indicating the four Sabin substitution sites (PDB 1RA6). The front and back views are indicated in A and B respectively. PV RdRp contains “palm”, “thumb”, and “fingers” subdomains and seven highly conserved motifs including motif A (green), B (hot pink), C (orange), D (cyan), E (blue), F (yellow), and G (red). The Sabin RdRp contains four substitutions including D53N, Y73H, K250E, and T362I (whose substitution sites are colored in purple) compared to WT RdRp. The general acid Lys359 is colored in cyan. The loop, where both D53N and Y73H substitutions occur, is colored in gold. This loop also contains Gly64. The substitution of Gly64 to Ser leads to an increase in enzyme fidelity. Pictures are created using Chimera.

Figure 3-2. Sabin substitutions do not affect the assembly and the stability of RdRp-RNA binary complexes. Reactions contained 50 mM HEPES 7.5, 10 mM BME, 5 mM MgCl₂, and 60 µM ZnCl₂. A. Sequence of sym/subU RNA. B. Reaction scheme of the assembly assay. RNA and ATP were preincubated for 5 min. Then, reactions were initiated by addition of RdRp and were quenched by addition of EDTA to a final concentration of 25 mM at different time points. Final
concentrations of RdRp, sym/subU RNA, and ATP were 1, 1, and 500 µM respectively. C. Overlapping of the formation of the assembled complexes for WT, Sabin, D53N, Y73H, K250E, and T362I RdRps suggested that these Sabin substitutions do not affect enzyme assembly. D. Reaction scheme of the dissociation assay. RdRp-RNA binary complexes were incubated for 90s followed by addition of trap. At different time points, ATP was added and was allowed to incorporate for 30s after which reactions were quenched by the addition of EDTA to a final concentration of 25 mM. Final concentrations of RdRp, sym/subU, and ATP were 1, 0.1, and 500 µM respectively. E. Overlapping of the dissociation of the enzyme complexes for WT, Sabin, D53N, Y73H, and T362I RdRps. The lines represent fits of the data to a single exponential function. The calculated dissociation rates ($k_{dis}$) are in panel F. Data for T362I RdRp was obtained from previous studies (see Chapter 2).

Figure 3-3. Sabin 1 RdRp has the lowest sugar fidelity but has nucleobase fidelity similar to WT level. Reactions contained 50 mM HEPES 7.5, 10 mM BME, 5 mM MgCl$_2$, and 60 µM ZnCl$_2$. Final concentrations of RdRp and sym/subU RNA were 0.5 and 0.5 µM respectively. A. Reaction scheme of a stopped flow experiment. RdRp and sym/subU RNA were preincubated for 2min followed by addition of equal volume of nucleotide to initiate reactions. Product formation was monitored by changes in fluorescence. B, C. Comparison of the cognate ATP (B) and noncognate 2’dATP (C) incorporations among WT, T362I, and Sabin 1 RdRps. The lines represent fits of the data to a hyperbola function. The obtained $k_{pol}$ and $K_{d,app}$ are listed in Table 3-1. D. Reaction scheme of a benchtop assay. RdRp and sym/subU RNA were preincubated for 5min followed by addition of equal volume of GTP to initiate reactions. Reactions were quenched by addition of EDTA to a final concentration of 25 mM. E. Comparison of the noncognate GTP incorporations among WT, T362I, and Sabin 1 RdRps. The lines represent fits of the data to a hyperbola function. The obtained kinetic parameters are listed in Table 3-1.

Figure 3-4. There is wide range of cooperativity among D53N, Y73H, K250E, and T362I substitutions in Sabin 1 RdRp and Sabin derived RdRps. Thermodynamic effect $\Delta \Delta G$ is calculated by the following equation $\Delta \Delta G_x = RT \ln\left(\frac{(k_{pol}/K_{d,app})_x}{(k_{pol}/K_{d,app})_{WT}}\right)$, where x represents a particular variant. A. The comparison of $\Delta \Delta G$ between Sabin 1 variant and sum of D53N, Y73H, K250E, and T362I single variants (whose substitutions together consist of the Sabin variant), between D53N/Y73H/T362I and sum of D53N, Y73H, and T362I single variants, and between D53N/T362I and sum of D53N and T362I single variants for ATP, 2’dATP, and GTP incorporations. Positive cooperativity represents a greater $\Delta \Delta G$ of a multiple variant than that of the sum of the corresponding single variants. Thermodynamic cycle is calculated for ATP (B), 2’dATP (C), and GTP (D) incorporations. On the left side is the thermodynamic cycle of Y73H, D53N/Y73H, and D53N/Y73H/K250E variants. On the right side is the thermodynamic cycle of K250E, D53N/Y73H/K250E, and Sabin 1 variants. The unit is cal/mol. E, F. Comparison of the sugar (E) and nucleobase (F) fidelities among WT, T362I, D53N/Y73H, D53N/Y73H/K250E, and Sabin 1 variants.

Figure 3-5. Motif D cooperates with other structural elements to account for RdRp fidelity of Sabin derived RdRp variants. Comparison of the [Methyl$^{13}$C]Met $^1$H$^{13}$C HSQC spectra between the cognate RdRp-RNA-UTP and noncognate RdRp-RNA-2’dUTP ternary complexes for WT (A), T362I (B), D53N/Y73H/T362I (C), and Sabin 1 RdRps. WT enzyme strictly distinguishes cognate and noncognate nucleotides by assigning motif D to the “closed” and “open” states respectively. T362I variant promiscuously recognizes noncognate nucleotide by assigning motif D partially to the “closed” state, which leads to a lower fidelity of this variant. Neither D53N/Y73H/T362I triple variant nor Sabin 1 RdRp assigns motif D to the “closed” state to the
same extent as the T362I variant for noncognate nucleotide, which suggests involvement of other structural elements in determining their fidelities. (D) Spectra are collected at 293K with 250 µM RdRp, 1 mM sym/subU RNA, and 4 mM UTP or 8 mM 2'dUTP.

Figure 3-6. The D53N and Y73H substitutions are able to affect the conformation of motif D (red) via a hydrogen bond network including motif A (purple), Gly64 (green), Asn65 (green), and the N terminus (orange). The Sabin substitution sites Asp53 and Tyr73 are colored in grey, and the structures that these two substitutions site are located are colored in yellow. The general acid Lys359 (cyan) and another substitution site Thr362 (grey) is located on motif D. The figure was edited using Chimera (PDB 3OL7). 

Figure 4-1. A. Reaction scheme for single nucleotide incorporation. Steps 1 to 5 are nucleotide binding, prechemistry conformational change, phosphoryl transfer, postchemistry conformational change, and pyrophosphate release respectively. Among the five steps, step 2 is an important fidelity checkpoint. B. Crystal structure of PV RdRp (PDB 1RA6). Motifs D and F are colored in cyan and hot pink respectively. The general acid Lys359, Lys167, and Arg174 are represented as spheres. C. Crystal structure of PV RdRp. D. Lys167 and Arg174 interact with the γ and α phosphate moiety of the incoming nucleotide respectively. Motif D is colored in hot pink, and the backbone and nucleobases of RNA are colored in grey and blue respectively. For Lys167, Arg174, and incoming NTP, heteroatoms are colored by blue (N), tan (C), orange (P), and red (O) respectively.

Figure 4-2. The K167R substitution does not affect assembly and disassociation of enzyme complexes. Reactions contained 50 mM HEPES 7.5, 10 mM BME, 5 mM MgCl₂, and 60 µM ZnCl₂. A. Sequence of sym/subU RNA. B. Reaction scheme for the assembly assay. RNA and ATP were preincubated for 5 min. Then, reactions were initiated by addition of RdRp and were quenched by addition of EDTA to a final concentration of 25 mM at different time points. Final concentrations of RdRp, sym/subU RNA, and ATP were 1, 1, and 500 µM respectively. C. Comparison of the assembled enzyme complexes of K167R variant (red circle) to those of WT enzyme (blue circle). D. Reaction scheme for the dissociation assay. RdRp-RNA binary complexes were incubated for 90s followed by addition of trap. At different time points, ATP was added and was allowed to incorporate in 30s after which reactions were quenched by addition of EDTA to a final concentration of 25 mM. Final concentrations of RdRp, sym/subU, and ATP concentrations were 1, 0.1, and 500 µM respectively. E. The decay of formed enzyme complexes of the K167R variant, compared to WT RdRp. The solid lines represent the fit of the data to a single exponential curve, which gives \( k_{\text{dis}} \) values of 0.00040 ± 0.00003 s⁻¹ and 0.00035 ± 0.00001 s⁻¹ for WT and K167R RdRps respectively.

Figure 4-3. The assembled complexes of R174K variant are slightly different from those of WT RdRp. Reactions were conducted in the same way as described in Figure 4-2. A, B, C and D. Comparison of the assembled complexes of R174K variants (red circle) to those of WT RdRp (blue circle). Reactions contained 500 µM ATP, either 1 or 2 µM sym/subU RNA for WT and R174K RdRps respectively, and either 0.25 (A), 0.5 (B), 1 (C), or 5 (D) µM RdRp.

Figure 4-4. The dissociation of enzyme complexes of R174K variant is slightly faster than that of WT RdRp. Reaction and data analysis was conducted in the same way as described in Figure 4-2 except that the concentration of sym/subU RNA was 0.2 µM. A. Comparison of the dissociation of enzyme complexes between R174K and WT RdRps. The calculated \( k_{\text{dis}} \) for R174K RdRp is 0.0011 ± 0.00008 s⁻¹.
Figure 4-5. Comparison of the AMP incorporation in AMPCPP assay for WT, K167R and R174K RdRps suggests that the R174K substitution significantly affects the prechemistry conformational change. A. Minimal kinetic mechanism of the dissociation of AMPCPP ternary complexes monitored by incorporation of AMP. Reactions contained 50 mM HEPES 7.5, 10 mM BME, 5 mM MnCl$_2$, and 60 µM ZnCl$_2$. Enzyme was preincubated with sym/subU RNA to form binary complexes for 90s. Then 200 µM AMPCPP or buffer was added and incubated for extra 5 min to form RdRp-sym/subU-AMPCPP ternary complexes. Reactions were initiated by addition of an excess amount of ATP. AMP incorporation was then monitored by fluorescent changes using a stopped flow apparatus. Final concentrations of ATP, sym/subU RNA, and RdRp were 1000, 1 or 2 µM (for WT and K167R RdRps and R174K RdRp respectively), and 1 µM respectively. B, C, and D. The comparison of AMP incorporation in the presence (blue) and absence (red) of AMPCPP for WT (B), K167R (C), and R174K RdRp (D). The solid lines represent the fit of the data into a single exponential curve and give the $k_{obs}$ values that are listed in Table 4-3.

Figure 4-6. Motifs D and F are functionally independent. The units are kcal/mol. Thermodynamic effect $\Delta \Delta G$ is calculated by the following equation $\Delta \Delta G_{x1} = RT \ln \left( \frac{k_{pol}/K_{d,app,x1}}{k_{pol}/K_{d,app,x2}} \right)$, where $x1$ represents the variant that an arrow points to and $x2$ represents the variant that the same arrow begins. The thermodynamic cycle is calculated for AMP incorporation by WT, K359L, R174K, and R174KK359L RdRps.

Figure 4-7. NMR scans at every step of sample preparation allows assignment of the protein bound triphosphate. A. Sample preparation for $^{31}$P NMR experiments. 3'dAMP is incorporated into sym/subU RNA by RdRp. The blocked 3'H RNA allows addition of the second nucleotide UTP to form the RdRp-sym/subU-UTP ternary complex. A desalting column is used to remove unincorporated free nucleotide. B. The triphosphate of bound and unbound nucleotide can be detected by $^{31}$P NMR. Not shown are the peaks for the phosphate-backbone of the RNA (-1.2 ppm). The spectra for the RdRp complexes (RdRp-sym/subU+3’dATP, RdRp(3’H), RdRp(3’H)+UTP and RdRp(3’H)-UTP (desalted) shown in orange, brown, purple, and green respectively) are compared to the spectra for free UTP in the presence (blue) and absence (red) of Mg$^{2+}$.

Figure 4-8. $^{31}$P NMR experiments indicate that the triphosphate conformations of the incoming nucleotide for K167R, R174K, and K359R ternary complexes are different from those of WT enzyme. A, B, and C. Comparison of the conformations of the triphosphate among variants and WT enzyme for cognate UTP (A), 2’ddeoxynucleotide (B), and noncognate nucleobase (C) ternary complexes respectively. Spectra of K359R, R174K, K167R, and WT ternary complexes are shown in purple, green, red, and blue respectively. Spectra were collected at 303 K with 250 µM RdRp, 500 µM RNA, and 4 mM UTP, 8 mM 2’dUTP or 12 -18 mM CTP.

Figure 4-9. Functions and conformations of motifs D and F are independent. A, B, and C. Comparisons of [Methyl-$^{13}$C]Met NMR spectra for K167R ternary complexes (red) with cognate nucleotide UTP (A), 2’dUTP (B), and noncognate nucleobase CTP (C) with those for WT ternary complexes (black). The K167R substitution does not perturb conformation of motif D for both cognate and noncognate nucleotides. D, E. Comparison of the [Methyl-$^{13}$C]Met NMR spectra for R174K ternary complexes (green) with cognate nucleotide UTP (D) and noncognate nucleotide (E) with those for WT ternary complexes (black). The R174K variant cannot form the same conformation of motif D as that of WT enzyme. F. Overlapping of cognate (black) and noncognate (green) ternary complexes for R174K variant indicates that the R174K substitution
significantly impairs the prechemistry conformational equilibrium that the “closed” state cannot form even with cognate nucleotide.......................................................... 116

Figure 4-10. The “occluded” (A) and “competent” states (B) of the active site in PV RdRp. Residues from highly conserved motifs A (green), B (tan), D (blue), and F (red) participate in the regulation of the equilibrium between “occluded” and “competent” states. The figure was reprinted from reference 37. © the American Society for Biochemistry and Molecular Biology. .......................................................................................................................... 121

Figure 5-1. 6-FAM-sym/subU RNA can be utilized for nucleotide incorporation by CVB3 WT RdRp. Reactions contained 50 mM Tris 7.0, 10 mM KCl, and 0.8 mM MgCl₂. Either radioactive sym/subU RNA (A) or fluorescent 6-FAM-sym/subU RNA (B) was incubated with CVB3 RdRp for 90 s at which time, reactions were initiated by the addition of ATP. The reaction was quenched by addition of EDTA to a final concentration of 25 mM. Final concentrations of RNA, RdRp, and ATP were 1, 1, and 500 µM respectively. PV RdRp was used as a control. CVB3 WT RdRps stored in 50% glycerol (lane CVB3) or without glycerol (lane CVB3NG) were also compared for enzyme activity.............................................................. 132

Figure 5-2. The dissociation constant of 6-FAM-sym/subU RNA to CVB3 WT RdRp is at the micromolar level. A. Kinetic measurements of 4 nM 6-FAM-sym/subU RNA were performed in a reaction buffer containing 50 mM Tris 7.0, 10 mM KCl, and 0.8 mM MgCl₂. B. 6-FAM-sym/subU RNA binding to CVB3 WT RdRp was measured under the same condition as before (see Figure 5-1) except that the salt concentration was fixed to 10 mM KCl and 15 mM NaCl. The concentration of CVB3 WT RdRp used to titrate RNA varied from 10 to 1700 nM. Measured mini-polarization is plotted as a function of CVB3 WT RdRp concentration. The solid line represents the fit of the data to a hyperbola function. The Kᵤ value equals 1.3 x 10⁻³ ± 0.3 µM. Experiments were conducted in triplicate.......................................................... 134

Figure 5-3. GPC-N114 inhibits UTP incorporation into the homopolymeric template. Reactions contained 50 mM Tris 7.0, 10 mM KCl, and 0.8 mM MgCl₂. Final concentrations of CVB3 WT RdRp, poly(rA)/dT15, and UTP* were 20 nM, 350 nM/2 µM, and 10 µM (2 µCi) respectively. Concentration of GPC-N114 ranged from 0 to 80 µM. CVB3 WT RdRp was firstly incubated with RNA and GPC-N114 for 2 min, at which time UTP was added to initiate the reaction. Reactions were quenched by the addition of 50 mM EDTA. A. Reaction scheme of UTP incorporation into homopolymeric template. B. GPC-N114 inhibited UTP incorporation in a concentration dependent manner. The solid lines represent the fit of the data to linear curves. C. The initial velocity obtain from B is plotted as a function of concentration of GPC-N114........... 140

Figure 5-4. Results of order-of-addition reactions indicated that GPC-N114 inhibits the reaction more efficiently when it was incubated with CVB3 WT RdRp before the addition of RNA. Reactions contained 50 mM Tris 7.0, 10 mM KCl, and 0.8 mM MgCl₂. Final concentrations of CVB3 WT RdRp, poly(rA)/dT15, and UTP* were 20 nM, 14 nM/80 nM, and 10 µM (2 µCi) respectively. The concentration of GPC-N114 ranged from 0 to 100 µM. CVB3 WT RdRp was either incubated with GPC-N114 before the addition of RNA (A) or was incubated with RNA prior to the addition of GPC-N114 (B). Reactions were initiated by the addition of UTP and were quenched by the addition of 50 mM EDTA. C. Percentage of activity relative to control is plotted as a function of the concentration of GPC-N114.......................................................... 142
Figure 5-5. GPC-N114 does not show competition against sym/subU RNA. Reactions contained 50 mM Tris 7.0, 10 mM KCl, and 0.8 mM MgCl₂. Final concentrations of CVB3 WT RdRp, sym/subU RNA, and ATP/NTPs were 1, 1, and 500 µM respectively. The concentration of GPC-N114 ranged from 0 to 100 µM. Reactions were conducted by adding CVB3 WT RdRp to initiate reactions (A, B), or by incubating CVB3 WT RdRp with sym/subU RNA and GPC-N114 followed by initiation with ATP addition (C, D), or by incubating CVB3 WT RdRp with NTPs and GPC-N114 followed by initiation with sym/subU RNA addition (E, F) or by incubating CVB3 WT RdRp with GPC-N114 followed by initiation with addition of sym/subU RNA and NTPs (G, H).

Figure 5-6. GPC-N114 does not show inhibition of template switching. A. Reaction scheme of template switching. Reactions contained 50 mM Tris 7.0, 10 mM KCl, and 0.8 mM MgCl₂. CVB3 WT RdRp (5 µM) and dT15/rA30 (10 µM) were incubated in the presence of UTP* (500 µM, 2 µCi) for 3 min. The formed complexes were diluted by 10 fold in reaction buffer containing 500 µM UTP, 10 µM heparin, and with (solid square) or without (circle) 50 µM GPC-N114 in the absence (B) or presence (C) of 100 µM GPC-N114.

Figure 6-1. Alignment of PV RdRp sequence with the sequences from CVB3, EV71, HRV14, and FMDV. Residues that are highly conserved across all the RdRps are highlighted with red boxes. Residues that are similar are indicated by pink boxes. Highly conserved motifs A to G are colored in cyan, yellow, purple, green, light blue, light purple and dark blue respectively.

Figure 6-2. Structural superposition of PV RdRp with RdRps from CVB3, EV71, HRV14, and FMDV. PV, CVB3, EV71, HRV and FMDV are colored in light blue, green, tan, orange, and light purple respectively (PDB 3OL6, 3DDK, 3N6L, 1XR5, and 1U09).

Figure 6-3. The structure of motif F and the incoming nucleotide in PV RdRp. A and B. Part of motif F is shown in grey with Arg174, Lys167, and Arg163 colored in yellow and Lys159, Glu161, Ser164, and Ile176 colored in magenta. Glu39 and Ala41 are located on the loop near the N terminus and are colored in green. The acceptor base to the incoming nucleotide is a square and colored in blue. The incoming nucleotide is colored by heteroatom (carbon (tan), oxygen (red), and phosphor (orange)). Cyan lines represent possible hydrogen bond interactions. The crystal structure is based on the PV RdRp ternary complex (PDB 3OLB).

Figure 6-4. Interactions of motif E with the 3' end of primer strand (PDB 3olB). Motif E is colored in light blue with Phe373 and Lys375 colored in dark green. Pro393 and Glu396 are colored in yellow. The α helix that contains Pro393 and Glu396 and motif C are both colored in grey. Asp238 on motif C is colored in magenta. The base of the -2 nucleotide in the primer strand is colored in blue. The backbone of the -1 and -2 nucleotides in the primer strand is colored by heteroatom (carbon (tan), oxygen (red), and phosphor (orange)). The cyan lines represent all the contacts between atoms.

Figure 6-5. Interactions of motif G with the backbone of template strand (PDB 3olB). Motif G is colored in blue with Asp111, Thr114, and Ser115 that are colored in yellow. Lys127 is colored in magenta. From top to bottom, the backbone of the +1 (blue), +2 (blue), and -1 nucleotides in the template strand is colored by heteroatom (carbon (tan), oxygen (red), and phosphor (orange)). Water molecules are colored in red. The cyan lines represent all the possible hydrogen bonds.
LIST OF TABLES

Table 2-1. Comparison of nucleotide incorporation for WT and T362I RdRps reveals that the T362I substitution decreases polymerase fidelity. The table was reprinted from reference 1. © The American Society for Biochemistry and Molecular Biology........................................43

Table 3-1. Comparison of the nucleotide incorporations for WT, Sabin, and Sabin derived RdRps reveals that they all have slightly different catalytic parameters and/or fidelity..........................68

Table 4-1. Nucleotide incorporations in the presence of Mg$^{2+}$ ........................................................99

Table 4-2. Nucleotide incorporations in the presence of Mn$^{2+}$ .........................................................102

Table 4-3. Comparison of the AMP incorporation in the presence and absence of AMPCPP for WT, K167R, and R174K RdRps respectively.................................................................105

Table 5-1. Comparison of the dissociation constants of 6-FAM-sym/subU RNA binding to CVB3, PV, and FMDV WT RdRps revealed that 6-FAM-sym/subU RNA has similar binding affinity for all RdRps. The reaction buffer was the same as before (see Figure 5-1) except that the salt concentration was fixed to 10 mM KCl and 15 mM NaCl.................................................................135

Table 5-2. Comparison of the effects of GPC-N114 on fluorescent intensity and mini-polarization in three different reaction conditions. Reaction buffer was the same as before (see Figure 5-1) except that salt concentration was fixed to 10 mM KCl and 15 mM NaCl. Both intensity and mini-polarization (mP) were increased with increasing concentration of GPC-N114 in the absence of RNA, or in the presence of RNA, or both RNA and CVB3 WT RdRp. The GPC-N114 concentration ranged from 0.1 to 100 µM. Final concentrations of 6-FAM-sym/subU and CVB3 WT RdRp were fixed to 4 nM and 1.3 µM respectively.........................................................137

Table 5-3. Comparison of the GPC-N114 inhibition of UTP incorporation between UTP initiation and UTP incubation suggested that UTP disturbs GPC-N114 binding to CVB3 WT RdRp although GPC-N114 does not occupy the nucleotide binding pocket as shown in the crystal structures of CVB3 WT RdRp-GPC-N114 complexes.........................................................141
**LIST OF ABBREVIATIONS**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACE</td>
<td>bis(2-acetoxyethoxy)methyl</td>
</tr>
<tr>
<td>AMPCPP</td>
<td>α, β-Methyleneadenosine 5’-triphosphate</td>
</tr>
<tr>
<td>bp</td>
<td>Base pairs</td>
</tr>
<tr>
<td>BME</td>
<td>β-Mercaptoethanol</td>
</tr>
<tr>
<td>cPVR</td>
<td>Poliovirus receptor</td>
</tr>
<tr>
<td>CVB3</td>
<td>Coxsackievirus B3</td>
</tr>
<tr>
<td>DdDp</td>
<td>DNA-dependent DNA polymerases</td>
</tr>
<tr>
<td>DdRp</td>
<td>DNA-dependent RNA polymerases</td>
</tr>
<tr>
<td>ds</td>
<td>Double stranded</td>
</tr>
<tr>
<td>EMCV</td>
<td>Encephalomyocarditis virus</td>
</tr>
<tr>
<td>EV71</td>
<td>Enterovirus 71</td>
</tr>
<tr>
<td>6-FAM</td>
<td>6-Carboxyfluorescein</td>
</tr>
<tr>
<td>FDA</td>
<td>U S Food and Drug Administration</td>
</tr>
<tr>
<td>FMDV</td>
<td>Foot-and-mouth disease virus</td>
</tr>
<tr>
<td>FP</td>
<td>Fluorescence polarization</td>
</tr>
<tr>
<td>GPC-N114</td>
<td>2,2’-[(4-chloro-1,2-phenylene)bis(oxy)]bis(5-nitro-benzonitrile)</td>
</tr>
<tr>
<td>HAV</td>
<td>Hepatitis A virus</td>
</tr>
<tr>
<td>HFMD</td>
<td>Hand-foot-and-mouth diseases</td>
</tr>
<tr>
<td>HIV</td>
<td>Human immunodeficiency viruses</td>
</tr>
<tr>
<td>HRV14</td>
<td>Human rhinovirus serotype 14</td>
</tr>
<tr>
<td>HRV16</td>
<td>Human rhinovirus serotype 16</td>
</tr>
<tr>
<td>HSQC</td>
<td>Heteronuclear single quantum coherence</td>
</tr>
<tr>
<td>IRES</td>
<td>Internal ribosomal entry site</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>MD</td>
<td>Molecular dynamics</td>
</tr>
<tr>
<td>MERS</td>
<td>Middle East respiratory syndrome</td>
</tr>
<tr>
<td>mP</td>
<td>mini-Polarization</td>
</tr>
<tr>
<td>NMR</td>
<td>Nuclear magnetic resonance</td>
</tr>
<tr>
<td>NNI</td>
<td>non-Nucleoside inhibitors</td>
</tr>
<tr>
<td>NP-40</td>
<td>Nonidet P-40</td>
</tr>
<tr>
<td>ORF</td>
<td>Open reading frame</td>
</tr>
<tr>
<td>PEI</td>
<td>Polyethyleneimine</td>
</tr>
<tr>
<td>per-RMSD</td>
<td>Per-residue mean square deviation</td>
</tr>
<tr>
<td>PMSF</td>
<td>Phenylmethylsulfonyl fluoride</td>
</tr>
<tr>
<td>pol v</td>
<td>Polymerase v</td>
</tr>
<tr>
<td>PV</td>
<td>Poliovirus</td>
</tr>
<tr>
<td>RdRp</td>
<td>RNA-dependent RNA polymerase</td>
</tr>
<tr>
<td>RMSD</td>
<td>Root-mean-square-deviation</td>
</tr>
<tr>
<td>RT</td>
<td>Reverse transcriptase</td>
</tr>
<tr>
<td>RTP</td>
<td>Ribavirin triphosphate</td>
</tr>
<tr>
<td>SARS</td>
<td>Severe acute respiratory syndrome</td>
</tr>
<tr>
<td>SDKIE</td>
<td>Solvent deuterium kinetic isotope effect</td>
</tr>
<tr>
<td>ss</td>
<td>Single stranded</td>
</tr>
<tr>
<td>Taq pol I</td>
<td>Thermus Aquaticus DNA polymerase I</td>
</tr>
<tr>
<td>TATase</td>
<td>Terminal adenyldyl transferase</td>
</tr>
<tr>
<td>T_{10}E_{1}</td>
<td>10 mM Tris and 1 mM EDTA</td>
</tr>
<tr>
<td>UTR</td>
<td>Untranslated region</td>
</tr>
<tr>
<td>VAPP</td>
<td>Vaccine associated paralytic poliomyelitis</td>
</tr>
<tr>
<td>VLP</td>
<td>Virus-like particles</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>-------------</td>
<td>---------------------------</td>
</tr>
<tr>
<td>VPg</td>
<td>Viral protein genome</td>
</tr>
<tr>
<td>WHO</td>
<td>World Health Organization</td>
</tr>
<tr>
<td>WT</td>
<td>Wild type</td>
</tr>
</tbody>
</table>
ACKNOWLEDGEMENTS

Firstly, I would like to express my sincere gratitude to my advisor Dr. David Boehr and Dr. Craig Cameron for the continuous support of my Ph.D. study and related research, for their patience, motivation, and knowledge. Their guidance helped me in all the time of research and writing of this dissertation. I could not have imagined having a better advisor and mentor for my Ph.D. study.

Besides my advisor, I would also like to thank the rest of my committee: Dr. John Golbeck and Dr. Scott Showalter, for their insightful comments and encouragement, but also for the hard question which incented me to broaden my horizon. Especially, Dr. Showalter opened the door of NMR for me and took me to navigate in the sea of knowledge.

My sincere thanks also goes to Dr. Jamie Arnold and Dr. Eric Smidansky who taught me how to study functions of enzyme using kinetic and thermodynamic methodologies. Especially, Dr. Jamie Arnold is always professional and answers my questions patiently.

I thank my fellow labmates for the stimulating discussions and for all the fun we have had in the last five years.

Most importantly, I would like to thank my family: my parents and my relatives who always support me unconditionally and love me forever.
CHAPTER 1
INTRODUCTION

1.1 Viruses

Viruses are significant infectious pathogens that can transmit among all types of living organisms from plants to insects to humans. Epidemic outbreaks caused by viruses often become serious public health issues. Some infectious diseases are able to rapidly spread within a country or between countries and kill thousands of people in just a few months. For example, the Ebola outbreak that is ongoing in West Africa has infected close to thirty thousand people and led to over 10,000 deaths from 2014 to 2015 according to the WHO (World Health Organization) report. Other viruses include influenza virus, severe acute respiratory syndrome (SARS) coronavirus, Middle East respiratory syndrome (MERS) coronavirus, coxsackievirus, poliovirus, foot-and-mouse disease virus, hepatitis C virus, dengue virus, and so on.

Viral infection is dependent on host cells. Without host cells, viruses are not able to replicate. Once entering the host cell, the virus immediately “highjacks” cellular metabolic tools to produce a large amount of viral products including viral proteins and genetic materials to produce thousands to millions of copies of itself. These new virus particles are then released to infect other healthy cells by lysing the cell, also termed as the lytic phase in the virus life cycle. The lytic phase is very fast. For example, poliovirus may take only about 6 hours to leave its original infected cells.

In the viral life cycle, genome replication is carried out by the virally encoded polymerase. This process is very error prone and may allow viruses to evolve into drug resistant mutants or immune escape mutants. Viruses can be divided into seven groups according to the
identity of the nucleic acid (DNA or RNA) strandedness (single stranded (ss) or double stranded (ds)), sense (positive (+) or negative (-)) and mode of synthesis (DNA polymerase, RNA polymerase, reverse transcriptase (RT)): I: dsDNA viruses (e.g. herpes virus and papillomavirus), II: ssDNA viruses (i.e. paroviruses), III: dsRNA viruses (e.g. reoviruses), IV: (+)ssRNA viruses (e.g. hepatitis C virus and foot-and-mouth disease virus ), V: (-)ssRNA viruses (e.g. influenza virus), VI: ssRNA-RT viruses (e.g. the human immunodeficiency virus (HIV)), and VII: dsDNA-RT viruses (e.g. hepatitis B virus)\textsuperscript{3,4}. This dissertation mainly focuses on the viruses of group IV.

1.2 Picornaviruses

Positive strand RNA viruses contain three major orders of viruses, including Tymovirales, Nidovirales, Picornavirales, and other unassigned viruses mainly targeting human, animals, and plants\textsuperscript{5}. These viruses can use their single-stranded genome as a template to guide viral protein synthesis\textsuperscript{6}. For genome replication, negative-strand RNA, complementary to the positive-strand RNA, must first be produced to act as a template for the further production of the viral genome\textsuperscript{6}. The family of positive-strand RNA viruses having the most viruses that infect animals is Picornaviridae in Picornavirales. Picornaviridae has 26 genera\textsuperscript{5}, among which enteroviruses, aphthoviruses, and cardioviruses are widely studied. Enteroviruses contain multiple common and significant pathogens threatening human lives including poliovirus (PV) that causes poliomyelitis\textsuperscript{7}, coxsackieviruses that mainly infect newborns and young children and cause hand-foot-and-mouth diseases (HFMD), herpangina, pleurodynia, myocarditis, conjunctivitis and pancreatitis\textsuperscript{8-13}, enterovirus 71 (EV71) that also causes HFMD and may also cause severe poliomyelitis and permanent paralysis\textsuperscript{14-16}, and rhinovirus that causes the common cold\textsuperscript{17}. The Aphthovirus genus contains the foot-and-mouth disease virus (FMDV) that infects cloven-hoofed animals. The pandemics in animals caused by FMDV results in animal slaughter, food security issues, and
huge economic loss in the farming industry\textsuperscript{18,19}. Altogether, positive-strand RNA viruses bring great public health issues, social problems, and tremendous financial loss.

This thesis will focus mostly on PV and coxsackievirus, so it is worthwhile to focus on the life cycle of picornaviruses, as some of the details with the other positive-strand viruses may be different. The life cycle of picornaviruses begins with interactions in the viral canyon\textsuperscript{20,21} The viral canyon is important for receptor specificity: it is a deep surface depression of the viral capsid, whose dimension is proposed to prevent antibody recognition but to specifically bind to cellular receptor at the surface of a cell\textsuperscript{20,21}. After receptor recognition, viruses enter into the cytoplasm and release their RNA (Figure 1-1). The length of the picornaviral genomes ranges from 7500 to 9000 base pairs (bp). The viral genome contains the protein VPg (viral protein genome) covalently linked to the 5’ end, poly(A) tail at the 3’ end, and a large open reading frame (ORF) encoding the polyprotein. The untranslated region (UTR) at the 5’ end contains a clover-leaf structure and internal ribosomal entry site (IRES) (Figure 1-1, 2)\textsuperscript{6}. Before protein translation, VPg is cleaved from the viral genome. Upon the binding of a cellular ribosome to IRES, the VPg free viral genome can directly act as a messenger RNA, so that it can be translated into a polyprotein (Figure 1-1, 2)\textsuperscript{6}. The newly synthesized polyprotein is then proteolytically processed into the structural protein P1, and non-structural proteins P2 and P3, which are then further cleaved into protein precursors and viral proteins important for viral translation, replication, and maturation (Figure 1-2). P1 is converted into structural capsid proteins VP1-VP4 and VP0 (the precursor of VP2 and VP4) (Figure 1-2)\textsuperscript{22}. P2 is converted into 2A\textsuperscript{pro} (proteinase), 2B and 2C\textsuperscript{NTPase}\textsuperscript{22}. P3 is converted into 3A, 3B (also VPg, the primer of RNA synthesis), 3AB (the precursor of 3A and 3B), 3C\textsuperscript{pro}, 3D\textsuperscript{pol} (RNA-dependent RNA polymerase (RdRp)) and 3CD\textsuperscript{pro} (the precursor of 3C\textsuperscript{pro} and RdRp) (Figure 1-2)\textsuperscript{22}. The RdRp, together with other viral proteins and host proteins, forms the replication machinery to synthesize viral genomes (Figure 1-1)\textsuperscript{22}. The positive-strand RNA genomes either re-enter cycles of genome replication or are packed into
viral virions (Figure 1-1). Finally, a large amount of viral particles will be released by cell lysis, and infect other cells (Figure 1-1).
Figure 1-1. The life cycle of picornavirus. The life cycle is initiated by the recognition of the cellular receptor at the surface of a cell. Once virus enters into a cell, RNA (which is indicated as a horizontal line with VPg and poly(A) tail at the 5’ and 3’ end) is released from the virion. The RNA is translated into a polyprotein, and then the polyprotein is proteolytically processed into multiple viral proteins and their precursors. The 3D (RdRp) with other viral proteins and host factors forms a replication organelle on the membrane to synthesize negative-strand RNA (which is indicated as a horizontal line with VPg and poly(U) at the 5’ and 3’ end) as an intermediate. The negative-strand RNA, in turn, acts as a template to allow synthesis of positive-strand RNA. The RdRp that catalyzes RNA synthesis is very error-prone, which leads to introduction of many mutations (which are indicated as small colored symbols on the RNA) into the viral genomes. Finally the RNA is packaged into viral capsid proteins to form viral particles that are released from the cell upon cell lysis.
Figure 1-1 A. The structure of viral genome in enteroviruses. The RNA contains one open reading frame that consists of genes encoding P1, P2 and P3 proteins. VPg and poly(A) tail are at the 5’ and 3’ end of the genome respectively. The 5’ untranslated region (5’ UTR) contains a cloverleaf structure and an internal ribosomal entry site (IRES). B. The polyprotein is then proteolytically processed into multiple precursors and fully mature viral proteins.
1.3 Historical perspective on antiviral strategies

Although great efforts have been made to develop antiviral drugs against enteroviruses, including poliovirus and coxsackievirus infection, we are far from prepared for the next pandemic outbreaks of viral infectious diseases or the emergence of new viruses. Particularly, there are no clinically approved antivirals for enteroviruses. Laboratories have, however, discovered a few antiviral inhibitors targeting viral proteins, including capsid protein (pleconaril and WIN compounds)\textsuperscript{23}, 2C (guanidine hydrochloride, MRL-1237, and TBZE-029)\textsuperscript{24-26}, 3A (TTP-8307, enviroxime, and itraconazole)\textsuperscript{27-29}, 3C (AG7088 and benzerade)\textsuperscript{30, 31}, and RdRp (DTriP-22, ATA, ribavirin, amiloride, GPC-N114, and 5 substituted cytidine derivatives)\textsuperscript{32-38}. Some potent inhibitors have been discovered. For example, in \textit{in vitro} cell culture studies, DTriP-22 inhibited replication of EV71 with an EC\textsubscript{50} value at the submicromolar level\textsuperscript{35}. Host factors can also be targeted. For example, inhibitors of the ubiquitin proteasome system show various degrees of viral inhibition\textsuperscript{39}. However, those inhibitor candidates are not satisfactory for clinical applications, partially due to the limited scope of viruses they can inhibit and occurrence of drug resistance. Development of antiviral drugs with broad spectra and high barriers to resistance are of great significance. A genus of viruses usually contains multiple species, among which one species may have multiple serotypes. Therefore, treating multiple members in one genus or across multiple genera is required for ideal antiviral inhibitors. One further complication is that viruses have high mutation rates and are able to evolve drug resistant mutants in infected cells.

Vaccines are a more preventative type of antiviral treatment are vaccines. In picornaviruses, only three viruses including PV, FMDV, and hepatitis A virus (HAV) have corresponding vaccines being developed\textsuperscript{40-42}. Vaccines prevent viral transmission and protect the host from viral infections. PV vaccines are composed of either inactivated viruses (i.e. the Salk vaccine) or activated but attenuated virus strains (i.e. the Sabin vaccine)\textsuperscript{42}. The attenuated viral vaccines are believed to be more efficient than the inactivated ones because attenuated viruses are
still able to proliferate slowly and be maintained in the long term in restricted tissues, and so can stimulate the full immune response. One classic way to select for attenuated viruses is to allow virulent viruses to replicate over generations under conditions different from those encountered in host environments (e.g. reduced temperature)\textsuperscript{43}. The viruses that adapt to the new environment may no longer replicate efficiently, or may be unable to induce disease in the host organisms. However, vaccine strains developed through this method usually contain large amounts of mutations, and so, it can become difficult to rationalize the efficacy of the vaccine. This complexity prohibits the transfer of this knowledge to related viruses. For example, the Sabin 1 vaccine strain has 55 point mutations compared to its parental Mahoney strain\textsuperscript{44}. Within these 55 point mutations, only a few mutations have been studied but functions (if any) of the remaining mutations are not clear\textsuperscript{45-49}. In this situation, it is hard to choose the specific combination of the mutations to transfer to other viruses without the loss of the attenuated phenotype. More rational strategies may aim to simplify development of the next generation of vaccines. One rational way to design viral vaccine is to produce virus-like particles (VLP) that are hollow capsids and are generated using self-assembly of expressed structural proteins\textsuperscript{50}. The VLP mimics the shape of the natural virus to stimulate immune responses\textsuperscript{50}. VLP vaccines that are developed by the Merck Company are approved to treat hepatitis B virus and human papillomavirus, and provide a promising method to develop vaccines offering protection against other viruses\textsuperscript{50}.

The RdRp is a particularly good target for drug discovery and vaccine development efforts. The RdRp is highly conserved in all RNA viruses, and is the core protein responsible for viral genome replication\textsuperscript{51, 52}. During RNA synthesis, the error-prone RdRp also introduces mutations into the genome, which can be critical for drug resistance and viral adaption\textsuperscript{53-55}. Lastly but most importantly, the RdRp fidelity determines the genetic diversity of a viral population, which opens a new pathway for the development of antiviral drugs and/or vaccines\textsuperscript{53-55}. 
1.4 New antiviral strategies targeting RdRp fidelity

A population of viruses is genetically heterogeneous and can be thought of as a ‘quasispecies’. This genetically diverse viral population can rapidly evolve to escape the immune response and/or drug treatment, which make viral diseases extremely hard to treat. The genetic heterogeneity in the viral quasispecies mainly derives from the introduction of mutations and/or genome recombination as catalyzed by the virally encoded RdRp. For example, the intrinsic error rate of the PV RdRp leads to one or two mutations in each cycle of genome replication. Some antiviral compounds, like ribavirin and 5-fluorouracil, increase the RdRp error rate. In these cases if the mutation rate rises over a certain threshold, viral extinction may occur because of an inability to maintain genetic information, denoted as ‘lethal mutagenesis’. Studies on FMDV replication in the presence of 5-fluorouracil in cell-based cultures showed that FMDV cannot survive after multiple passages due to the accumulation of surplus mutations. Amino acid changes within the RdRp that reduce polymerase fidelity would likely lead to the same consequence. Antiviral drugs leading to lethal mutagenesis might be amenable to a broad spectrum of viruses considering the structural, functional and mechanistic similarities of RdRps in the family of picornaviruses.

Reduction of the genetic variability within the viral quasispecies also leads to viral attenuation, likely due to limited viral adaptability to the host response. Studies on a ribavirin resistant mutant of PV recovered a variant containing a Ser mutation at Gly64 in the coding region of RdRp. The G64S RdRp is a high fidelity variant that is able to discriminate more successfully against ribavirin and also generally leads to a lower rate of nucleotide misincorporation. However, the G64S PV is more sensitive to other types of inhibitors (i.e. inhibitors targeting capsid) than wild type (WT) PV in cell culture studies. G64S PV is also attenuated in terms of neurovirulence in the mouse model. These studies suggest that attenuated virus strains may be attained by encoding viruses with higher fidelity RdRps. For example, PV
encoding the high-fidelity K359R RdRp does not lead to disease in a mouse model of PV, but provides immunoprotection against virulent WT virus to the same extent as the Sabin vaccine strain\textsuperscript{75}. This result suggests that the single K359R mutation in the viral genome has comparative efficacy to that of the 55 mutations in the Sabin strain. Additionally, this residue is highly conserved among other RNA viruses, and so unlike the situation with the Sabin strain, the lessons learned here are more readily applied to other viruses.

### 1.5 PV as a model system for understanding RdRp function

Altogether, the fidelity of the RdRp is a critical consideration in the virulence of the virus, and as such, a mechanistic understanding of nucleotide discrimination is of great value for any drug pipelines targeting RdRp fidelity. The PV RdRp has been selected as a great model to study the fidelity mechanism of RdRps due to many reasons. First, many crystal structures of PV RdRp and its precursor 3CD\textsuperscript{pro} have been published\textsuperscript{76-79} and the complete kinetic mechanism of PV RdRp in the presence of different metal ions has been detailed for more than a decade\textsuperscript{74, 80-86}. Additionally, a combination of site-directed mutagenesis and overexpression of PV RdRp at milligrams level provide a very convenient way to study residues important for enzyme catalysis and fidelity. Also, purification of PV RdRp with an authentic N terminus is achieved\textsuperscript{87}. NMR experiments and computer simulations are already used to study the structural dynamics of PV RdRp\textsuperscript{88-91}. Finally, transgenic mice that express the human PV receptor are available to test viral attenuation and vaccine candidates in an animal model\textsuperscript{92-95}. 
1.6 Catalysis of PV RdRp

Like other template-direct nucleic acid polymerases, PV RdRp has a canonical right hand conformation with “palm”, “thumb”, and “fingers” subdomains (Figure 1-3)\textsuperscript{76,79}. The N-terminus region of PV RdRp bridges the “fingers” and “palm” subdomains, and limits large conformational changes observed in some DNA polymerase that lack this structural motif\textsuperscript{76,79}. PV RdRp has seven highly conserved motifs A to G (Figure 1-3)\textsuperscript{76,79}. The “palm” subdomain has five highly conserved motifs A to E, among which the motifs A to D form the core structure comprised of four antiparallel β sheets and two α helices beneath these β sheets (Figure 1-3)\textsuperscript{76,79}. The core structure is highly conserved in all nucleic acid polymerases\textsuperscript{76,79}. Motifs A and C have two conserved aspartates to coordinate two divalent ions (i.e. Mg\textsuperscript{2+} (physiologically relevant) or Mn\textsuperscript{2+}), which is the canonical catalytic strategy in nucleic acid polymerases (Figure 1-4)\textsuperscript{76,79}. The coordination sites of the two metal ions are similar and involve the two aspartates in motif A (or motif C) and the first aspartate in the motif C (or motif A)\textsuperscript{76,79}. The remaining coordination sites are reserved for the phosphate moiety of the incoming nucleotide and the 3’ OH of the primer (Figure 1-4)\textsuperscript{78}. The metal ion lowers the pKa of the primer 3’ OH to attack the α phosphate of the incoming nucleotide and drives the formation of the phosphodiester bond (Figure 1-4)\textsuperscript{51}. One general acid and one general base have been proposed to facilitate catalysis, termed as the two-proton-transfer mechanism (Figure 1-4)\textsuperscript{51}. The putative general base withdraws the proton from the primer 3’ OH. The general acid in motif D protonates the pyrophosphate leaving group (Figure 1-4)\textsuperscript{85,86,90}. Motif B has a highly conserved Asn297 important for the selection of ribonucleotide against deoxynucleotide\textsuperscript{81,84}. Motifs F and G are in the “fingers” subdomain. Some positively charged residues (such as Lys167 or Arg174) in motif F interact with the remaining non-bridged O in the triphosphate of the incoming nucleotide to further stabilize the nucleotide binding\textsuperscript{78}. The function of the motif G is unclear.
The complete kinetic mechanism of single nucleotide incorporation by PV RdRp has been delineated by the Cameron lab (Figure 1-5). This mechanism starts from the binary complex with RdRp and RNA duplex and has five subsequent elementary steps including (1) nucleotide binding, (2) the prechemistry conformational change to rearrange the ternary complex into a catalytically competent conformation, (3) phosphoryl transfer, (4) the postchemistry conformational change including translocation of the substrate RNA, and (5) release of the pyrophosphate (Figure 1-5). For cognate nucleotide incorporation in the presence of Mg$^{2+}$, both the prechemistry conformational change and the phosphoryl transfer are rate limiting and are important fidelity checkpoints. In the presence of Mn$^{2+}$, only the phosphoryl transfer is rate-limiting.
Figure 1-2 The crystal structure of PV RdRp (PDB 1RA6). A. A ribbon representation of PV RdRp with “fingers” (cyan), “palm” (magenta), and “thumb” (green) subdomains. B. Highly conserved functional motifs in PV RdRp include motifs A (green), B (hot pink), C (orange), D (cyan), E (blue), F (yellow) and G (red).
Figure 1-3 Chemical mechanism of PV RdRp. Nucleotide incorporation is the consequence of formation of phosphodiester bond between the 3’ OH of primer nucleotide (blue) and the α phosphate of the incoming nucleotide (green). Aspartates from motifs A and C coordinate with metal ion B and A respectively. Metal A lowers the pKa of the 3’ OH of primer strand, and an unknown base (B) deprotonates the same hydroxyl group. Lys359 from motif D acts as a general acid to protonate the pyrophosphate group (green) and facilitates the formation of the phosphodiester bond. This figure was adapted with permission from reference 88. Copyright 2012 by the Elsevier Ltd..

Figure 1-4 The pre-steady-state kinetic mechanism of single nucleotide incorporation by PV RdRp. When nucleotide binds to the enzyme-RNA binary complex (step 1), the incoming nucleotide rearranges into a catalytically competent conformation (step 2). Then, the phosphodiester bond forms between the primer nucleotide and the incoming nucleotide. Pyrophosphate (PP₁) also forms (step 3). The competent state relaxes to a nonproductive state accompanied by RNA translocation (step 4). The PP₁ releases (step 5). The RdRp-RNA complex is then ready for the next incoming nucleotide.
1.7 Structure and function relationship in the fidelity of PV RdRp

Different nucleotides bind with similar energetics, suggesting that initial binding of the incoming nucleotide mainly derives from interactions with the undiscriminative triphosphate group of the incoming nucleotide. The prechemistry conformational change is the most important fidelity checkpoint in RdRps since this step allows Watson-Crick base pairing and rearrangement of the ribose ring and the triphosphate of the incoming nucleotide into an optimal conformation to facilitate phosphoryl transfer\textsuperscript{82, 83}. The nucleotide selection model developed by Dr. Cameron’s lab suggests that both motifs A and B participate in the selection of cognate nucleotide through an extensive hydrogen bond network\textsuperscript{84}. In this model, Asp238 on motif A is the liaison between motif A (reorienting the triphosphate moiety) and Asn297 (checking the 2’ OH of the sugar) on motif B\textsuperscript{84}. Perturbations of Asn297 (i.e. binding of nucleotide with a noncognate sugar or nucleobase) transmit to motif A through Asp238, which impairs the rearrangement of the triphosphate and reduces catalysis\textsuperscript{84}. This model is consistent with kinetic and thermodynamic data\textsuperscript{84}. Amino acid changes at positions 238 and 297 lead to reductions in sugar fidelity (i.e. more ability to incorporate nucleotides with 2’deoxy sugars)\textsuperscript{84}. This fidelity strategy is likely conserved in polymerases of other viruses due to the high conservation of motifs A and B\textsuperscript{96}.

Our lab and the Cameron lab have also suggested that conformational rearrangements in motif D that helps to re-position the general acid Lys359 represents another fidelity determinant\textsuperscript{85, 86, 90, 97}. NMR studies on the RdRp-RNA-nucleotide ternary complexes showed that Lys359 (as probed by the chemical shift positions of Met354) forms different conformations depending on the identity of the incoming nucleotide\textsuperscript{90}. Only when cognate nucleotide binds does motif D form the “closed” state to stabilize the ternary complex and to facilitate catalysis\textsuperscript{90}. When noncognate nucleotide binds, motif D stays in the “open” state (Figure 1-6)\textsuperscript{90}. Additionally, substitution of Lys359 for Arg increases fidelity by five fold\textsuperscript{90}.
Besides conserved motifs, some residues remote from the active site indirectly affect catalysis and fidelity of RdRp polymerase. This remote control suggests that RdRp is highly dynamic and that distant residues can regulate the active site residues to form the correct catalytic conformation. For example, the residue Gly64 is more than 10 Å away from the nucleoside-binding pocket of PV RdRp. Substitution of Gly64 for Ser enhances enzyme fidelity via a three-fold reduction in the equilibrium constant for the prechemistry conformational change (i.e. the G64S variant favors the “open” state). Gly64 is proposed to be involved in a conserved hydrogen-bond network containing the N terminus and motif A. Thus, perturbations of this hydrogen-bond network (i.e. a substitution of the Gly64 to Ser) lead to changes in motif A that eventually affects fidelity. These observations suggest that there is a long-range network of interactions between residues near to and distant from the active site and that the long-range network also determines enzyme fidelity.
Figure 1-5 The conformational change of motif D can be monitored by NMR. A. The conformational change of motif D in the X-ray crystal structure of norovirus RdRp complexes with nucleotide (green) and without nucleotide (orange). B. The conformational change of motif D (as sensed by the changes in chemical shift of Met354) of PV RdRp ternary complexes with cognate nucleotide (black) and noncognate nucleotide (red) can be monitored by NMR. The NMR spectra of ternary complexes with cognate and noncognate nucleotides are colored black and red respectively. When noncognate nucleotide binds, motif D stays at the “open” state. The picture was adapted with permission from reference 95. Copyright 2014 by the Elsevier B.V.
1.8 Summary

Although great efforts have been made to investigate the nucleotide selection mechanism of the RdRp, some details are still poorly understood. Previous experiments have highlighted the roles of residues in structural motifs A, B, C and D (see above). In particular, motif D has been shown to undergo an “open” to “closed” state transition to reposition the Lys359 general acid for catalysis. Interestingly, the Sabin vaccine strain encodes one motif D substitution (i.e. T362I). The location of this amino acid substitution so close to the general acid might suggest that this modification would induce changes in the catalysis and fidelity of the RdRp. The other three substitutions in the Sabin RdRp (i.e. D53N, Y73H, and K250E) may also change polymerase function. Delineating the effects of these Sabin substitutions may be important to understanding the biology and pathogenicity of the Sabin 1 vaccine strain.

The functions of other structural motifs are not as clearly understood. For example, motif F has significant contacts with the triphosphate group of the incoming nucleotide, and so residues in motif F may also be involved in nucleotide selection and catalysis. It is well established that the rearrangement of the triphosphate group of the incoming nucleotide is a critical component in the nucleotide selection mechanism of other DNA and RNA polymerases. However, there are currently no detailed investigations of how residues on motif F are involved in the nucleotide selection mechanism of RdRps. Understanding the roles of motifs D and F will lead to a more complete molecular mechanism of RdRp fidelity and will also provide valuable insights into the rational development of vaccine candidates and antiviral drugs.

The subsequent chapters try to clarify some of these issues. Chapter 2 discusses that the T362I substitution derived from the Sabin 1 vaccine strain changes polymerase fidelity via changing the closed/open equilibrium of motif D. Chapter 3 describes cooperative effects among the four substitutions in the Sabin RdRp on fidelity. Chapter 4 describes the triphosphate
reorientation as a new fidelity checkpoint, and the roles of residues on motifs D and F play.

Chapter 5 describes the mechanism of action of a new primer/template inhibitor for CVB3 RdRp.

1.9 References


CHAPTER 2

VACCINE-DERIVED MUTATION IN MOTIF D OF POLIOVIRUS RNA-DEPENDENT RNA POLYMERASE LOWERS NTP INCORPORATION FIDELITY


In this chapter, I was responsible for protein overexpression, purification and all the kinetic studies. Dr. Xiaorong Yang was responsible for protein preparation and all the related NMR experiments. Dr. Ibrahim Moustafa conducted the MD simulation for T362I RdRp. Dr. Cheri Lee performed the plaque assay and pathogenesis studies in mice.

2.1 Introduction

Positive strand RNA viruses are important etiologic agents of many human and animal diseases2-10. Viral persistence and adaption in hosts relies heavily on rapid selection of escape mutants (i.e. immune escape mutants or drug resistance mutants) out of a pool of viral mutants with related genomes (termed as the quasispecies)11-13. These escape mutants may evolve into even more robust ones, which make viral diseases hard to treat11-13. The genetic diversity of the quasispecies mainly derives from the incorporation of noncognate nucleotides into the genome by the virally encoded error-prone RNA-dependent RNA polymerase (RdRp)11-13. By contrast, when the mutation rate of the RdRp increases or reduces over a certain threshold, lethal mutagenesis or viral attenuation occurs due to an inability to maintain the genetic information or because of the
limited adaptability of the virus\textsuperscript{11-13}. Consequently, viral fitness declines. In fact, two fold changes in the RdRp fidelity lead to viral attenuation\textsuperscript{14,15}. Based on these observations, it was suggested that manipulation of RdRp fidelity would offer a new design strategy for antiviral drugs or vaccines. For example, in studies of the effects of fidelity on PV pathogenicity, attenuated strains that restrict population diversity provide immune protection in animal models\textsuperscript{16,17}. Additionally, the designed vaccine candidates have comparable efficacy to the clinical Sabin 1 vaccine strain\textsuperscript{16,17}.

The RNA-dependent RNA polymerase (RdRp) belongs to a polymerase superfamily including DNA-dependent DNA polymerases (DdDp), DNA-dependent RNA polymerases (DdRp), and RNA-dependent DNA polymerases (reverse transcriptase, RT)\textsuperscript{18-21}. All these polymerases utilize a conserved two-metal-ion mechanism (i.e. Mg\textsuperscript{2+} or Mn\textsuperscript{2+}) for phosphodiester bond formation\textsuperscript{20,22}. Metal B coordinates to the triphosphate group of a free nucleotide in solution and enters into the active site when the nucleotide binds to a polymerase\textsuperscript{20,22}. Metal A, which comes to the active site later, lowers the pKa of the 3’OH of the nascent primer and drives nucleophilic attack of the primer hydroxyl to the α phosphate of the incoming nucleotide to form a phosphodiester bond\textsuperscript{20,22}. The two proton transfers are proposed to help accelerate the reaction\textsuperscript{20,22}. In the reaction, one proton is transferred from the primer hydroxyl to an unknown general base. The second proton is added at the β phosphate of the incoming nucleotide by a general acid to facilitate release of the pyrophosphate leaving group\textsuperscript{20,22}. Proton inventory experiments, pH profile studies, and mutagenesis studies indicate that the general acid is a highly conserved lysine residue located on motif D in RdRps and RTs, O helix in A family polymerases and P helix in B family polymerases\textsuperscript{23,24}. Substitution of the general acid to leucine leads to a reduction of catalytic efficiency by 50 to 2000 fold (depending on the type of the polymerase)\textsuperscript{23}.

The close relationship between the general acid and the polymerase fidelity is reflected in studies of the A family polymerases. In A family polymerases, the highly conserved O helix is
an essential fidelity determinant. The conformational change of O helix from an “open” to a catalytically competent “closed” state is very important for catalysis. Crystal structure alignment between apoenzymes and enzyme complexes indicate that when nucleotide binds, the O helix undergoes a conformational change to rearrange the nucleotide substrate to form a proper conformation in the active site and to bring the general acid on the O helix close to the β phosphate\(^{25-27}\). Studies indicate that only when cognate nucleotide binds, the “closed” state forms\(^{25-27}\). The O helix denies noncognate nucleotide incorporation by forming various “ajar” states between “open” and “closed” states such that the noncognate nucleotide no longer forms a productive conformation in the active site\(^{25-27}\). Additionally, perturbations of the O helix by substitutions also lead to changes in fidelity\(^{28-34}\). For example, a substitution at Lys679 in the O helix of an error-prone polymerase v (pol v) enhances the fidelity by 8 to 20 fold\(^{34}\). A glutamate substitution at Ala661 in the O helix of *Thermus Aquaticus* DNA polymerase I (*Taq* pol I) decreases fidelity by at least 5 fold depending on the type of nucleotide incorporation\(^{30}\).

As suggested above, a mechanistic understanding of RdRp fidelity has direct consequences for antiviral strategies. One strategy to modify the polymerase function has been to modify the general acid on motif D. We hypothesized that motif D in RdRps/RTs performs an analogous function as the O helix\(^{23, 35-37}\). Previous NMR experiments and kinetic studies in our lab and the Cameron lab suggest that when cognate nucleotide binds, motif D in PV RdRp closes the active site (termed as the “closed” state) to reposition the general acid Lys359 toward the β phosphate of the incoming nucleotide\(^{36}\). Binding of noncognate nucleotide or a mismatched primer terminus retains motif D in the “open” state and retards phosphodiester bond formation\(^{36}\). Additionally, replacement of Lys359 with Arg leads to a reduction in catalysis by 10 fold and enhances fidelity by 5 fold\(^{36}\). These observations suggest that motif D in PV RdRp is an equivalent element as the O/P helices in the A/B-family polymerases and is an important fidelity determinant.
Substitutions on motif D in PV RdRp may alter enzyme fidelity. One particularly noteworthy substitution is the T362I substitution derived from the RdRp of the Sabin 1 vaccine strain\textsuperscript{38}. The Sabin vaccine was a very successful effort for PV eradication in the 20th century. Its complete mechanism of viral biology and attenuation is not clear although diminished protein translation and RNA synthesis have been suggested to largely contribute to viral attenuation\textsuperscript{39-42}. Here, we hypothesized that the T362I substitution in the Sabin RdRp alters the equilibrium of motif D to disrupt polymerase function and fidelity, which may further affect fitness and biology of the Sabin 1 vaccine strain.

In this Chapter, we demonstrated that the T362I substitution lowers RdRp fidelity by changing the conformational equilibrium between the “open” and “closed” states for more efficient noncognate nucleotide incorporation. The decrease in the RdRp fidelity induced by the T362I substitution likely contributes to viral attenuation of the strain carrying this genetic mutation. These studies highlight motif D as a potential and rational target for modifying the RdRp fidelity to generate attenuated viruses as vaccine candidates.

2.2 Methods

2.2.1 Materials

\[\gamma^{32}\text{P}]\text{ATP} \text{ and } [\alpha^{32}\text{P}]\text{UTP} (>7000\text{Ci/mmol}) \text{ were from VWR-MP Biomedical. Nucleoside 5'}\text{-triphosphates and 2'}\text{-deoxynucleoside 5'}\text{-triphosphates (all nucleotides were ultrapure solutions) were from GE Healthcare. 3'}\text{-Deoxyadenosine 5'}\text{-triphosphate (cordycepin) was from Trilink Biotechnologies. Oligonucleotide dT}_{15} \text{ was from Integrated DNA Technologies, Inc. All RNA oligonucleotides were from Dharmacon Research, Inc. (Boulder, CO). T4 polynucleotide kinase was from New England Biolabs, Inc. [Methyl-13C]Methionine was from} \]
Cambridge Isotope Laboratories. HisPur Ni-NTA resin was from Thermo Scientific. Q-Sepharose fast flow was from Amersham Pharmacia Biotech, Inc. Polyethyleneimine-cellulose TLC plates were from EM Science. The QuickChange site-directed mutagenesis kit was from Stratagene. The plasmid DNA isolation Miniprep kit was from Qiagen. All other reagents were of the highest grade available from Sigma or Fisher.

2.2.2 Plasmid construction

Plasmid pET26-6H-SUMO-T362I-RdRp was made by David Lum. The plasmid pET26-
6H-WT-RdRp was from the Cameron lab.

2.2.3 Expression and purification of RdRp

The autoinducing system was used to overexpress [Methyl-\(^{13}\)C]methionine labeled protein\(^{43,44}\). Briefly, the pET plasmid was transformed into Met-auxotroph Escherichia coli B834 (DE3) pRARE cells in which the pRARE is a plasmid expressing tRNA for rare codons in the cell and the DE3 lysogen carries a gene of T7 polymerase under control of the lacUV5 promoter. The expression of the target protein is under control of a lacT7 promoter, which is specifically recognized by the T7 polymerase. The cells were firstly grown in a non-inducing M9 minimal media with glucose to prevent pre-mature induction and then were used to inoculate into a large scale autoinducing M9 minimal media with lactose to unblock both the lacUV5 and lacT7 promoter and begin protein expression.

For the overexpression of [Methyl-\(^{13}\)C]methionine labeled PV RdRp, the pSUMO-RdRp plasmid was transformed into B834 (DE3) pRARE cells. The cells containing the plasmid were used to inoculate 50 mL non-inducing M9 minimal media and shaken at 250 rpm at 37 °C until
the optical density at 600 nanometer (OD\textsubscript{600}) reaches 1.0. The cells were then inoculated into a 500 mL autoinducing M9 minimal media with OD\textsubscript{600} starting at 0.05 at the same temperature and shaking speed until OD\textsubscript{600} reached 1.0 again, after which temperature was reduced to 25 °C to allow the overexpression of PV RdRp for 20 hours.

The cells were harvested by centrifugation at 6000 rpm for 30 min at 4 °C, washed with 10 mM Tris and 1 mM EDTA buffer (T\textsubscript{10}E\textsubscript{1}), and harvested again by centrifugation in the same condition. The cell pellets were stored at -80 °C until further use.

For protein purification, the protocol was adjusted based on previous procedures.\textsuperscript{45} The frozen cells were thawed and suspended in 50 mL lysis buffer (100 mM potassium phosphate pH 8.0, 20 % glycerol, 5 mM β-mercaptoethanol (BME), 60 µM ZnCl\textsubscript{2}, 1.4 µg/mL pepstatin, and 1 µg/mL leupeptin). The cells were lysed by passing through a French Press cell at 10,000 psi for three times. After each time, 100 mM phenylmethylsulfonyl fluoride (PMSF) was added to comprise one third of the total volume. Nonidet P-40 (NP-40) was added after lysis. The final concentrations of PMSF and NP-40 were 1 mM and 0.1 % (v/v) respectively. Polyethylenimine (PEI) was slowly added over 30 min to precipitate the nucleic acid to a final concentration of 0.25 % PEI. The lysate was stirred for an additional 15 min and was centrifuged (Beckmann) at 16,700 rpm for 30 min at 4 °C. The supernatant was poured into another container. After decanting the supernatant, ammonium sulfate (to a final concentration of 60 % saturation) was slowly added over 30 min to precipitate protein. The supernatant was stirred for an additional 15 min and was centrifuged at 16,700 rpm for 30 min at 4 °C. The supernatant was decanted and the pellets were resuspended in buffer B (100 mM potassium phosphate pH 8.0, 20 % glycerol, 5 mM BME, 60 µM ZnCl\textsubscript{2}, 500 mM NaCl, 5 mM imidazole, 1 mM PMSF and 0.1 % NP40) to a final concentration of 5.4 mg/mL crude protein. The suspended protein was loaded onto a Ni-NTA column pre-equilibrated with the buffer B at a flow rate of 1 mL/min (approximately 1 mL resin/50 mg crude protein). The Ni-NTA column was washed the first time with 10 column
volumes of the buffer B and was washed the second time to the baseline (monitored by absorbance at 280 nm) with buffer C (100 mM potassium phosphate pH 8.0, 20 % glycerol, 5 mM BME, 60 µM ZnCl₂, and 500 mM NaCl) containing 5 mM imidazole and 1 mM PMSF. The column was washed the third time with the buffer C containing 50 mM imidazole and 1 mM PMSF with 5 column volumes. The protein was then eluted with 10 column volumes of the buffer C containing 500 mM imidazole and 1 mM PMSF. The eluted protein was dialyzed (12-14,000 Da MWCO membrane) overnight against 1 L buffer D (50 mM Tris pH 8.0, 500 mM NaCl, 5 mM BME, 60 µM ZnCl₂, and 20 % glycerol) with Ulp 1 enzyme (1 µg / 1 mg SUMO fusion) to cleave the SUMO fusion. The dialyzed protein was dialyzed a second time for 2 h against 1 L buffer E (100 mM potassium phosphate pH 8.0, 5 mM BME, 60 µM ZnCl₂, and 20 % glycerol). The dialyzed protein was passed through the second Ni-NTA column pre-equilibrated with the buffer E at a flow rate of 1 mL/min to remove the cleaved SUMO fusion. The flow-through was collected and dialyzed the third time with 1 L buffer F (50 mM Tris pH 8.0, 1 mM DTT, and 20 % glycerol) containing 25 mM NaCl. The dialyzed protein solution was then loaded onto the Q-Sepharose column (1 bed vol / 20 mg protein) pre-equilibrated with the buffer F containing 25 mM NaCl. The column was washed and eluted with the buffer F containing 25 mM and 500 mM NaCl respectively. The protein purification was monitored by SDS-PAGE. The molar extinction coefficient of RdRp at 280 nm was 75,750 M⁻¹cm⁻¹. The purified protein was aliquoted and stored at -80 °C.

2.2.4 Purification, 5'-³²P end labeling and annealing of sym/sub

For kinetic studies, the RNA oligonucleotide sym/sub was gel purified as described previously⁴⁶. Briefly, 2'-ACE (bis(2-acetoxyethoxy)methyl) protected sym/sub RNA was purified by 23 % denaturing PAGE gel containing 1.7 % bisacrylamide, 7 M urea and TBE buffer (89 mM
Tris base, 89 mM boric acid, and 2 mM EDTA). The sym/sub RNA was visualized via short length UV. The gel containing sym/sub RNA was cut and chopped. Nucleic acid sym/sub was electroeluted from the gel in TBE buffer by Elutrap. The sym/sub RNA was desalted by Sep Pak column (Millipore) following the instructions from the manufacturer. The sym/sub RNA was dried by vacuum evaporation and was stored at -20 °C. Prior to use, the purified sym/sub RNA was deprotected following the instructions from the manufacturer (Dharmacon) and desalted by passing through a 1 mL G25 column.

The procedure for $^{32}$P labeling of sym/sub at the 5’ end was described previously\textsuperscript{46}. Briefly, $^{32}$P γ-labeled ATP and T4 polynucleotide kinase were added into the sym/sub RNA in the T4 reaction buffer provided by the manufacturer to final concentrations of 1 µM and 0.4 unit/µL respectively. The reaction was performed at 37 °C for 1 h and was quenched by incubating at 65 °C for 5 min.

For annealing of sym/sub RNA\textsuperscript{46}, the sym/sub RNA was diluted to a certain concentration in T\textsubscript{10}E\textsubscript{0.1} buffer (10 mM Tris pH 8.0 and 0.1 mM EDTA). The annealing was conducted in a thermocycler following a program with denaturing at 90 °C for 1 min and decreasing to 10 °C with a rate of 5 °C/min.

For NMR studies, RNA oligonucleotide sym/sub was directly deprotected without any further purification\textsuperscript{43}. The RNA product was dried by SpeedVac and diluted in D\textsubscript{2}O to the proper concentration.

### 2.2.5 PV RdRp assay\textsuperscript{46}

Reactions contained 50 mM HEPES 7.5, 10 mM BME, 5 mM MgCl\textsubscript{2} or MnCl\textsubscript{2}, and 60 µM ZnCl\textsubscript{2}. Reactions were firstly incubated at 30 °C for 5 min and were initiated by the addition of either RdRp or ATP. RdRp enzyme was diluted immediately prior to the reaction with enzyme
dilution buffer (20% glycerol, 50 mM HEPES, 10 mM BME and 60 µM ZnCl₂) and was added to a volume equal or less than one tenth of the total volume of the reaction. Reactions were quenched by addition of EDTA to a final concentration of 25 mM at each time point. Final concentrations of RdRp, end-labeled sym/sub RNA, and nucleotides, along with any deviations from above, are indicated in the appropriate figure legend.

2.2.6 Stopped flow experiments

The reaction buffer used in the stopped flow experiments was the same as the buffer in the PV RdRp assay (see 2.2.5). Reactions were performed at 30 °C using a circulating water bath. Binary complexes of RdRp-RNA were formed by incubating at room temperature for 2 min. Reactions were initiated by rapidly mixing the newly formed binary complexes with an equal volume of nucleotides of specific concentrations. Radioactive sym/subU and sym/subUA RNA were substituted with fluorescent nucleotide analog 2-aminopurine and pyrrolo-cytosine at the 5’ side of the template respectively. Fluorescence change corresponding to product formation was monitored during the chain elongation reaction. For sym/subU RNA with the 2-aminopurine modifications, the excitation and emission wavelengths were 313 and 370 nm respectively. For sym/subUA RNA with the pyrrolo-cytosine modifications, the excitation and emission wavelengths were 350 and 450 nm respectively. Final concentrations of RdRp, end-labeled sym/sub, and nucleotides, along with any deviations from above, are indicated in the appropriate figure legend.
2.2.7 Benchtop assay for slow nucleotide incorporation

Reaction buffer used in the benchtop experiment was the same as the buffer in the PV RdRp assay (see 2.2.5). Binary complexes of RdRp and end-labeled sym/sub RNA were formed by incubating the complexes at room temperature for 3 min firstly followed by 30 °C for 2 min. Reactions were initiated by the addition of an equal volume of nucleotides with a certain concentration. Reactions were quenched by addition of EDTA to a final concentration of 25 mM at different time points. Final concentrations of RdRp, end-labeled sym/sub RNA, and nucleotides, along with any deviations from above, are indicated in the appropriate figure legend.

2.2.8 Product analysis and data analysis

Product analysis and all data fitting were followed by previous procedures. Briefly, quenched product samples were preheated at 65 °C for 5 min and were isolated via denaturing PAGE gel at constant 90 watts for 2.5 h. The gel was then exposed to the screen, visualized by PhosphorImager, and quantified by ImageQuant software (Molecular Dynamics). The data is fit into different functions by KaleidoGraph software (Synergy) as indicated in the appropriate figure legend.

2.2.9 Sample preparation for NMR spectroscopy

Purified protein with [Methyl-13C]methionine labeling was concentrated to 200 - 400 µM, desalted using a spin desalting column (Thermo scientific), and exchanged into NMR buffer containing 10 mM HEPES 8.0, 200 mM NaCl, 5 mM MgCl2, 20 µM ZnCl2, and 0.02% NaN3 in D2O. Binary complexes of RdRp-sym/subU were formed by incubating RdRp (200 - 250 µM) with sym/subU RNA (200 - 500 µM) and 3’ dATP (3.1 mM). The purpose of adding of 3’dATP
was for it to be incorporated into primer sym/subU RNA at the 3’ end by the RdRp in situ such that the newly formed binary complex with the 3’ blocked end would facilitate trapping of the ternary complex after the addition of the next incoming nucleotide. The formed 3’ block complexes were passed through a desalting column to remove any excess amount of 3’dATP prior to the addition of the second nucleotido. Concentration of the second nucleotidio ranged from 4 to 12 mM.

2.2.10 \(^1\)H - \(^{13}\)C HSQC NMR experiment\(^{43}\)

NMR experiments were performed on a Bruker Avance III 600 MHz spectrometer equipped with a 5 mm inverse detection triple resonance \(\(^1\)H/^{13}\)C/\(^{15}\)N) single axis gradient TCI cryoprobe. \(^1\)H - \(^{13}\)C heteronuclear single quantum coherence (HSQC) spectra were generally acquired at 64 \((t_1) \times 512 \(t_2)\) complex matrix, with 64 - 128 scans per increment and 1.0 s recovery delay.

2.3 Results

2.3.1 The T362I substitution lowers RdRp fidelity by making the pre-chemistry conformational change for misincorporation more efficient.

The T362I substitution is derived from the Sabin 1 vaccine strain and is located in motif D in the “palm” subdomain of PV RdRp (Figure 2-1)\(^{48,49}\). The single substitution may change enzyme fidelity via interfering with conformational changes in motif D including those that help to reposition the general acid Lsy359. Before testing this hypothesis, it is important to know whether the substitution affects RNA binding affinity and stability of RdRp-RNA binary complexes. I performed the assembly and dissociation assays to look at the formation and the
stability of enzyme complexes, respectively, for both T362I and WT RdRps. In kinetic assays, two RNA primer/templates were used, those being sym/subU and sym/subUA RNAs, which contain six and seven complimentary base pairs with four- and three-nucleotide overhangs respectively (Figure 2-2). In the sym/subU RNA, U is the templating nucleotide and ATP is the corresponding cognate nucleotide. In the sym/subUA RNA, A is the templating nucleotide and UTP is the corresponding complimentary nucleotide (Figure 2-2). For the assembly assay, the T362I RdRp formed a comparative amount of catalytically competent enzyme complexes as WT RdRp regardless of the RNA primer/template. When the sym/subUA RNA was the primer/template, the enzyme complexes were formed in more than the stoichiometric quantity, which suggested that both strands in the sym/subUA duplex could act as a primer substrate. This finding was different from the sym/subU RNA where only one strand in the duplex could act as a primer substrate (Figure 2-2).

In the dissociation assay, the T362I RdRp complexes were as stable as those of WT RdRp regardless of the RNA primer/template (Figure 2-2). The RdRp-sym/subUA complexes were more stable than the RdRp-sym/subU complexes since the dissociation rate increased by more than 10 fold (Figure 2-2). The half-lives of the RdRp-RNA complexes ranged from 0.5 to 11 h, which was sufficiently long for the subsequent transient kinetic studies. Altogether, the T362I substitution did not interfere with the assembly and the stability of enzyme complexes regardless of the RNA primer/template used.
Figure 2-1. The T362I substitution may impact the structural dynamics of motif D and other regions to affect RdRp catalysis and fidelity. A, B. Crystal structure of PV RdRp (PDB 1RA6). Colored motifs include motif A (green), B (hot pink), C (orange), D (cyan), E (blue), F (yellow) and G (red). Thr362, Lys359, and probe Met354 are shown in purple, cyan, and purple respectively. The Met probes that are sensitive to RNA or nucleotide binding are shown in gold. C. Kinetic schemes of single nucleotide incorporation for both Mg$^{2+}$ and Mn$^{2+}$ include 5 steps. From steps 1 to 5, those are nucleotide binding, pre-chemistry conformational change, chemical reaction, post-chemistry conformational change, and release of the pyrophosphate group. For cognate nucleotide insertion, step 2 (red) is rate limiting only in Mg$^{2+}$, and step 3 (light blue) is rate limiting in both Mg$^{2+}$ and Mn$^{2+}$.
I then used single nucleotide incorporation assays to study whether the T362I substitution changes enzyme fidelity. In these kinetic assays, fidelity is defined as the ratio of the catalytic efficiency of cognate nucleotide over that of noncognate nucleotide. The catalytic efficiency in these assays equals the ratio of the maximal rate constant $k_{pol}$ over the apparent dissociation constant $K_{d,app}$. The results indicated that the T362I substitution lowered enzyme fidelity for both sugar and nucleobase selection regardless of the RNA primer/template (Table 2-1). The reduction in enzyme fidelity was due to more efficient incorporation of the noncognate nucleotide without significant changes in the incorporation of the cognate nucleotide (Figure 2-3). T362I RdRp had a very similar catalytic efficiency as WT RdRp for AMP incorporation ($k_{pol}/K_{d,app}$ values were 1.6 and 2.0 $\mu$M$^{-1}$s$^{-1}$ for WT and T362I RdRps respectively) (Figure 2-3, Table 2-1). However, T362I RdRp facilitated 2’dAMP misincorporation by nearly two-fold over that observed with WT enzyme ($k_{pol}/K_{d,app}$ values were $6.7 \times 10^{-3}$ and $1.3 \times 10^{-2}$ $\mu$M$^{-1}$s$^{-1}$ for WT and T362I RdRps respectively) (Figure 2-3, Table 2-1).

To test whether T362I RdRp also has reduced fidelity for other nucleotides and their analogs, I determined the kinetic values for ribavirin incorporation. Ribavirin was the earliest FDA (US Food and Drug Administration) approved antiviral drug to treat hepatitis C infection when combined with interferon $\alpha$. Upon entering the cell, ribavirin is firstly transformed to ribavirin triphosphate (RTP) via phosphorylation and then serves as a guanidine triphosphate analog, such that it can become incorporated into the viral genome through the action of the RdRp and can form Watson-Crick base pairing with either U or C$^{15}$. This process can eventually lead to lethal mutagenesis of the virus. The results showed that T362I RdRp more readily utilized RTP than WT enzyme (Table 2-1), which suggests that PV encoding this mutant may be more sensitive to ribavirin treatment.
The T362I substitution did not affect the apparent dissociation constant $K_{d,\text{app}}$ for any of the nucleotides, but significantly changed the maximal rate constant $k_{\text{pol}}$, compared to WT enzyme (Table 2-1). These observations suggested that the T362I substitution interferes with the intrinsic conformational change and/or the chemistry reaction after nucleotide binding, which contribute to the $k_{\text{pol}}$ value (Figure 2-1). In the presence of Mg$^{2+}$, both the pre-chemistry conformational change and the chemistry step are rate-limiting steps and fidelity checkpoints. In order to test which step acts as the main fidelity checkpoint here, I also determined the Mn$^{2+}$ and solvent deuterium kinetic isotope effects (SDKIE).

The SDKIE is a sensor of the chemistry step, which reflects the relative contribution of the chemistry and the pre-chemistry conformational change steps to $k_{\text{pol}}$. The SDKIE equals the quotient of the maximal rate constants in water and heavy water (i.e. D$_2$O) (i.e. SDKIE = $k_{\text{H}_2\text{O}}/k_{\text{D}_2\text{O}}$). For T362I RdRp, the SDKIE slightly increased for both the cognate (2.4 ± 0.02) and noncognate (1.8 ± 0.07) nucleotide incorporations compared to WT RdRp (cognate nucleotide, 2.18 ± 0.02; noncognate nucleotide, 1.51 ± 0.06) (Table 2-1). The increase in the SDKIE values suggested that the chemistry reaction was more rate-limiting in T362I RdRp.

I also performed single nucleotide incorporation assays under the same reaction condition as before but used Mn$^{2+}$ as the catalytic metal ions. In the presence of Mn$^{2+}$, only the pre-chemistry conformational change step determines enzyme specificity, and the chemistry step is more rate-limiting compared to those in Mg$^{2+}$. The results indicated that T362I RdRp had the same catalytic rate as WT RdRp for the cognate nucleotide incorporation and that T362I RdRp increased the catalytic rate for noncognate nucleobase incorporation ($k_{\text{pol, GTP, T362I}}/k_{\text{pol, GTP, WT}} = 1.4$) (Figure 2-4). The promiscuous incorporation of the noncognate nucleobase suggested that the T362I substitution shifts the pre-chemistry conformational change more to the “closed” state, which is consistent with the SDKIE result that indicates that the chemistry reaction was more rate limiting. Surprisingly, the catalytic rate of 2’deoxynucleotide incorporation for T362I RdRp was
almost the same as that for WT RdRp ($k_{pol, 2'dATP, T362I}/k_{pol, 2'dATP, WT} = 1.1$) (Figure 2-4). This finding may be because the T362I substitution only leads to a small perturbation of the pre-chemistry conformational change that is not reported by the kinetic parameters. Altogether, the T362I substitution lowers enzyme fidelity by shifting the pre-chemistry conformational change to the “closed” state.
Figure 2-2. Formation and stability of RdRp-RNA complexes are not affected by the T362I substitution in motif D. Reactions contained 50 mM HEPES 7.5, 10 mM BME, 5 mM MgCl$_2$ and 60 µM ZnCl$_2$. A. Sequence of both sym/subU and sym/subUA RNAs. B, C. In the assembly assay, ATP and RNA were first incubated for 5 min followed by addition of RdRp to initiate the reaction. Reactions were quenched by addition of 25 mM EDTA. Final concentrations of RdRp, RNA, and ATP were 1, 1, and 500 µM respectively. The data for formed WT RdRp-RNA complexes (empty circle) and T362I RdRp-RNA complexes (solid circle) is plotted as a function of time. C, D. In the dissociation assay, RdRp-RNA binary complexes were formed in 90 s followed by the addition of trap (100 µM). At each time point, ATP was allowed to react for 30 s followed by the addition of 25 mM EDTA to quench the reaction. Final concentrations of RdRp, RNA, and ATP were 1, 0.1, and 500 µM respectively. Percentage of the remaining binary complex is plotted as a function of time for both WT (solid circle) and T362I (empty circle) RdRps. The solid lines represent the fit of the data into single exponential function. The $k_{dis}$ of WT enzyme for sym/subU and sym/subUA RNAs are $4.00 \pm 0.09 \times 10^{-4}$ s$^{-1}$ and $1.70 \pm 0.07 \times 10^{-5}$ s$^{-1}$ respectively. The $k_{dis}$ of T362I RdRp for sym/subU and sym/subUA RNAs are $3.90 \pm 0.02 \times 10^{-4}$ s$^{-1}$ and $2.70 \pm 0.03 \times 10^{-5}$ s$^{-1}$ respectively. The figure was reprinted from reference 1. © The American Society for Biochemistry and Molecular Biology.
Figure 2-3. The T362I substitution lowers fidelity in the presence of Mg$^{2+}$. Reactions contained 50 mM HEPES 7.5, 10 mM BME, 5 mM MgCl$_2$, and 60 µM ZnCl$_2$. RdRp and RNA were first incubated for 2 min to form RdRp-RNA binary complexes. Reactions were initiated by the addition of an equal volume of nucleotide at various concentrations. Changes in fluorescence were recalculated as product formed in time. Final concentrations of RdRp and RNA were 0.5 and 0.5 µM respectively. A. Reaction scheme of single nucleotide incorporation in the presence of Mg$^{2+}$. B. For cognate ATP incorporation, T362I RdRp (empty circle) is as efficient as WT RdRp (solid circle). C, D. For noncognate 2’dATP and GTP incorporations, T362I RdRp (empty circle) is more efficient than WT RdRp (solid circle). Solid lines represent the fit of the data into hyperbola function. The obtained $k_{pol}$ and $K_{d,app}$ for each nucleotide incorporation are shown in table 2-1. The figure was reprinted from reference 1. © The American Society for Biochemistry and Molecular Biology.
Figure 2-4. The T362I substitution lowers sugar and nucleobase selection by making the pre-chemistry conformational change more efficient. Reactions contained 50 mM HEPES 7.5, 10 mM BME, 5 mM MnCl$_2$, and 60 µM ZnCl$_2$. RdRp and RNA were first incubated for 2 min to form RdRp-RNA binary complexes. Reactions were initiated by the addition of an equal volume of nucleotide at various concentrations. Changes in fluorescence were recalculated as product formed. Final concentrations of RdRp and RNA were 0.5 and 0.5 µM respectively. A, B. For cognate ATP and noncognate 2’dATP incorporations, T362I RdRp (empty circle) is as efficient as WT RdRp (solid circle). C. For noncognate GTP incorporation, T362I RdRp (empty circle) is more efficient than WT RdRp (solid circle). Solid lines represent the fit of the data into hyperbola function. The obtained $k_{pol}$ and $K_{d,app}$ for each nucleotide incorporation are shown in Table 2-1. The figure was reprinted from reference 1. © The American Society for Biochemistry and Molecular Biology.
Table 2-1. Comparison of nucleotide incorporation for WT and T362I RdRps reveals that the T362I substitution decreases polymerase fidelity. The table was reprinted from reference 1. © The American Society for Biochemistry and Molecular Biology.

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Metal</th>
<th>RNA</th>
<th>NTP</th>
<th>( k_{pol} ) (s(^{-1}))</th>
<th>( K_{D,app} ) (µM)</th>
<th>( k_{pol}/K_{D,app} ) (µM(^{-1})s(^{-1}))</th>
<th>( k_{pol,correct} ) ( k_{pol,incorrect} )</th>
<th>( (k_{pol}/K_D)_{correct} )</th>
<th>( (k_{pol}/K_D)_{incorrect} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>ATP</td>
<td>UTP</td>
<td>130 ± 7</td>
<td>142 ± 36</td>
<td>0.91</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>T362I</td>
<td>Mg(^{2+}) S/S-UA</td>
<td>UTP</td>
<td>131 ± 2</td>
<td>114 ± 9</td>
<td>1.1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>WT</td>
<td>ATP</td>
<td>UTP</td>
<td>16 ± 1</td>
<td>1.7 ± 0.2</td>
<td>9.4</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>T362I</td>
<td>ATP</td>
<td>UTP</td>
<td>17 ± 1</td>
<td>1.6 ± 0.1</td>
<td>10.6</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>WT</td>
<td>Mn(^{2+}) S/S-U</td>
<td>2’dATP</td>
<td>7.4 ± 0.1</td>
<td>6.1 ± 0.2</td>
<td>1.2</td>
<td>2</td>
<td>2.2</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>T362I</td>
<td>Mn(^{2+}) S/S-U</td>
<td>2’dATP</td>
<td>8.0 ± 0.1</td>
<td>3.9 ± 0.3</td>
<td>2</td>
<td>2.2</td>
<td>5</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>WT</td>
<td>GTP</td>
<td>UTP</td>
<td>0.78 ± 0.01</td>
<td>98 ± 3</td>
<td>8.0 x 10(^{-3})</td>
<td>21</td>
<td>1200</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>T362I</td>
<td>GTP</td>
<td>UTP</td>
<td>1.1 ± 0.1</td>
<td>100 ± 6</td>
<td>1.1 x 10(^{-2})</td>
<td>15</td>
<td>960</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
2.3.2 Solution-state NMR indicates that the T362I substitution induces a structural rearrangement within motif D to lower RdRp fidelity.

Our previous solution state NMR experiments investigated the conformational changes along the catalytic pathway of RdRp by following chemical shift changes of $^{13}$C-Methyl]Met resonances. Met354 is located on motif D and so, it is especially sensitive to conformational changes in the motif-D active-site loop. In these experiments, the sym/subU RNA was the primer/template. In order to capture the RdRp-RNA-nucleotide ternary complex, 3’dATP was incorporated first and the ternary complex was formed with the addition of the next nucleotide. In the cases here, the templating nucleotide was A and so, the corresponding cognate nucleotide was UTP. When cognate nucleotide binds, a conformational change takes place to form the competent “closed” conformation (Figure 2-5). However, when 2’deoxynucleotide (2’dUTP) or noncognate nucleobase (ATP) binds, the same conformational changes are not induced (Figure 2-5). The nonproductive conformation of the noncognate ternary complexes was denoted as the “open” state (Figure 2-5). These observations suggested that the “open” to “closed” state switch of motif D is sensitive to the identity of the incoming nucleotide.

With the T362I variant, there was a small chemical shift change to the Met354 resonance in the absence (data not shown) and presence of RNA and nucleotide (Figure 2-5). When cognate nucleotide bound, the T362I RdRp formed a similar “closed” state as observed with WT RdRp (Figure 2-5). However, when 2’deoxynucleotide bound, resonances associated with the “open” and “closed” states were both observed (Figure 2-5). This observation suggested that the T362I variant more readily adopts the “closed” conformation in the presence of 2’deoxynucleotide. This finding is consistent with the SDKIE and the Mg$^{2+}$/Mn$^{2+}$ effects that suggest that the prechemistry conformational change more readily occurs with the T362I variant even in the presence of noncognate nucleotide (Figure 2-4, Table 2-1). It should be kept in mind that the concentration of
nucleotide (2’dUTP = 10 mM, ATP = 8 mM) was identical for both T362I and WT RdRps, and binding saturation was reached given the $K_{d,app}$ of the noncognate nucleotides in Table 2-1.
Figure 2-5. The T362I substitution alters the motif D equilibrium between the open and closed states as revealed by NMR spectra of [Methyl-\textsuperscript{13}C]Met-labeled RdRp. A, B, and C. WT RdRp clearly distinguishes cognate UTP (black) from noncognate 2’dUTP (red) and ATP (blue) by binding only UTP into the “closed” state and remaining the “open” state of 2’dUTP or ATP. D, E, and F. Motif D in T362I RdRp completely or partially shifts to the “closed” state when UTP (black) or 2’dUTP (red) is bound respectively. However, T362I RdRp is unable to significantly populate the “closed” state when ATP binds (blue). The figure was reprinted from reference 1. © The American Society for Biochemistry and Molecular Biology.
2.3.3 The T362I substitution results in long-range structural changes that may allow for more efficient misincorporation.

To gain more insights into how the T362I substitution changes the structure and/or the internal dynamics of the RdRp to further impact catalytic function and fidelity, an in silico model was constructed for the T362I variant (as there was not a crystal structure available for this particular protein). A 20 ns all-atom MD simulation was performed. The intention of these studies was to focus on interactions that were perturbed by the T362I substitution. The MD simulation of the T362I variant was compared with that of the WT enzyme from previous studies. It should be emphasized that these MD simulations are for RdRp in the form of apo enzyme. Nonetheless, it has been suggested that the structural dynamics of the apo enzyme reflects the ensemble of conformations that is sampled by RdRp in its catalytic cycle. Additionally, previous MD simulations have given insights into the function and fidelity of RdRp, which have been consistent with NMR studies.

The MD simulation results revealed that the T362I substitution induces changes to the protein structures at sites both local and distant from the substitution site. In the simulated T362I structure, the side chain of Ile362 appeared to be stabilized by van der Waals interaction with the side chain of the nearby Phe363 (Figure 2-6). The average distance between the two side chains was 4.6 Å. The T362I substitution greatly affected the conformation of motif D, especially that of Glu364 (Figure 2-6). In WT enzyme, the Glu364 residue interacts electrostatically with Lys228, which precedes motif A (amino acids 229-240) in the WT RdRp (Figure 2-6). With the T362I variant, the Glu364 is flipped over to the other side and interacts with Asn370 and Thr367 (by H bonding interactions) and Lys359 (by electrostatics) (Figure 2-6). The conformational change of Glu364 also led to smaller changes in other motif D residues including Thr365, Glu337, Asp339, and Ser341 (Figure 2-6).
In addition to these local conformational changes, the T362I substitution also impacted more distant residues (Figure 2-6). These conformational changes included residues Ser288 and Asp238 in motif A, which play important roles in sugar selection and triphosphate rearrangement, and Asp328 and Asp329 in motif C, which are partly responsible for catalytic metal ion coordination (Figure 2-6).

To look at the perturbations of global dynamics by the T362I substitution, the per-residue mean square deviation (per-RMSD) was calculated. The results showed a restrained global dynamics of RdRp on the nanosecond timescale for the T362I variant compared to the WT enzyme (Figure 2-6). Interestingly, the T362I substitution led to reduced dynamics in regions encompassing some of the other Sabin 1 substitution sites (Figure 2-6).
Figure 2-6. The T362I substitution in motif D leads to structural and dynamics changes as revealed by MD simulations, consistent with changes to RdRp function. A. The average structure of WT RdRp from the previous study (grey) and T362I RdRp (blue). The structures are overlapped at the Cα atoms. Four substitution sites (Asp53, Tyr73, Lys250, and Thr362) in RdRp of the Sabin I vaccine strain are shown as sticks. B. Zoom-in view of the surrounding regions at Thr362 depicted in panel A. Conformational changes of select residues (shown as sticks) near Thr362 or Ile362. C. Zoom-in view of the active site distant from the substitution site T362I in A. Conformations of many conserved, functional residues (shown as sticks) in motif A (Asp238), B (Thr293), C (Asp328 and Asp329), and D (Lys359) are changed by the single substitution of T362I. D. The per residue root mean square deviation (per RMSD) calculated from the last 10 ns of WT (grey) and T362I (black) trajectories. The per RMSD corresponding to the four substitution sites in the Sabin I vaccine strain are shown in dark red diamond in WT RdRp and dark red triangle in T362I RdRp. The T362I substitution constrains the overall dynamics of RdRp and is also able to communicate with other Sabin sites via changing their structural dynamics. The figure was reprinted from reference 1. © The American Society for Biochemistry and Molecular Biology.
2.3.4 T362I PV expresses a mutator phenotype in vivo and has reduced virulence in the mouse model.

To determine how the biophysical and biochemical changes associated with the T362I substitution impact virus multiplication in cell culture, we constructed a PV genome encoding the T362I RdRp by changing the ACA codon encoding Thr to the ATC codon encoding Ile. Passage of T362I PV at a low multiplicity of infection showed no evidence of reversion to the wild-type codon even after four consecutive passages, suggesting that there was no strong fitness consequence for this mutation in cell culture.

In kinetic studies, the T362I RdRp incorporated RTP (the active form of ribavirin) more readily than the WT enzyme. To determine the consequences of this in vivo, we compared the sensitivity to ribavirin between WT, T362I, and G64S PVs (the virus showed antiviral resistance against ribavirin) (Figure 2-7). In plaque assays, T362I PV was two-fold more sensitive to ribavirin than WT PV (Figure 2-7). The IC₅₀ values for WT and T362I PVs were 0.67 ± 0.06 mM and 0.30 ± 0.02 mM respectively (Figure 2-7). G64S PV was the least sensitive to ribavirin considering the IC₅₀ was 2.15 ± 0.75 mM (Figure 2-7). To determine the impacts of the T362I polymerase on the viral virulence in animals, we conducted virulence tests in transgenic mice expressing the human PV receptor. In this system, the infection by WT PV is generally lethal. In the case of infection at the highest dose (1.6 x 10¹¹ genomes), T362I PV resulted in the delay of onset of double-limb paralysis by 1 day compared to WT PV (Figure 2-7). At the lower dose (1.6 x 10¹⁰ genomes), the onset of double-limb paralysis showed even more gradual procession (Figure 2-7). These results are consistent with the suggestion that suboptimal of fidelity leads to viral attenuation.
Figure 2-7. PV encoding the T362I substitution in motif D expresses a mutator phenotype in cell culture and is attenuated in vivo. A. In the ribavirin sensitivity assay, HeLa cells were pretreated with ribavirin at various concentrations and were infected with WT, G64S, and T362I PVs at 50 plaque forming unit. Plaque numbers relative to control (without ribavirin) are plotted as a function of ribavirin concentration. The solid lines represent the fit of the data to a sigmoidal dose response equation (four parameter logistic model). The IC₅₀ values of WT, G64S, and T362I PVs were 0.67 ± 0.06, 2.15 ± 0.75, and 0.3 ± 0.02 mM respectively. B, C. T362I PV was attenuated in the cPVR (poliovirus receptor) mouse model. cPVR mice (4 to 6 weeks old) were infected via intraperitoneal route with serum-free media containing WT and T362I PVs at the indicated genome copy and observed for 10 days. Mice that developed symptoms of double limb paralysis or paralysis that compromised the ability to get food and water were euthanized. The figure was reprinted from reference 1. © The American Society for Biochemistry and Molecular Biology.
2.4 Discussion

Previous studies have shown that mutations in the coding region of RdRp derived from the Sabin vaccine strain lead to viral attenuation\(^{40\text{-}42, 53}\). Here, we have shown that the T362I substitution in PV RdRp reduces fidelity by making the pre-chemistry conformational change more efficient. NMR experiments and MD simulations showed the impacts of the T362I substitution on the conformational dynamics of both motif D and active site residues distant from motif D. PV carrying the low fidelity RdRp variant enhanced the sensitivity to antiviral drug ribavirin in cell culture and had reduced virulence in a cPVR (poliovirus receptor) mouse model. These observations highlight the importance of RdRp fidelity to viral fitness. We have proposed that motif D is equivalent to the O/P helices in A/B family polymerases\(^{23, 35, 36}\). There are many similarities: both the regions contain a highly conserved lysine that acts as a general acid to protonate the leaving group pyrophosphate, contribute significantly to enzyme structural dynamics, and help determine polymerase fidelity\(^{23}\). Substitutions at Lys359 in PV RdRp lead to a decrease in catalysis by 10 fold and an increase in fidelity by 5 fold\(^{23}\). Substitutions in the O/P helices in DNA polymerases lead to similar results\(^{28\text{-}33}\). Here, we showed that substitutions elsewhere in motif D change RdRp fidelity further demonstrating that motif D is a key structural element for catalysis and fidelity.

It has been suggested that RNA viruses encoding RdRp enzymes with altered fidelity may serve as vaccine strains. Previous studies have shown that PV encoding a high fidelity RdRp variant (i.e. K359R\(^{17}\)) provides immunoprotection almost as good as the Sabin 1 vaccine in the PV mouse model. Here, we have shown that by changing other residues in motif D (i.e. T362I), we can lower RdRp fidelity, and viruses encoding the RdRp variant are attenuated. These results suggest that the fidelity can be engineered by modifying motif D, and that viruses encoding these variants may serve as live, attenuated vaccine candidates.
In picornaviruses, similar interactions as those highlighted by the current MD simulations may govern the motion of motif D in other polymerases and hence be important for this fidelity checkpoint. Although Thr362 is not conserved among picornoviral polymerases, the amino acid equivalent to this position is also polar in coxsackievirus B3 (Cys363), foot-and-mouth disease virus (Lys372), and human rhinovirus serotype 16 (HRV16; Glu361). Modifying these positions to something more non-polar may allow them to interact with the well conserved motif D Phe and lead to similar structural and/or dynamic changes that we have proposed for the T362I PV RdRp. The non-polar substitutions may also disrupt the local environment in a similar way due to similarities in structure. For example, in HRV16 RdRp, there is an electrostatic interaction between Asp228/Lys363 comparable to the Lys228/Glu364 in PV RdRp. Asn369 in HRV16 RdRp is comparable to Asn370 in PV RdRp. The substitution of Glu361 in HRV16 RdRp may again lead to an interaction between Lys363 and Asn369 that would lower RdRp fidelity.

One intriguing feature of PV RdRp is that amino acid substitutions remote from the active site can alter polymerase catalysis and fidelity. These findings suggest that there are long-range interaction networks for the function and fidelity of the RdRp. The T362I substitution also leads to long-range structural and dynamic changes, including a reduction in nanosecond dynamics of Asp53, the location of another Sabin amino acid substitution. Interestingly, previous MD simulations of WT RdRp have also indicated that the motions of the α-helix containing Tyr73 are anti-correlated with the motions of motif D. These correlations between regions encompassing the Sabin substitution may suggest that there is cooperativity among the Sabin sites. These Sabin amino acid substitutions contribute to additional RdRp functions: The Y73H and the T362I PV mutants display temperature sensitivity phenotypes. The Y73H substitution also interferes with the initiation of genome replication. The K250E substitution may interfere with protein structure and/or RNA binding affinity through the drastic charge inversion. Multiple
functions of the Sabin substitutions lead to complex mechanisms of viral attenuation in the Sabin vaccine strain.

2.5 References

CHAPTER 3

SELECTIVE PRESSURE TO RE-ESTABLISH RNA-DEPENDENT RNA POLYMERASE FIDELITY IN THE SABIN VACCINE STRAIN


In this Chapter, I was responsible for plasmid construction and kinetic assays for all the Sabin derived RdRp variants. Derek Musser and I overexpressed and purified the RdRp proteins. Drs. David Boehr and Xiaorong Yang performed NMR experiments for D53N/Y73H/T362I RdRp and Sabin RdRp complexes, respectively. Dr. Cheri Lee performed cell culture assays.

3.1 Introduction

Positive strand RNA viruses are common causative agents of human disease including the common cold, viral myocarditis, viral encephalitis, hepatitis, and paralytic poliomyelitis. Paralytic poliomyelitis caused by poliovirus (PV) has been successfully controlled by the introduction of both the live, attenuated Sabin vaccine and the inactivated Salk vaccine in the late 1950s to early 1960s. PV belongs to the enterovirus genus in the Picornaviridae family and has three distinct serotypes, those being PV1 (i.e. Mahoney strain), PV2 (i.e. Lansing strain), and PV3 (i.e. Leon strain). The virion is comprised of an icosahedral protein shell and a single-stranded RNA genome with positive polarity and length around 7.5 k base pairs (bp). The viral genome only has one large open reading frame encoding both the structural proteins and the functional proteins including the RNA-dependent RNA polymerase (RdRp).
protein genome’ (VPg), which acts as a primer during RNA synthesis, and a polyadenylated tail are covalently linked to the 5’ and 3’ ends of the viral genome respectively. In the 5’-untranslated region, there is an internal ribosomal entry site (IRES) important for initiating protein translation. When viruses infect a host cell, genome replication completely depends on the virally encoded RdRp. The RdRp first transcribes a positive strand RNA into the complimentary negative strand RNA. The intermediate negative strand RNA then acts as the template to synthesize additional positive strand viral genomes. Therefore, the RdRp is a core enzyme during RNA synthesis and viral infection.

The Sabin vaccine formulation contains three attenuated strains (Sabin 1, 2, and 3) corresponding to the three PV serotypes. Among the three attenuated strains, Sabin 1 is most studied partially due to its safety credits for vaccinations. The Sabin 1 strain is different from its parental PV1 Mahoney strain by 55 point mutations, among which leads to four amino acid substitutions (including D53N, Y73H, K250E, and T362I) in the RdRp.

Unlike Mahoney strains, attenuated Sabin 1 strains only efficiently replicate in the gut and do not efficiently infect the nervous system once they enter into humans. However, the immune responses, initiated by Sabin 1 strains in the intestine, provide superior protection against infections of the virulent parental strains. The molecular mechanism of attenuation for Sabin 1 strains has been suggested to mainly derive from mutations in two areas of the viral genome, the IRES and the 3’ end of the viral genome including both the RdRp and 3’-untranslated regions. The mutations in the IRES reduce the efficacy of protein translation through weakening the binding affinity between cellular ribosomes and the IRES. These mutations are proposed to be the major attenuating factor.

The mutations in the RdRp region lead to viral attenuation and temperature sensitivity. The effects of the RdRp mutations on viral attenuation and temperature sensitivity have been mainly studied through PV recombinants of Mahoney and Sabin 1 strains, and site directed
mutants of both strains\textsuperscript{18, 19, 21, 22}. Neurovirulence tests in multiple animal models (i.e. the monkey model, mouse model encoding the human PV receptor, and mouse model with mice-adapted PV strains) determined that the Y73H mutation with or without additional mutations in 3'-untranslated region contributes to viral attenuation and the temperature sensitive phenotype\textsuperscript{21, 22}. Additionally, plaque morphology analyses in HeLa cells showed that D53N and T362I mutations contribute to temperature sensitive phenotypes\textsuperscript{18, 19}.

Besides the temperature sensitivity, other functions of the Sabin 1 RdRp may be changed, which could have some effect on the pathogenicity and/or efficacy of the Sabin 1 vaccine strain\textsuperscript{20}. It is now established that viruses must maintain a certain degree of genetic diversity within their populations to escape the immune response and other environmental challenges\textsuperscript{25-30}. The error-prone RdRp can provide this genetic diversity. However, it should be kept in mind that some degree of RNA replication fidelity must be maintained or viruses may experience extinction due to lethal mutagenesis\textsuperscript{25-30}. Understanding the fidelity of the Sabin 1 RdRp may provide additional insights into viral replication and attenuation on the molecular level.

Our previous kinetic and solution-state nuclear magnetic resonance (NMR) studies demonstrated that the T362I single substitution lowers enzyme fidelity through shifting the conformational equilibrium of motif D towards a more catalytically competent “closed” state in which the general acid Lys359 is poised to catalyze the phosphodiester reaction, even in the presence of incorrect nucleotide\textsuperscript{31}. Molecular dynamics (MD) simulations also indicated that the T362I substitution changes the conformation of motif D and also, impacts more distant residues, including the other substitution sites in the Sabin 1 RdRp\textsuperscript{31}. For example, The T362I substitution constrains the dynamics of Asp53 located in the “fingers” subdomain\textsuperscript{31}. The MD simulation of the WT enzyme also indicated that the motions of motif D, which contains Thr362, are anti-correlated with respect to the motions experienced by the α-helix containing Tyr73\textsuperscript{32}. The helices containing Tyr73 and Lys250 are also closely packed together. Here, we hypothesized that the
T362I substitution interacts with the D53N, Y73H, and K250E substitutions to perturb or rescue the fidelity and the function of the Sabin 1 RdRp.

In this Chapter, we studied single nucleotide incorporation kinetics of single and multiple variants derived from the Sabin 1 RdRp. The single and multiple variants did not significantly impair the assembly and the stability of the enzyme complexes. The Sabin 1 RdRp had the lowest sugar fidelity but had similar nucleobase fidelity as WT enzyme. The contributions of the four amino acid substitutions to the function of the Sabin 1 RdRp were not simply additive. For example, the Y73H substitution only changed fidelity when it coexisted with the D53N and T362I substitutions. NMR experiments showed that conformational changes of some structures (i.e. N terminus or motif A) contributed to the separation of sugar and nucleobase fidelity of the Sabin 1 RdRp. These results suggested that WT-levels of nucleobase fidelity may be essential for the fitness of Sabin 1 vaccine strains, and cooperative interactions among Sabin 1 substitutions help to regulate RdRp fidelity.

3.2 Methods

3.2.1 Materials

[γ-32P]ATP and [α-32P]UTP (>7000Ci/mmol) were from VWR-MP Biomedical. Nucleoside 5’-triphosphates and 2’-deoxynucleoside 5’-triphosphates were ultrapure solutions from GE Healthcare. 3’-Deoxyadenosine 5’-triphosphate (cordycepin) was from Trilink Biotechnologies. All RNA oligonucleotides were from Dharmaco Research, Inc. (Boulder, CO). T4 polynucleotide kinase was from New England Biolabs, Inc. [methyl-13C]Methionine was from Cambridge Isotope Laboratories. HisPur Ni-NTA resin was from Thermo Scientific. Q-Sepharose fast flow resin was from Amersham Pharmacia Biotech, Inc. Polyethyleneimine-cellulose TLC
plates were from EM Science. The QuickChange site-directed mutagenesis kit was from Stratagene. The plasmid DNA isolation Miniprep kit was from Qiagen. All other reagents were of the highest grade available from Sigma or Fisher.

3.2.3 Plasmid construction

All RdRp mutants including D53N, Y73H, K250E, D53N/T362I, D53N/Y73H/T362I and D53N/Y73H/K250E/T362I (Sabin mutant) were made based on the pSUMO-WT RdRp plasmid with the design of forward and reverse primers following the instructions of the QuickChange site-directed mutagenesis kit. The plasmid products were sent for DNA sequencing at the Nucleic Acid Facility at the Pennsylvania State University. It should be noted that all RdRp variants also contain two additional interface I mutants (L446D and L455D) to reduce protein oligomerization.

3.2.3 Overexpression and protein purification

Overexpression and protein purification were conducted by following the procedures previously described (see Methods in Chapter 2).

3.2.4 Kinetic assays

Kinetic assays including active site titration assays, assembly assays, and dissociation assays were conducted as described previously (see Methods in Chapter 2)\textsuperscript{31, 33}. Briefly, reactions contained 50 mM HEPES pH 7.5, 10 mM 2-mercaptoethanol (BME), 5 mM MgCl\textsubscript{2}, and 60 μM ZnCl\textsubscript{2}. Stopped flow experiments and the benchtop assays for nucleotide incorporation were
conducted as described previously\textsuperscript{31,34,35}. Reactions were incubated at 30 °C and were quenched by the addition of equal volume of EDTA to a final concentration of 25 mM. Specific concentrations of RdRp, sym/subU RNA, and nucleotides are indicated in the corresponding figure legend.

3.2.5 Determination of kinetic constants ($K_{d,\text{app}}$ and $k_{\text{pol}}$) for nucleotide incorporation catalyzed by RdRp

Data analysis of kinetic constants was conducted by following procedures and equations previously described (see Methods in Chapter 2). Briefly, quenched product samples from kinetic assays were analyzed by 23% highly cross-linked denaturing polyacrylamide gel. Gels were visualized through PhosphorImager and quantified through the ImageQuant software (Molecular Dynamics). Data were fit into different curves using Kaleidagraph (Synergy Software, Reading, PA).

3.2.6 NMR sample preparation and spectroscopy

NMR sample preparation and collection of $^1$H - $^{13}$C heteronuclear single quantum coherence (HSQC) spectra were described previously (see Methods in Chapter 2). Most of the experiments were performed on a Bruker Avance III 600 MHz spectrometer equipped with a 5 mm inverse detection triple resonance ($^1$H/$^{13}$C/$^{15}$N) single axis gradient TCI cryoprobe. Final concentrations of RdRp and sym/subU RNA were 250 and 1000 µM respectively. Final concentrations of nucleotide varied (i.e. 4 mM UTP, 8 mM 2’dUTP, or 16 mM ATP).
3.3 Results

3.3.1 The three single variants D53N, Y73H, and K250E RdRps, and Sabin RdRp do not change assembly and stability of enzyme complexes

The four substitutions in Sabin RdRp may affect assembly and stability of enzyme complexes (Figure 3-1). The D53N substitution is located near the N terminus of RdRp and perturbations on the N terminus decrease enzyme activity (Figure 3-1)

The K250E substitution is close to the highly conserved motif B (Figure 3-1), which contains residues important for sugar selection. Structural dynamics of the loop at the N terminus of motif B is also essential for RNA binding.

I have already shown that the T362I substitution does not weaken the interactions between enzyme and RNA compared to WT RdRp (see Chapter 2). Here, I conducted the same assembly and dissociation assays for the other three single variants including D53N, Y73H, and K250E RdRps, and Sabin 1 RdRp (which contains all four Sabin substitutions together).

The rate and the yield of the formation of active RdRp-RNA complexes for D53N, Y73H, K250E, and Sabin 1 RdRps were highly similar to that of WT RdRp (Figure 3-2). The dissociation rates of the formed RdRp-RNA complexes for D53N, Y73H, K250E, and Sabin 1 RdRps were also very slow and comparable to that of WT RdRp (Figure 3-2). The dissociation rates ranged from $2 \times 10^{-4}$ to $4 \times 10^{-4}$ s$^{-1}$ (Figure 3-2). Altogether, the stability of the complexes of the RdRp variants ensures that I could proceed with single nucleotide incorporation assays.
The structure of PV RdRp indicating the four Sabin substitution sites (PDB 1RA6). The front and back views are indicated in A and B respectively. PV RdRp contains “palm”, “thumb”, and “fingers” subdomains and seven highly conserved motifs including motif A (green), B (hot pink), C (orange), D (cyan), E (blue), F (yellow), and G (red). The Sabin RdRp contains four substitutions including D53N, Y73H, K250E, and T362I (whose substitution sites are colored in purple) compared to WT RdRp. The general acid Lys359 is colored in cyan. The loop, where both D53N and Y73H substitutions occur, is colored in gold. This loop also contains Gly64. The substitution of Gly64 to Ser leads to an increase in enzyme fidelity. This figure was created using Chimera.
Figure 3-2. Sabin substitutions do not affect the assembly and the stability of RdRp-RNA binary complexes. Reactions contained 50 mM HEPES 7.5, 10 mM BME, 5 mM MgCl$_2$, and 60 µM ZnCl$_2$. A. Sequence of sym/subU RNA. B. Reaction scheme for the assembly assay. RNA and ATP were preincubated for 5 min. Then, reactions were initiated by addition of RdRp and were quenched by addition of EDTA to a final concentration of 25 mM at different time points. Final concentrations of RdRp, sym/subU RNA, and ATP were 1, 1, and 500 µM respectively. C. Overlapping of the formation of the assembled complexes for WT, Sabin, D53N, Y73H, K250E, and T362I RdRps suggested that these Sabin substitutions do not affect enzyme assembly. D. Reaction scheme for the dissociation assay. RdRp-RNA binary complexes were incubated for 90s followed by addition of trap. At different time points, ATP was added and was allowed to incorporate for 30s after which reactions were quenched by the addition of EDTA to a final concentration of 25 mM. Final concentrations of RdRp, sym/subU, and ATP were 1, 0.1, and 500 µM respectively. E. Overlapping of the dissociation of the enzyme complexes for WT, Sabin, D53N, Y73H, and T362I RdRps. The lines represent fits of the data to a single exponential function. The calculated dissociation rates ($k_{\text{dis}}$) are in panel F. Data for T362I RdRp was obtained from previous studies (see Chapter 2).
3.3.2 D53N, Y73H, and K250E RdRps have different sugar and nucleobase fidelities

Single nucleotide incorporation assays were used to evaluate all the Sabin-derived RdRp variants. In these assays, two kinetic parameters are obtained, those being the maximal rate constant $k_{pol}$ and the apparent dissociation constant $K_{d,\text{app}}$ for nucleotide$^{35}$. The catalytic efficiency for each type of nucleotide incorporation equals the ratio of $k_{pol}$ over $K_{d,\text{app}}^{35}$. The fidelity, that is the ability of RdRp to select a cognate against a noncognate nucleotide, is defined as the ratio of the catalytic efficiency of the cognate nucleotide incorporation over that of noncognate nucleotide incorporation (i.e. $(k_{pol}/K_{d,\text{app}})^{\text{cognate}}/(k_{pol}/K_{d,\text{app}})^{\text{noncognate}}$)$^{35}$. For all the Sabin variants, the apparent dissociation constants were similar to WT RdRp (Table 3-1), which suggests that the ground-state binding of nucleotide does not significantly contribute to changes in enzyme fidelity for these variants. Therefore, I used a simplified fidelity equation in this Chapter, which is the ratio of the maximal rate constant of the cognate nucleotide over that of the noncognate nucleotide (i.e. $(k_{pol})^{\text{cognate}}/(k_{pol})^{\text{noncognate}}$). Previous kinetic results showed that T362I RdRp is a low fidelity variant with respect to both sugar and nucleobase groups. This decrease in the fidelity derives from an increase in the catalytic rate for incorporation of noncognate nucleotides without significantly changing the catalytic rate for the incorporation of cognate nucleotide (i.e. compared to WT RdRp). The fidelity of T362I RdRp decreases up to 1.5 fold for both the sugar and nucleobase selection. I have previously shown that the decrease in the fidelity by this single substitution is template independent (see Chapter 2).

To examine kinetic effects of the other three Sabin substitutions, I used single nucleotide incorporation assays to evaluate the D53N, Y73H, and K250E single variants. D53N RdRp had similar single nucleotide incorporation rates as WT RdRp for both the cognate nucleotide and 2’deoxynucleotide, which indicates that the D53N substitution did not change the sugar fidelity (Table 3-1). However, D53N RdRp had a much slower rate than WT RdRp for noncognate
nucleobase incorporation (Table 3-1). Thus, the nucleobase fidelity of D53N RdRp increased by 1.6 fold. The Y73H substitution did not change the fidelity for either sugar or nucleobase selections although the nucleotide incorporation rates were slightly slower than those of WT RdRp (Table 3-1). K250E RdRp had higher nucleotide incorporation rates, for both cognate nucleotide and 2’deoxynucleotide compared to WT RdRp (Table 3-1). Consequently, its sugar fidelity decreased slightly (Table 3-1). By contrast, the K250E substitution led to an increase in the nucleobase fidelity by 1.4 fold (Table 3-1). Altogether, the D53N, Y73H, and K250E substitutions led to small alterations in the rates and fidelity of nucleotide incorporation.
Table 3-1. Comparison of the nucleotide incorporation rates for WT, Sabin, and Sabin derived RdRps reveals that they all have slightly different catalytic parameters and/or fidelity.

<table>
<thead>
<tr>
<th>Variant</th>
<th>nucleotide</th>
<th>k\textsubscript{pol} (s\textsuperscript{-1})</th>
<th>K\textsubscript{d,app} (µM)</th>
<th>k\textsubscript{pol}/K\textsubscript{d,app} (µM\textsuperscript{-1} s\textsuperscript{-1})</th>
<th>k\textsubscript{pol,correct}/k\textsubscript{pol, incorrect}</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td></td>
<td>59±1</td>
<td>36±2</td>
<td>1.6</td>
<td>-</td>
</tr>
<tr>
<td>D53N</td>
<td>ATP</td>
<td>62±1</td>
<td>39±2</td>
<td>1.6</td>
<td>-</td>
</tr>
<tr>
<td>Y73H</td>
<td>ATP</td>
<td>46±1</td>
<td>40±2</td>
<td>1.2</td>
<td>-</td>
</tr>
<tr>
<td>K250E</td>
<td>ATP</td>
<td>76±1</td>
<td>58±3</td>
<td>1.3</td>
<td>-</td>
</tr>
<tr>
<td>T362I\textsuperscript{a}</td>
<td>ATP</td>
<td>67±1</td>
<td>33±2</td>
<td>2.0</td>
<td>-</td>
</tr>
<tr>
<td>D53N/T362I</td>
<td>ATP</td>
<td>72±1</td>
<td>44±2</td>
<td>1.6</td>
<td>-</td>
</tr>
<tr>
<td>D53N/Y73H/T362I</td>
<td>ATP</td>
<td>84±2</td>
<td>47±4</td>
<td>1.8</td>
<td>-</td>
</tr>
<tr>
<td>Sabin</td>
<td>ATP</td>
<td>74±1</td>
<td>39±2</td>
<td>1.9</td>
<td>-</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variant</th>
<th>nucleotide</th>
<th>k\textsubscript{pol} (s\textsuperscript{-1})</th>
<th>K\textsubscript{d,app} (µM)</th>
<th>k\textsubscript{pol}/K\textsubscript{d,app} (µM\textsuperscript{-1} s\textsuperscript{-1})</th>
<th>k\textsubscript{pol,correct}/k\textsubscript{pol, incorrect}</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>2’dATP</td>
<td>0.89±0.01</td>
<td>134±4</td>
<td>6.7 x 10\textsuperscript{-3}</td>
<td>70</td>
</tr>
<tr>
<td>D53N</td>
<td>2’dATP</td>
<td>0.93±0.02</td>
<td>101±9</td>
<td>9.3 x 10\textsuperscript{-3}</td>
<td>70</td>
</tr>
<tr>
<td>Y73H</td>
<td>2’dATP</td>
<td>0.73±0.01</td>
<td>117±5</td>
<td>6.2 x 10\textsuperscript{-3}</td>
<td>60</td>
</tr>
<tr>
<td>K250E</td>
<td>2’dATP</td>
<td>1.4±0.1</td>
<td>174±6</td>
<td>8.0 x 10\textsuperscript{-3}</td>
<td>50</td>
</tr>
<tr>
<td>T362I\textsuperscript{a}</td>
<td>2’dATP</td>
<td>1.5±0.1</td>
<td>112±4</td>
<td>1.3 x 10\textsuperscript{-2}</td>
<td>40</td>
</tr>
<tr>
<td>D53N/T362I</td>
<td>2’dATP</td>
<td>1.5±0.1</td>
<td>132±6</td>
<td>1.1 x 10\textsuperscript{-2}</td>
<td>50</td>
</tr>
<tr>
<td>D53N/Y73H/T362I</td>
<td>2’dATP</td>
<td>1.4±0.1</td>
<td>101±9</td>
<td>1.4 x 10\textsuperscript{-2}</td>
<td>60</td>
</tr>
<tr>
<td>Sabin</td>
<td>2’dATP</td>
<td>2.0±0.1</td>
<td>145±4</td>
<td>1.4 x 10\textsuperscript{-2}</td>
<td>40</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variant</th>
<th>nucleotide</th>
<th>k\textsubscript{pol} (s\textsuperscript{-1})</th>
<th>K\textsubscript{d,app} (µM)</th>
<th>k\textsubscript{pol}/K\textsubscript{d,app} (µM\textsuperscript{-1} s\textsuperscript{-1})</th>
<th>k\textsubscript{pol,correct}/k\textsubscript{pol, incorrect}</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>GTP</td>
<td>0.011±0.001</td>
<td>142±15</td>
<td>7.7 x 10\textsuperscript{-5}</td>
<td>5400</td>
</tr>
<tr>
<td>D53N</td>
<td>GTP</td>
<td>0.0073±0.0008</td>
<td>91±35</td>
<td>8.0 x 10\textsuperscript{-5}</td>
<td>8400</td>
</tr>
<tr>
<td>Y73H</td>
<td>GTP</td>
<td>0.0082±0.0005</td>
<td>154±31</td>
<td>5.3 x 10\textsuperscript{-5}</td>
<td>5600</td>
</tr>
<tr>
<td>K250E</td>
<td>GTP</td>
<td>0.0099±0.0008</td>
<td>160±41</td>
<td>6.1 x 10\textsuperscript{-5}</td>
<td>7700</td>
</tr>
<tr>
<td>T362I\textsuperscript{a}</td>
<td>GTP</td>
<td>0.018±0.001</td>
<td>149±25</td>
<td>1.2 x 10\textsuperscript{-4}</td>
<td>3700</td>
</tr>
<tr>
<td>D53N/T362I</td>
<td>GTP</td>
<td>0.012±0.001</td>
<td>115±21</td>
<td>1.0 x 10\textsuperscript{-4}</td>
<td>6000</td>
</tr>
<tr>
<td>D53N/Y73H/T362I</td>
<td>GTP</td>
<td>0.011±0.001</td>
<td>128±12</td>
<td>8.6 x 10\textsuperscript{-5}</td>
<td>7600</td>
</tr>
<tr>
<td>Sabin</td>
<td>GTP</td>
<td>0.013±0.001</td>
<td>127±16</td>
<td>1.0 x 10\textsuperscript{-4}</td>
<td>5700</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variant</th>
<th>nucleotide</th>
<th>k\textsubscript{pol} (s\textsuperscript{-1})</th>
<th>K\textsubscript{d,app} (µM)</th>
<th>k\textsubscript{pol}/K\textsubscript{d,app} (µM\textsuperscript{-1} s\textsuperscript{-1})</th>
<th>k\textsubscript{pol,correct}/k\textsubscript{pol, incorrect}</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>2’C-methyl</td>
<td>1.2±0.1</td>
<td>160±9</td>
<td>7.5 x 10\textsuperscript{-3}</td>
<td>50</td>
</tr>
<tr>
<td>Sabin</td>
<td>ATP</td>
<td>1.9±0.1</td>
<td>129±5</td>
<td>1.5 x 10\textsuperscript{-2}</td>
<td>40</td>
</tr>
</tbody>
</table>

\textsuperscript{a} Data is obtained from previous studies (see Chapter 2).
3.3.3 Sabin 1 RdRp has lower sugar selectivity but similar nucleobase selectivity as WT enzyme.

I have shown that all four Sabin amino acid substitutions affect catalysis and fidelity of RdRp. To determine the impacts of all four amino acid substitutions in the Sabin 1 RdRp, I conducted single nucleotide incorporation assays using both cognate and noncognate nucleotides. When cognate nucleotide is incorporated, Sabin 1 RdRp had a higher rate than both WT and T362I RdRps (Table 3-1). Sabin 1 RdRp, like T362I RdRp, had a faster rate for 2’deoxynucleotide incorporation than WT RdRp (Figure 3-3, Table 3-1). As such, Sabin 1 and T362I RdRps had the lowest sugar fidelity among all the Sabin derived variants (Table 3-1). By contrast, the nucleobase fidelity of Sabin 1 RdRp was very similar to that of WT enzyme (Table 3-1).

The thermodynamic effects were calculated for all of the Sabin derived variants to further investigate the relationship among the four Sabin substitutions. Similar to the methods of Fersht and Mildvan, we determined ΔΔG values where ΔΔG_X = RT ln((k_{pol}/K_{d,app})_X/(k_{pol}/K_{d,app})_{WT}) where X is a particular enzyme variant (in this analysis, when a variant has a higher catalytic efficiency (k_{pol}/K_{d,app}), the ΔΔG_X will be positive). By comparing the ΔΔG_X value of a multiple variant with the theoretical thermodynamic effect that is equal to the sum of the ΔΔG_X values of all the variants whose substitutions make up the multiple variant, cooperativity among these substitutions can be defined. For example, if a calculated ΔΔG_X is greater than its theoretical thermodynamic effect, additional stabilization energy in RdRp variant was generated. The additional energy due to multiple substitutions is referred to as positive cooperativity. On the other hand, if a calculated ΔΔG_X is less than its theoretical thermodynamic effect, the deficit of energy due to multiple substitutions is referred to as negative cooperativity. The thermodynamic analysis indicated that the four substitutions were not simply additive for the Sabin 1 RdRp, but instead, the four substitutions appeared to be interdependent, in that ΔΔG_{Sabin} did not equal the...
sum of the thermodynamic contributions of the four individual amino acid substitutions (i.e. \( \Delta \Delta G_{D53N} + \Delta \Delta G_{Y73H} + \Delta \Delta G_{K250E} + \Delta \Delta G_{T362I} \)) (Figure 3-4). For example, in the noncognate nucleobase incorporation, the four substitutions were positively cooperative (\( \Delta \Delta G_{Sabin} > \Delta \Delta G_{D53N} + \Delta \Delta G_{Y73H} + \Delta \Delta G_{K250E} + \Delta \Delta G_{T362I} \)) (Figure 3-4). However, the extent of the positive cooperativity was small; the difference between \( \Delta \Delta G_{Sabin} \) and \( \Delta \Delta G_{D53N} + \Delta \Delta G_{Y73H} + \Delta \Delta G_{K250E} + \Delta \Delta G_{T362I} \) was only around 0.2 ~ 0.3 kcal/mol (Figure 3-4).
Figure 3-3. Sabin 1 RdRp has the lowest sugar fidelity but has nucleobase fidelity similar to WT level. Reactions contained 50 mM HEPES 7.5, 10 mM BME, 5 mM MgCl₂, and 60 µM ZnCl₂. Final concentrations of RdRp and sym/subU RNA were 0.5 and 0.5 µM respectively. A. Reaction scheme for a stopped flow experiment. RdRp and sym/subU RNA were preincubated for 2min followed by the addition of an equal volume of nucleotide solution to initiate reactions. Product formation was monitored by changes in fluorescence. B, C. Comparison of the cognate ATP (B) and noncognate 2'dATP (C) incorporation rates among WT, T362I, and Sabin 1 RdRps. The lines represent fits of the data to a hyperbola function. The obtained kinetic parameters are listed in Table 3-1. D. Reaction scheme of a benchtop assay. RdRp and sym/subU RNA were preincubated for 5min followed by the addition of an equal volume of GTP solution to initiate reactions. Reactions were quenched by the addition of EDTA to a final concentration of 25 mM. E. Comparison of the noncognate GTP incorporation rates among WT, T362I, and Sabin 1 RdRps. The lines represent fits of the data to a hyperbola function. The obtained kinetic parameters are listed in Table 3-1.
3.3.4 Cooperative interactions among the D53N, Y73H, K250E, and T362I substitutions are small

Previous MD simulations have suggested that the T362I substitution affects protein structural dynamics near the Asp53 site\textsuperscript{31}. To further evaluate the relationship between these two substitution sites, I constructed the double variant D53N/T362I RdRp. In single nucleotide incorporation assays, D53N/T362I RdRp incorporated the cognate nucleotide and the 2’deoxy nucleotide faster than WT RdRp, but incorporated nucleotide with noncognate nucleobase near a WT level (Table 3-1). The sugar fidelity for the double variant D53N/T362I was similar to the fidelity of T362I RdRp, but the nucleobase fidelity was between that of T362I RdRp and D53N RdRp (Table 3-1). We also calculated the thermodynamic effects of the double variant and compared it to those of the two single variants (D53N RdRp and T362I RdRp). The results indicated that the D53N and T362I substitutions were negatively cooperative for all types of nucleotide incorporations (i.e. \( \Delta \Delta G_{D53N/T362I} < \Delta \Delta G_{D53N} + \Delta \Delta G_{T362I} \)) (Figure 3-4).

The Y73H substitution may also affect the D53N and T362I substitutions. To evaluate the relationship among the three substitutions, I constructed the triple variant D53N/Y73H/T362I RdRp. Another reason to study the triple variant is that PV carrying the four Sabin mutations in the background of the Mahoney strain is not stable during passages in HeLa cells. The mutation K250E reverts back to Lys (data not shown). Thus, it was interesting to know if D53N/Y73H/T362I RdRp could replace Sabin 1 RdRp in studies of viral fitness in cell culture. Unfortunately, the kinetic parameters of the triple variant were substantially different from those of Sabin 1 RdRp such that further cell culture assays were not attempted (Table 3-1).

With the noncognate nucleobase, the thermodynamic effects of the triple variant nearly equaled the sum of the effects of the three single variants (i.e. D53N, Y73H, and T362I RdRps). The “independence” of the three substitutions in the triple variant may be a coincidence of cancellation of the positive and negative cooperativity among the three substitutions, since the
T362I and D53N substitutions contribute to negative cooperativity (as observed in the studies of D53N/T362I RdRp). The Y73H substitution may have positive cooperativity with the D53N/T362I substitution. To test this hypothesis, I calculated thermodynamic cycles of Y73H, D53N/T362I, and the triple variants and thermodynamic cycles of K250E, D53N/Y73H/T362I, and Sabin 1 variants (Figure 3-4). As I expected, the results indicated that there was a positive cooperativity between the Y73H substitution and the D53N/T362I substitution ($\Delta\Delta G_{D53N/Y73H/T362I} > \Delta\Delta G_{Y73H} + \Delta\Delta G_{D53N/T362I}$) for both noncognate nucleobase and 2’deoxy nucleotide incorporations (Figure 3-4). The extent of the positive cooperativity in the 2’deoxy nucleotide incorporation was greater than that in the noncognate nucleobase incorporation, though cooperativity was small (the difference between $\Delta\Delta G_{D53N/Y73H/T362I}$ and $(\Delta\Delta G_{D53N/T362I} + \Delta\Delta G_{Y73H})$ is around 0.1 to 0.3 kcal/mol) (Figure 3-4). The K250E substitution also had small cooperativity with the D53N/Y73H/T362I substitution (~ 0.1 to 0.2 kcal/mol) for all types of nucleotide incorporations (Figure 3-4).
Figure 3-4. There is wide range of cooperativity among D53N, Y73H, K250E, and T362I substitutions in Sabin 1 RdRp and Sabin derived RdRps. Thermodynamic effect ΔΔG is calculated by the following equation $\Delta \Delta G_x = RT \ln(\frac{K_{d,app}}{k_{pol}})_x/\frac{K_{d,app}}{k_{pol}}_{WT}$, where $x$ represents a particular variant. A. The comparison of ΔΔG between Sabin 1 variant and sum of D53N, Y73H, K250E, and T362I single variants (whose substitutions together consist of the Sabin variant), between D53N/Y73H/T362I and sum of D53N, Y73H, and T362I single variants, and between D53N/T362I and sum of D53N and T362I single variants for ATP, 2’dATP, and GTP incorporations. Positive cooperativity represents a greater ΔΔG of a multiple variant than that of the sum of the corresponding single variants. Thermodynamic cycle is calculated for ATP (B), 2’dATP (C), and GTP (D) incorporations. On the left side is the thermodynamic cycle of Y73H, D53N/Y73H, and D53N/Y73H/K250E variants. On the right side is the thermodynamic cycle of K250E, D53N/Y73H/K250E, and Sabin 1 variants. The unit is cal/mol. E, F. Comparison of the sugar (E) and nucleobase (F) fidelities among WT, T362I, D53N/Y73H, D53N/Y73H/K250E, and Sabin 1 variants.
3.3.5 Other conformational changes compensate for the changes of motif D to regulate the fidelity of Sabin RdRp

Our previous NMR experiments have suggested that the T362I variant has lower fidelity because this substitution allows the motif D active-site loop to more readily fluctuate into a “closed” state even with noncognate nucleotide\textsuperscript{31}. To determine if similar structural dynamic changes occur for other variants, we conducted $[\text{Methyl}^{13}\text{C}]$Met NMR experiments for D53N/Y73H/K250E triple variant and Sabin 1 variant and compared the NMR spectra with those for WT and T362I RdRps (see Chapter 2). The NMR spectra for the cognate RdRp-RNA-nucleotide ternary complexes were similar among all variants (Figure 3-5). However, NMR spectra for the RdRp-RNA-2’dUTP ternary complexes were substantially different. The most informative change is likely that of the Met354 resonance, which we believe reports on the conformational state of the motif D active-site loop. In the RdRp-RNA-2’dUTP ternary complex, WT RdRp stays primarily in the nonproductive “open” state, but the T362I variant fluctuates between the “closed” and “open” states. In the triple variant D53N/Y73H/T362I, the enzyme is primarily in the nonproductive “open” state, just like WT enzyme. This finding was surprising since the triple variant contained the T362I substitution and had the same catalytic rate as T362I variant for the 2’deoxyribonucleotide incorporation (Table 3-1). In the Sabin 1 variant, the enzyme was able to fluctuate into the “closed” state, but not nearly to the extent of the T362I variant (Figure 3-5). Considering that the Sabin 1 RdRp has lower sugar fidelity than T362I RdRp, these results suggest that there must be other structural/dynamic changes that govern the fidelity changes of the Sabin 1 RdRp. There may be conformational changes at the N terminus and/or motif A that are also important for nucleotide selection in the Sabin 1 RdRp.
Figure 3-5. Motif D cooperates with other structural elements to account for RdRp fidelity of Sabin derived RdRp variants. Comparison of the [Methyl\(^{13}\)C]Met\(^{1\text{H},13}\)C HSQC spectra between the cognate RdRp-RNA-UTP and noncognate RdRp-RNA-2’dUTP ternary complexes for WT (A), T362I (B), D53N/Y73H/T362I (C), and Sabin 1 RdRps. WT enzyme strictly distinguishes cognate and noncognate nucleotides by assigning motif D to the “closed” and “open” states respectively. T362I variant promiscuously recognizes noncognate nucleotide by allowing motif D to partially fluctuate into the “closed” state, which leads to a lower fidelity for this variant. Neither D53N/Y73H/T362I triple variant nor Sabin 1 RdRp allows motif D to populate the “closed” state to the same extent as the T362I variant for noncognate nucleotide, which suggests the involvement of other structural elements in determining their fidelities. (D). Spectra are collected at 293K with 250 µM RdRp, 1 mM sym/subU RNA, and 4 mM UTP or 8 mM 2’dUTP.
3.3.6 The Sabin RdRp is more susceptible to 2’ modified nucleotides

The incorporation of 2’C-methyl nucleotide was also evaluated for both WT and Sabin 1 RdRps. It is well established that nucleotide analogs can be very efficient antiviral drugs to treat viral infectious diseases. Results indicated that the Sabin 1 variant more readily incorporated this analog than WT enzyme, suggesting that Sabin 1 derived strains may be more sensitive to treatment with 2’ modified nucleotide analogs (Table 3-1).

3.4 Discussion

Previous studies indicated that the T362I single substitution lowers polymerase fidelity by shifting the motif D active-site loop more towards the catalytically competent “closed” state. Here, I demonstrated that all four Sabin 1 amino acid substitutions affect the fidelity and catalysis of RdRp (Table 3-1). When the four substitutions were added together, they cooperatively contribute to the fidelity of the Sabin 1 variant, such that Sabin 1 RdRp had nucleobase fidelity similar to WT enzyme but had the lowest sugar fidelity among all the variants tested here (Figure 3-3, 3-4). NMR studies suggested that the changes in fidelity among RdRp variants were a result of the changes in multiple structural elements (i.e. motif D, the N terminus, and motif B) (Figure 3-5). The NMR studies may underestimate contributions of other structural elements to the enzyme fidelity.

The selection of viral mutants that have nucleobase replication fidelity similar to WT PV may be important for the fitness of the Sabin 1 vaccine strain. Once a virus enters a cell, viral genomes are synthesized in the cytoplasm of the infected cell via the RdRp. For every cycle of nucleotide incorporation in the viral genome, the RdRp faces the challenge of selecting for the cognate nucleotide over that of 2’deoxynucleotide and ribonucleotide with noncognate
nucleobase. The selection of the noncognate nucleobase is likely more stringent since in the cytoplasm the concentration of ribonucleotides (mM) is much greater than that of deoxynucleotides (µM). The low concentration of deoxynucleotides provides a sufficient discrepancy for the RdRp. By contrast, the accumulation of an excess amount of noncognate nucleobases may result in loss of genetic information and extinction of the virus, which viruses avoid vigorously. As such, there is likely more selective pressure to maintain favorable nucleobase selection rather than sugar fidelity.

The low sugar fidelity of Sabin 1 RdRp likely has implications for patients with vaccine associated paralytic poliomyelitis (VAPP). There is still the controversy over the safety of the Sabin vaccine, and there are ongoing discussions to replace the Sabin vaccine with Salk vaccine, even in developing countries. One major concern is that one can get poliomyelitis by being inoculated by the Sabin vaccine, though the chance of VAPP is admittedly very low. Studies of PV in patients with VAPP indicate that His73 tends to revert to Tyr in the RdRp area of Sabin 1 strain. We have shown that Glu250 in Sabin 1 RdRp in the background of Mahoney strain tends to revert to Lys. Regardless of these reversions, the remaining RdRp variants have low sugar fidelity and should be sensitive to the 2’ modified nucleotide analogs, like sofosbuvir that is now a blockbuster drug for the treatment of Hepatitis C.

Our NMR experiments not only represent the dynamics of RdRp but also suggest that there is a communication among multiple functional elements including motif D, motif A, and the N terminus. For example, in the [Methyl-13C]Met NMR spectra of RdRp-RNA-2’dUTP ternary complex, the triple variant did not shift motif D to the “closed” state compared to T362I variant, but gave rise to a similar catalytic rate of 2’deoxynucleotide incorporation as the T362I variant. This finding suggests that there is communication between motif D and the two substitutions D53N and Y73H. Our previous MD simulations of the T362I variant indicate that the T362I substitution not only affects local interactions of residues on motif D but also perturbs
interactions of residues (i.e. Asp238) in motif A (see Chapter 2)\textsuperscript{31}. Interestingly, the D53N and Y73H substitutions are close to a loop (from Val63 to Asp71) on which Gly64 and Asn65 are located. These two substitutions likely perturb motif A through structural dynamic changes in Gly64 and Asn65 that are involved in a hydrogen bond network that also includes Ala239 and Leu241 on motif A and Gly1 on the N terminus (Figure 3-6). Therefore, the D53N and Y73H substitutions may drive motif D to the “open” state via perturbing motif A in the triple variant. Motif A may synergistically shift more to the “closed” state to compensate for the changes in motif D in order to maintain the rate of 2’deoxynucleotide incorporation at T362I RdRp type levels. This hypothesis can be tested by observing different conformations of motif A in NMR spectra of the ternary complexes of T362I and triple variants. However, we would need additional NMR probes in motif A. These observations indicate that Asp53 and Tyr73 may be important coordinators between motif D, motif A, and the N terminus. These observations also suggest that these residues are important in the catalysis and fidelity of RdRp.
Figure 3-6. The D53N and Y73H substitutions are able to affect the conformation of motif D (red) via a hydrogen bond network including motif A (purple), Gly64 (green), Asn65 (green), and the N terminus (orange). The Sabin substitution sites Asp53 and Tyr73 are colored in grey, and the structures that these two substitutions site are located are colored in yellow. The general acid Lys359 (cyan) and another substitution site Thr362 (grey) are located on motif D. The figure was edited using Chimera (PDB 3OL7).
3.5 References


CHAPTER 4

TRIPHOSPHATE RE-ORIENTATION OF THE INCOMING NUCLEOTIDE AS A FIDELITY CHECKPOINT IN THE VIRAL RNA-DEPENDENT RNA POLYMERASE


* The two authors contributed equally to the paper.

In this Chapter, Derek Musser and I were responsible for plasmid construction, protein overexpression, and protein purification for kinetic studies. I conducted all the kinetic assays for WT and other RdRp variants. Dr. Xiaorong Yang was responsible for protein preparation and all of the NMR experiments.

4.1 Introduction

The polymerase superfamily has four categories including RNA-dependent RNA polymerases (RdRp), DNA-dependent DNA polymerases (DdDp), DNA-dependent RNA polymerases (DdRp, transcriptases), and RNA-dependent DNA polymerases (reverse transcriptases (RT))\(^1\)\(^4\). Kinetic studies have suggested that for many of these polymerases, a conformational change prior to phosphoryl transfer is at least partially rate limiting and thus, a very important fidelity checkpoint (Figure 4-1)\(^3\)\(^5\)\(^6\). In the A family of DNA polymerases, crystal structures have shown that when cognate nucleotide binds, the O helix transfers from an “open” to a “closed” state helping to reposition the general acid and align the 3’ OH of the primer RNA with the \(\alpha\)-phosphate moiety of the incoming nucleotide\(^7\)\(^9\). By contrast, when nucleotide with
either noncognate base or sugar binds, the O helix fails to close, but instead forms unstable “ajar”
states, which tend to return to the “open” state and dissociate the noncognate nucleotide. The
significance of the O helix for polymerase fidelity is also confirmed by many studies that
demonstrate that O helix substitutions lead to either high or low fidelity variants. However,
the reasons for the changes in fidelity are complex due to the versatile nature of the O helix.

For example, in bacteriophage T7 RNA polymerase, a swinging motion of the O helix permits
nucleobase pairing and the closure of the active site. A highly conserved Lys631 in the O helix
acts as the general acid. Positively charged residues, such as Arg627 and Arg632 in the O helix,
participate in interactions with the triphosphate of the incoming nucleotide through interacting
with the α and γ phosphate moieties. Tyr639 in the O helix filters out the 2’ H group of the
incoming nucleotide. Additionally, a glutamate residue in the same O helix filters out the 2’
hydroxyl group of the incoming nucleotide through a steric gate mechanism. Perturbations of
any of these functions could lead to changes in fidelity.

There is no O helix in RTs and RdRps. Instead, residues on motifs D and F may be
playing similar roles as those in the O helix. Considering the encircled structure between the
“fingers” and “thumb” subdomains in all RdRps, there are likely smaller conformational changes
upon binding cognate nucleotide. It has been suggested that motif D in RTs and RdRps performs
an analogous function as the O helix in terms of repositioning the general acid (Figure 4-1).

For example, in PV RdRp, Lys359 in motif D is the general acid. Previous NMR experiments and
kinetic studies from our lab and the Cameron lab indicated that motif D fluctuates into the
“closed” state to protonate the β phosphate moiety of the cognate nucleotide. When noncognate
nucleotide binds or noncognate primer terminus exists, motif D stays in the “open” state. A
substitution of the general acid to Arg with a presumed pKa leads to a reduction in catalysis by 10
fold and an increase in fidelity by 5 fold. Additionally, changes in the motif D equilibrium
through substitutions (e.g. T362I substitution derived from the Sabin 1 vaccine) on this motif also
lead to decrease in fidelity\textsuperscript{27}. Motif D in RTs shows similar results as PV RdRp. For example, in HIV-1 RT, a substitution of the general acid Lys220 to Gln leads to an increase in fidelity\textsuperscript{28}.

However, motif D does not contain the positively charged residues that interact with the $\alpha$ and $\gamma$ phosphate moieties of the incoming nucleotide, like the $\alpha$ helix of DNA polymerases. Other positively charged residues that may interact with the triphosphate of the incoming nucleotide reside in the “fingers” subdomains of RTs and RdRps. In RTs, perturbations of Arg (interacting with the $\alpha$ phosphate) and Lys (interacting with the $\gamma$ phosphate) residues by a phosphate modified nucleotide leads to enhanced polymerase fidelity\textsuperscript{29, 30}. In PV RdRp, the corresponding positively charged residues are Arg174 and Lys167, which interact with the $\alpha$ and $\gamma$ phosphate groups respectively (Figure 4-1). Substitutions at Arg174 and Lys167 sites may lead to changes in fidelity due to perturbations in triphosphate binding.

In this Chapter, we found that amino acid substitutions at positions 167 and 174 lead to changes in fidelity through perturbations of the pre-chemistry conformational equilibrium according to kinetic studies. In particular, these substitutions perturb the triphosphate conformations for both cognate and noncognate nucleotide incorporation according to $^{31}$P NMR experiments. We also studied the effects of substitutions in motif F on the conformation of motif D that is also a fidelity checkpoint. Results indicated that motif D does not necessarily coordinate with motif F. Altogether, those findings suggested that reorientation of the triphosphate group of the incoming nucleotide is an important fidelity checkpoint and is aided by residues on motif F.
Figure 4-1. A. Reaction scheme for single nucleotide incorporation. Steps 1 to 5 are nucleotide binding, prechemistry conformational change, phosphoryl transfer, postchemistry conformational change, and pyrophosphate release respectively. Among the five steps, step 2 is an important fidelity checkpoint. B. Crystal structure of PV RdRp (PDB 1RA6). Motifs D and F are colored in cyan and hot pink respectively. The general acid Lys359, Lys167, and Arg174 are represented as spheres. C. Crystal structure of PV RdRp. D. Lys167 and Arg174 interact with the γ and α phosphate moiety of the incoming nucleotide respectively. Motif D is colored in hot pink, and the backbone and nucleobases of RNA are colored in grey and blue respectively. For Lys167, Arg174, and incoming NTP, heteroatoms are colored by blue (N), tan (C), orange (P), and red (O) respectively.
4.2 Methods

4.2.1 Materials

\[\gamma^{32}\text{P}]\text{ATP} \text{ and } [\alpha^{32}\text{P}]\text{UTP} (>7000\text{Ci/mmol}) \text{ were from VWR-MP Biomedical. Nucleoside 5'}\text{-triphosphates and 2'}\text{-deoxynucleoside 5'}\text{-triphosphates were ultrapure solutions from GE Healthcare. 3'}\text{-Deoxyadenosine 5'}\text{-triphosphate (cordycepin) was from Trilink Biotechnologies. } \alpha, \beta\text{-Methyleneadenosine 5'}\text{-triphosphate lithium salt (HPLC, solid) was from Sigma. All RNA oligonucleotides were from Dharmacon Research, Inc. (Boulder, CO). T4 polynucleotide kinase was from New England Biolabs, Inc. [Methyl-^{13}\text{C}]\text{methionine was from Cambridge Isotope Laboratories. HisPur Ni-NTA resin was from Thermo Scientific. Q-Sepharose fast flow was from Amersham Pharmacia Biotech, Inc. The polyethylenimine-cellulose TLC plates were from EM Science. The QuickChange site-directed mutagenesis kit was from Stratagene. Plasmid DNA isolation Miniprep kit was from Qiagen. Elutrap apparatus was from Schleicher & Schull. All other reagents were of the highest grade available from Sigma or Fisher.}

4.2.2 Plasmid construction

The plasmids encoding RdRp variants including K167R, K167M, R174K, and R174L were made based on the pET26-6H-SUMO-RdRp (pSUMO-RdRp) plasmid. The plasmid pSUMO-K359L-RdRp was from the Cameron lab. The plasmid pSUMO-R174KK359L-RdRp was made based on pSUMO-K359L-RdRp as the template. The plasmid products were checked using DNA sequencing at the Nucleic Acid Facility at the Pennsylvania State University.
4.2.3 Expression and purification of RdRp

Overexpression and purification of RdRp variants was described previously (see Methods in Chapter 2).

4.2.4 Purification, 5'-32P end labeling and annealing of sym/sub

Purification, 32P labeling of nucleic acid at the 5' end, and reannealing of nucleic acid followed the procedures as described previously (see Methods in Chapter 2).

4.2.5 PV RdRp Assay

Reactions contained 50 mM HEPES 7.5, 10 mM BME, 5 mM MgCl₂ or MnCl₂, and 60 µM ZnCl₂. Reactions were firstly incubated at 30 °C for 5 min and were then initiated by addition of either RdRp or nucleotides. RdRp enzyme was diluted immediately prior to reaction with enzyme dilution buffer (20% glycerol, 50 mM HEPES, 10 mM BME and 60 µM ZnCl₂) and was added to equal or less than one tenth of the total volume of the reaction. Reactions were quenched by addition of EDTA to a final concentration of 25 mM at each time point. Final concentrations of RdRp, sym/sub RNA, and nucleotides, along with any deviations from the above, are indicated in the appropriate figure legend.

4.2.6 Stopped flow experiment

Reaction buffer used in stopped flow experiment was the same as the buffer in the PV RdRp assay (see 4.2.5). Reactions were performed at 30 °C using a circulating water bath. RdRp-RNA binary complexes were formed and incubated at room temperature for 2 min. Reactions
were initiated by rapidly mixing the newly formed binary complexes with an equal volume of nucleotide solution. Nucleic acid sym/subU was substituted with a fluorescent nucleotide analog 2-aminopurine at the 5’ end. Fluorescence changes corresponding to product formation were monitored during the chain elongation reaction. For sym/subU RNA with modified 2-aminopurine, the excitation and emission wavelengths were 313 and 370 nm respectively. Final concentrations of RdRp, sym/subU RNA, and nucleotides, along with any deviations from the above, are indicated in the appropriate figure legend.

4.2.7 Benchtop assay for slow nucleotide incorporation

Reaction buffer used in the benchtop experiment was the same as the buffer in the PV RdRp assay (see 4.2.5). RdRp-RNA binary complexes were formed by incubating at room temperature for 3 min and then incubating at 30 °C for 2 min. Reactions were initiated by the addition of an equal volume of nucleotide solution at a specific concentration. Reactions were quenched by the addition of EDTA to a final concentration of 25 mM at different time points. Final concentrations of RdRp, end-labeled sym/sub RNA, and nucleotide, along with any deviation from the above, are indicated in the appropriate figure legend.

4.2.8 Fluorescence polarization experiment

Reactions contained 50 mM HEPES pH 7.5, 10 mM BME, 5 mM MgCl₂, 60 µM ZnCl₂, and 10 mM NaCl. Reactions were conducted at room temperature. Different concentrations of enzyme were added immediately from ice and the mini-polarization was measured after 10 min. Nucleic acid sym/subU was modified with 6-carboxyfluorescein (6-FAM) at the 5’ end. The concentration of 6-FAM-sym/subU RNA was fixed to 1 nM. The volume of enzyme solution
added into the reaction was no more than one tenth of the total assay volume. Final concentrations of RdRp and RNA are indicated in the appropriate figure legend.

4.2.9 Product analysis and data analysis

Product analyses and data fits followed previous procedures (see Methods in Chapter 2). Briefly, quenched product samples were preheated at 65 °C for 5 min and were isolated via denaturing PAGE gel at a constant 90 watts for 2.5 h. Then, gel was exposed to the screen, visualized by PhosphorImager, and quantified by ImageQuant software (Molecular Dynamics). Data was fit into different functions by KaleidoGraph software (Synergy) as indicated in the appropriate figure legend.

4.2.10 Sample preparation for NMR spectroscopy

Purified protein with/without [Methyl-\(^{13}\)C]methionine labeling was concentrated to 200 - 400 µM, desalted using spin columns (Thermo scientific), and exchanged into NMR buffer containing 10 mM HEPES 8.0, 200 mM NaCl, 5 mM MgCl\(_2\), 20 µM ZnCl\(_2\), and 0.02% NaN\(_3\) in D\(_2\)O. Binary RdRp-RNA complexes were formed by incubating RdRp (200 - 250 µM) with sym/subU RNA (200 - 500 µM) and 3’ dATP (3.1 mM). The purpose of adding 3’dATP was to form 3’ blocked primer terminus, which would then facilitate trapping of the ternary complexes after addition of the next incoming nucleotide. Concentrations of nucleotide ranged from 4 to 12 mM.
4.2.11 \(^1\)H - \(^{13}\)C HSQC NMR experiment

Most of experiments were performed on a Bruker Avance III 600 MHz spectrometer equipped with a 5 mm inverse detection triple resonance (\(^1\)H/\(^{13}\)C/\(^{15}\)N) single axis gradient TCI cryoprobe. \(^1\)H - \(^{13}\)C heteronuclear single quantum coherence (HSQC) spectra were generally acquired at 64 (\(t_1\)) x 512 (\(t_2\)) complex matrix, with 64 - 128 scans per increment and 1.0 s recovery delay.

4.2.12 \(^{31}\)P NMR experiment

\(^{31}\)P NMR spectra were collected using a Bruker Avance III 600 MHz NMR spectrometer with a \(^{31}\)P selective probe at spectrometer frequencies of 600.07 MHz for \(^1\)H and 242.91 MHz for \(^{31}\)P. Experiments were performed with fully calibrated 90° pulses and used a recycle delay of 1 s to suppress signal from free-floating RNA and nucleotide. Protons were decoupled during data acquisition using the Waltz16 sequence. \(^{31}\)P chemical shifts are expressed with reference to an 85% phosphoric acid external standard.

4.3 Results

4.3.1 The RNA binding affinity and stability of the RdRp-RNA binary complexes are slightly different between K167R and R174K RdRps.

We hypothesized that substitution at Lys167 and Arg174 in motif F would lead to changes in fidelity by perturbing the triphosphate conformation of the incoming nucleotides. In order to test the hypothesis, we designed two single variants: K167R and R174K. The rationale for these substitutions was to examine the effects of the side chain configuration, while
maintaining positive charge to avoid a great loss of nucleotide binding affinity. Before assessing any fidelity changes induced by these amino acid substitutions, I first characterized assembly and dissociation kinetics for the RdRp-RNA binary complexes. Unless otherwise noted, all the experiments were conducted in the presence of Mg$^{2+}$ ions.

In the assembly assay, I characterized the enzyme concentration dependence on the production of the RdRp-RNA binary complexes. The enzyme concentration varied from 0.25 to 5 µM. The concentration of sym/subU RNA was fixed to 1 µM and 2 µM for K167R and R174K RdRps respectively. It is worth mentioning that the concentration of sym/subU RNA for R174K RdRp used for the following kinetic experiments was twice that used for WT and K167R RdRps. Reactions were initiated by the addition of enzyme. For K167R RdRp, the rate and yield of the assembled complexes were similar to WT RdRp, which suggests that the Arg substitution did not affect RNA binding affinity (Figure 4-2). For R174K RdRp, the yield of the assembled complexes was different from that of WT RdRp, when enzyme concentration increased from 0.25 to 5 µM (Figure 4-3). This observation suggests that R174K RdRp had weaker binding affinity for sym/subU RNA than WT RdRp.

I next conducted dissociation assays to test whether the weak binding affinity of R174K RdRp to sym/subU RNA resulted from changes in the dissociation rate. The RNA dissociation rate ($k_{\text{dis}}$) was not significantly affected by the K167R and R174K substitutions (Figure 4-2, 4). Specifically, K167R RdRp ($k_{\text{dis}} = 3.5 \times 10^{-4}$ s$^{-1}$) had the same dissociation rate as WT RdRp ($k_{\text{dis}} = 3.5 \times 10^{-4}$ s$^{-1}$) (Figure 4-2). The dissociation rate of R174K RdRp ($k_{\text{dis}} = 11 \times 10^{-4}$ s$^{-1}$) increased by three fold, compared to WT RdRp (Figure 4-4). The half lives of the complexes for the two variants were more than 10 min, which suggested that the formed binary complexes were stable enough for subsequent single nucleotide incorporation assays.
Figure 4-2. The K167R substitution does not affect assembly and dissociation of enzyme complexes. Reactions contained 50 mM HEPES 7.5, 10 mM BME, 5 mM MgCl$_2$, and 60 µM ZnCl$_2$. A. Sequence of sym/subU RNA. B. Reaction scheme for the assembly assay. RNA and ATP were preincubated for 5 min. Reactions were then initiated by the addition of RdRp and were quenched by the addition of EDTA to a final concentration of 25 mM at different time points. Final concentrations of RdRp, sym/subU RNA, and ATP were 1, 1, and 500 µM respectively. C. Comparison of the assembled enzyme complexes of the K167R variant (red circle) to those of WT enzyme (blue circle). D. Reaction scheme for the dissociation assay. RdRp-RNA binary complexes were incubated for 90s followed by the addition of trap. At different time points, ATP was added and was allowed to incorporate in 30 s after which reactions were quenched by the addition of EDTA to a final concentration of 25 mM. Final concentrations of RdRp, sym/subU, and ATP concentrations were 1, 0.1, and 500 µM respectively. E. The decay of formed enzyme complexes of the K167R variant, compared to WT RdRp. The solid lines represent fits of the data to single exponential curves, which give k$_{\text{dis}}$ values of 0.00040 ± 0.00003 s$^{-1}$ and 0.00035 ± 0.00001 s$^{-1}$ for WT and K167R RdRps respectively.
Figure 4-3. The assembled complexes of R174K variant are slightly different from those of WT RdRp. Reactions were conducted in the same way as described in Figure 4-2. A, B, C and D. Comparison of the assembled complexes of R174K variants (red circle) to those of WT RdRp (blue circle). Reactions contained 500 µM ATP, either 1 or 2 µM sym/subU RNA for WT and R174K RdRps respectively, and either 0.25 (A), 0.5 (B), 1 (C), or 5 (D) µM RdRp.
Figure 4.4. The dissociation of enzyme complexes of R174K variant is slightly faster than that of WT RdRp. Reaction and data analyses were conducted in the same way as described in Figure 4-2 except that the concentration of sym/subU RNA was 0.2 µM. A. Comparison of the dissociation of enzyme complexes between R174K and WT RdRps. The calculated $k_{\text{dis}}$ for R174K RdRp is $0.0011 \pm 0.00008 \text{ s}^{-1}$. 
4.3.2 The single Arg substitution at Lys167 results in low fidelity, but the single Lys substitution at Arg174 results in high fidelity.

Single nucleotide incorporation assays were conducted for both K167R and R174K RdRps. The maximum rate (k_{pol}) and the apparent dissociation constant (K_{d,app}) were calculated for cognate (ATP) and noncognate nucleotides (2’dATP or GTP). The fidelity was defined as the ratio of the efficiency (k_{pol}/K_{d,app}) of the cognate nucleotide incorporation over that of the noncognate nucleotide incorporation. For example, if a variant gave rise to a higher ratio ((k_{pol}/K_{d,app})_{cognate}/(k_{pol}/K_{d,app})_{noncognate}) compared to that of WT RdRp, the variant was said to be of high fidelity. Results indicated that K167R RdRp (2.1 µM⁻¹s⁻¹) incorporated the cognate nucleotide as efficiently as WT RdRp (1.6 µM⁻¹s⁻¹), but incorporated both the 2’deoxynucleotide (1.1 x 10⁻² µM⁻¹s⁻¹) and the noncognate nucleobase (1.7 x 10⁻⁴ µM⁻¹s⁻¹) faster than WT RdRp (6.6 x 10⁻³ µM⁻¹s⁻¹ for the 2’deoxynucleotide and 7.7 x 10⁻⁵ µM⁻¹s⁻¹ for the noncognate nucleobase) (Table 4-1). These observations indicated that K167R RdRp was a low fidelity variant. The Lys167 substitution appeared to also be important for noncognate nucleotide binding since the apparent dissociation constant increased by 2 fold (Table 4-1).

For R174K RdRp, the catalytic efficiency of the cognate nucleotide incorporation was reduced by 7 fold (0.22 µM⁻¹s⁻¹) compared to WT RdRp (Table 4-1). Unexpectedly, the catalytic efficiency for 2’deoxynucleotide incorporation was reduced by 7000 fold (9.2 x 10⁻⁷ µM⁻¹s⁻¹). For the noncognate nucleobase incorporation, K_{d,app} could not be determined, but the k_{obs} dropped by more than 350 fold when the concentration of noncognate nucleobase was 5 mM (Table 4-1). These findings suggest that fidelity increased by 261 and 26 fold for the 2’deoxynucleotide and noncognate nucleobase incorporations respectively. The fidelity increases were likely due to the impaired reaction in the active site since the K_{d,app} value of R174K RdRp was similar to WT RdRp (Table 4-1). The R174K variant is the highest fidelity variant that has been reported so far.
Previously, the highest fidelity variant was K359R variant, in which a 5-fold increase in fidelity was observed.
<table>
<thead>
<tr>
<th>RdRp</th>
<th>nucleotide</th>
<th>Mg^{2+}</th>
<th>k_{pol} (s^{-1})</th>
<th>K_{d,app} (µM)</th>
<th>Fidelity in k_{pol}</th>
<th>Fidelity in k_{pol}/K_{d,app}</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>ATP</td>
<td>59 ± 1</td>
<td>36 ± 2</td>
<td></td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>ATP (D2O)</td>
<td>27.3 ± 0.4</td>
<td>15 ± 1</td>
<td></td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>2’dATP</td>
<td>0.89 ± 0.01</td>
<td>134 ± 4</td>
<td>70</td>
<td>250</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GTP</td>
<td>0.011 ± 0.001</td>
<td>142 ± 15</td>
<td>5400</td>
<td>21000</td>
<td></td>
</tr>
<tr>
<td>R174K</td>
<td>ATP</td>
<td>4.2 ± 0.2</td>
<td>19 ± 4</td>
<td></td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>ATP (D2O)</td>
<td>1.7 ± 0.1</td>
<td>2.5 ± 0.1</td>
<td></td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>2’dATP</td>
<td>2.4 ± 0.2 x 10^{-4}</td>
<td>26 ± 6</td>
<td>17500</td>
<td>24000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GTP</td>
<td>3.0 ± 0.1 x 10^{-5}</td>
<td>-</td>
<td>140000</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>K167R</td>
<td>ATP</td>
<td>56 ± 1</td>
<td>27 ± 1</td>
<td></td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>ATP (D2O)</td>
<td>11.7 ± 0.4</td>
<td>7.5 ± 1.1</td>
<td></td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>2’dATP</td>
<td>3.4 ± 0.1</td>
<td>321 ± 25</td>
<td>20</td>
<td>200</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GTP</td>
<td>0.037±0.01</td>
<td>223 ± 20</td>
<td>1500</td>
<td>12500</td>
<td></td>
</tr>
<tr>
<td>K167M</td>
<td>ATP</td>
<td>26 ± 1</td>
<td>683 ± 29</td>
<td></td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>K359L</td>
<td>ATP</td>
<td>0.62 ± 0.01</td>
<td>272 ± 15</td>
<td></td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>R174L</td>
<td>ATP</td>
<td>0.13 ± 0.03</td>
<td>169 ± 15</td>
<td></td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>R174KK359L</td>
<td>ATP</td>
<td>0.011 ± 0.001</td>
<td>148 ± 24</td>
<td></td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
4.3.3 The pre-chemistry conformational change in both K167R and R174K variants significantly contributes to changes in fidelity.

In the presence of Mg\(^{2+}\), both the pre-chemistry conformational change and the phosphoryl transfer step are partially rating limiting and help determine the maximum rate. In order to check whether the pre-chemistry conformational change was perturbed in the two single variants, solvent deuterium kinetic isotope effects (SDKIEs) and Mn\(^{2+}\) effects were determined. The K167R variant had much higher SDKIE values (4.7 and 6.8 using Mg\(^{2+}\) and Mn\(^{2+}\) respectively) than WT RdRp (2.2 and 3.1 using Mg\(^{2+}\) and Mn\(^{2+}\) respectively) regardless of divalent metals. The R174K variant also had higher SDKIE values (2.5 and 4.3 for Mg\(^{2+}\) and Mn\(^{2+}\) respectively) but not as high as those for the K167R variant. The increase in the SDKIE for both K167R and R174K variants suggests that the chemistry was more rate limiting for cognate nucleotide incorporation.

I also conducted single nucleotide incorporation assays in the presence of Mn\(^{2+}\) for the K167R and R174K variants to examine perturbations of the pre-chemistry conformational change. By comparing rates in Mg\(^{2+}\) and Mn\(^{2+}\), we can get indirect insights into the pre-chemistry conformational equilibrium. Previous studies indicated that the kinetics of RdRp in Mn\(^{2+}\) is completely different from Mg\(^{2+}\). Only the pre-chemistry conformational change determines substrate specificity although the chemistry step is a rate-limiting step in the cognate nucleotide incorporation. Therefore, comparison of rates (k_{pol}) of nucleotide incorporation between WT RdRp and the RdRp variants allows an understanding of the contributions of the conformational equilibrium to catalysis and fidelity of RdRp.

In Mn\(^{2+}\), the K167R RdRp had a similar rate as WT RdRp for cognate nucleotide incorporation, but had higher rates for 2’deoxynucleotide and noncognate nucleobase incorporation (Table 4-2). Consequently, both the sugar and nucleobase fidelity of K167R RdRp decreased by two fold compared to WT RdRp. The decrease in fidelity likely derives from the
formation of more stable complexes with noncognate nucleotide prior to the chemistry step. These results are also consistent with the larger SDKIE values in the K167R variant compared to WT enzyme.

The R174K variant had a very slow rate for the cognate nucleotide incorporation relative to WT RdRp in Mn$^{2+}$, which is consistent with the rate reduction in Mg$^{2+}$ (15 and 10 fold reduction for Mg$^{2+}$ and Mn$^{2+}$ respectively) (Table 4-2). This observation suggests that both the pre-chemistry conformational change and chemistry steps are affected significantly. For the 2’deoxyribonucleotide incorporation, the rate decreased even more (i.e. ~ 15 fold), compared to the rate for WT RdRp (Table 4-2). Consequently, the sugar fidelity of R174K RdRp increased by 1.5 fold. The increased fidelity likely derives from formation of more unstable complexes with 2’deoxyribonucleotide prior to the chemistry step. For noncognate nucleobase incorporation, the rate decreased by 8 fold compared to WT RdRp (Table 4-2).
Table 4-2. Kientic rate constants for nucleotide incorporations in the presence of Mn\(^{2+}\)

<table>
<thead>
<tr>
<th>RdRp</th>
<th>nucleotide</th>
<th>(k_{pol}) (s(^{-1}))</th>
<th>(K_{D,app}) (µM)</th>
<th>Fidelity in (k_{pol})</th>
<th>Fidelity in (k_{pol}/K_{D,app})</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>ATP</td>
<td>16.2 ± 0.4</td>
<td>1.7 ± 0.1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>ATP (D2O)</td>
<td>5.3 ± 0.1</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>2’dATP</td>
<td>7.4 ± 0.1</td>
<td>6.1 ± 0.1</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>GTP</td>
<td>0.8 ± 0.1</td>
<td>98 ± 3</td>
<td>20</td>
<td>1200</td>
</tr>
<tr>
<td>R174K</td>
<td>ATP</td>
<td>1.7 ± 0.1</td>
<td>0.20 ± 0.03</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>ATP (D2O)</td>
<td>0.4 ± 0.1</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>2’dATP</td>
<td>0.5 ± 0.1</td>
<td>7.6 ± 0.5</td>
<td>3</td>
<td>130</td>
</tr>
<tr>
<td></td>
<td>GTP</td>
<td>0.10 ± 0.01</td>
<td>57 ± 5</td>
<td>15</td>
<td>4400</td>
</tr>
<tr>
<td>K167R</td>
<td>ATP</td>
<td>21.4 ± 0.5</td>
<td>4.1 ± 0.3</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>ATP (D2O)</td>
<td>3.1 ± 0.1</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>2’dATP</td>
<td>14.9 ± 0.4</td>
<td>9.5 ± 0.7</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>GTP</td>
<td>1.9 ± 0.1</td>
<td>168 ± 19</td>
<td>11</td>
<td>460</td>
</tr>
</tbody>
</table>
The pre-chemistry conformational change in the R174K variant is significantly damaged due to an inability to form a stable “closed” state.

The kinetic studies suggested that the R174K substitution might perturb the open-to-closed conformational change. To gain more insight, I conducted the AMPCPP (α, β-methyleneadenosine 5’-triphosphate) assay for WT RdRp and the two single variants. In this assay, AMPCPP (or water as control) was preincubated with RdRp-RNA binary complexes for 5 min to trap the “closed” state by forming RdRp-RNA-AMPCPP ternary complexes, and then the reaction was initiated by the addition of an excess amount of cognate nucleotide. The reaction occurs only when AMPCPP is exchanged with the cognate nucleotide. The rate for nucleotide incorporation was determined with and without AMPCPP. The delay in nucleotide incorporation induced by AMPCPP is a measure of the reverse rate of the pre-chemistry conformational change. For example, the longer the delay is, the more stable the “closed” state. For WT and K167R RdRps, the rates were delayed by 40 to 50 fold (Figure 4-5, Table 4-3). However, in R174K RdRp, the rate was delayed by only three fold, which was significantly lower than both WT and K167R RdRps (Figure 4-5, Table 4-3). These results suggest that the R174K substitution de-stabilized the “closed” state of the ternary complex. This finding was also consistent with the previous Mn$^{2+}$ effect studies that suggested that the pre-chemistry conformational change was shifted towards the “open” state when cognate nucleotide was bound and was shifted even more toward the “open” state when noncognate nucleotide was bound. I did not observe a longer delay for K167R RdRp compared to WT RdRp as I expected. This might be because the equilibrium shift occurred only in the noncognate nucleotide incorporations or the shift was too small to be measurable by the AMPCPP-trap method.
Figure 4-5. Comparison of the AMP incorporation in the AMPCPP assay for WT, K167R and R174K RdRps suggests that the R174K substitution significantly perturbs the prechemistry conformational change. A. Minimal kinetic mechanism of the dissociation of AMPCPP ternary complexes monitored by incorporation of AMP. Reactions contained 50 mM HEPES 7.5, 10 mM BME, 5 mM MnCl$_2$, and 60 µM ZnCl$_2$. Enzyme was preincubated with sym/subU RNA to form binary complexes for 90s. Then 200 µM AMPCPP or buffer was added and incubated for extra 5 min to form RdRp-sym/subU-AMPCPP ternary complexes. Reactions were initiated by addition of an excess amount of ATP. AMP incorporation was then monitored by fluorescent changes using a stopped flow apparatus. Final concentrations of ATP, sym/subU RNA, and RdRp were 1000, 1 or 2 µM (for WT and K167R RdRps and R174K RdRp respectively), and 1 µM respectively. B, C, and D. The comparison of AMP incorporation in the presence (blue) and absence (red) of AMPCPP for WT (B), K167R (C), and R174K RdRps (D). The solid lines represent fits of the data into single exponential curves and give the $k_{obs}$ values that are listed in Table 4-3.
Table 4-3. Comparison of the AMP incorporation in the presence and absence of AMPCPP for WT, K167R, and R174K RdRps respectively.

<table>
<thead>
<tr>
<th></th>
<th>k_{obs} (s^{-1})</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>WT</td>
</tr>
<tr>
<td>Control</td>
<td>30.62 ± 0.13</td>
</tr>
<tr>
<td>AMPCPP</td>
<td>0.62 ± 0.01</td>
</tr>
<tr>
<td>Fold reduction</td>
<td>49</td>
</tr>
</tbody>
</table>
4.3.5 Motif F and motif D are independent of each other during the pre-chemistry conformational change.

I have shown that amino acid substitutions on motif F lead to high or low fidelity by perturbing the equilibrium of the pre-chemistry conformational change. Previous NMR and biochemical results from our lab and Dr. Cameron’s lab suggest that motif D also helps determine the polymerase fidelity via repositioning of the general acid Lys359 on this motif\(^{26}\). The substitution of Thr362 to Ile lowers the polymerase fidelity by shifting the equilibrium of motif D more to the “closed” state (see Chapter 2). The observations in this Chapter also suggest that motif F is important in nucleotide selection, in part through interactions with the triphosphate of the incoming nucleotide. Here, it was interesting to know the relationship between motifs D and F. Are the two motifs working independently or cooperatively? What is the sequential order during the prechemistry conformational change, or does the sequential order even matter?

One way to study the relationship between motifs D and F is through a thermodynamic cycle comparing R174K, K359L, and R174K/K359L variants. The calculation of the thermodynamic cycle was described previously (see Chapter 3). Arg174 was changed to Lys not Leu because the catalytic activity of the R174L variant was too weak to be measured accurately (data not shown).

The thermodynamic analysis for single nucleotide incorporation required that the dissociation rate of the enzyme complex did not affect these assays. I also constructed additional variants, including K167M and R174L, and examined the RNA binding affinities and dissociation rate for WT RdRp and all the Lys167 and Arg174 RdRp derivatives. The RNA binding affinity was measured using fluorescence polarization (Beacon 2000, Panvera). Briefly, the enzyme was added immediately from ice and preincubated with 6-FAM-sym/subU at the 5’ end, under room temperature, for 10 min to reach thermodynamic equilibrium followed by a read out of the mini-polarization (mP) value for each enzyme concentration. The final concentration of the enzyme
ranged from 12.5 to 3000 nM. The highest enzyme concentration was limited by enzyme precipitation. The results indicated that substitutions at both Lys167 and Arg174 position affected the binding affinity of RNA (Table 4-4). Specifically, K167M (K_D is 254 ± 28 nM) and K359L variants (K_D is 397 ± 28 nM) had stronger binding affinity than WT enzyme (K_D is 575 ± 68 nM). K167R (K_D is 767 ± 203 nM) and R174K variants (K_D is 603 ± 35 nM) had comparable binding affinity with WT enzyme. R174K/K359L (K_D is 1024 ± 95 nM) and R174L variants (K_D is 1951 ± 304 nM) had weaker binding affinity. The dissociation constants for the R174K/K359L and R174L RdRp were reduced by two and three fold respectively, compared to WT enzyme.

To look at whether the reduction in the binding affinity results from changes to the association and dissociation rates, I conducted the dissociation assay for WT enzyme and all the Lys167 and Arg174 RdRp variants. Most variants had similar dissociation rates. The R174K/K359L (4.2 ± 0.7 x 10^-4 s^-1) and K359L variants (1.4 ± 0.1 x 10^-4 s^-1) showed comparable or stronger stability respectively compared to WT enzyme for the RdRp-RNA binary complex. The half lives ranged from 30 min to 1.5 hour, which ensured that the dissociation of the binary complex would not hinder the time course for the single nucleotide incorporation assays.

The single nucleotide incorporation assays were conducted for both the R174K/K359L and K359L variants under the same condition as WT RdRp and R174K variant (Figure 4-6). The k_pol/K_d,app values were calculated for all the variants in order to establish the thermodynamic cycle. The ΔΔG values of R174K and K359L variants were -1.2 and -4.0 kcal/mol respectively (Figure 4-6). The ΔΔG value of R174K/K359L variant was -6.0 kcal/mol (Figure 4-6). The sum (-5.2 kcal/mol) of the ΔΔG values of the two single variants only slightly differed from that of the double variant, which suggests that motif D and motif F were functioning independently.
Figure 4-6. Motifs D and F are functionally independent. The units are kcal/mol. Thermodynamic effect $\Delta\Delta G$ is calculated by the following equation $\Delta\Delta G_{x1} = RT \ln (k_{pol}/K_{d,app})_{x1}/(k_{pol}/K_{d,app})_{x2}$, where $x1$ represents the variant that an arrow points to and $x2$ represents the variant that the same arrow begins$^{31}$. The thermodynamic cycle is calculated for AMP incorporation by WT, K359L, R174K, and R174KK359L RdRps.
I have shown that substitutions on motif F lead to changes in fidelity through perturbing the pre-chemistry conformational change. Specifically, the K167R substitution decreased fidelity through stabilizing the noncognate ternary complexes prior to the chemistry step. The R174K substitution significantly affected the active site and de-stabilized the noncognate ternary complex more to increase enzyme fidelity. These observations suggest that there are different interactions between the RdRp variants and nucleotide substrates compared to those interactions found in the WT complexes. It is likely that multiple interactions between the incoming nucleotide and residues in the nucleotide binding pocket and the active site have to be satisfied in order to check the nucleobase, ribose, and triphosphate group during the pre-chemistry conformational change. Substitutions on motif F likely affect the interactions with the triphosphate moiety to lead to changes in fidelity. It is also possible that other interactions are perturbed since it has been suggested that the polymerase is globally dynamic and different motifs work coordinately or sequentially. Therefore, it is interesting to know the contributions of different motifs to the fidelity mechanism. Here, we mainly considered the triphosphate orientation on motif F and the general acid reposition on motif D by using $^{31}$P NMR and [Methyl-$^{13}$C]Met labeled NMR respectively.

The sample preparation for active enzyme complexes for $^{31}$P NMR was the same as [Methyl-$^{13}$C]Met labeled NMR as described previously (see Methods in Chapter 2). Acquiring $^{31}$P NMR spectra at every step of sample preparation for the WT RdRp complexes allowed us to track the formation of enzyme complexes and complete assignment of the triphosphate of bound nucleotide. Note that the peaks for the phosphate in RNA were at 0-1.2 ppm (not discussed here), which was distant from the peaks for the triphosphate in the nucleotide (-5 to -20 ppm) (Figure 4-7). Here, we will focus on the triphosphate peaks. When an excess amount of 3’dATP was added,
two sets of peaks were observed. The set of intense peaks was assigned for the free 3’dATP (Figure 4-7). The other set of weak peaks was assigned for the protein bound 3’dATP due to the presence of nonincorporated 3’dATP in the active site. Passing the complex through the desalting column removed all the free 3’dATP and protein bound 3’dATP as indicated by the disappearance of both sets of peaks (Figure 4-7). When cognate nucleotide was added, two sets of peaks were observed again (Figure 4-7). The set of three intense peaks (δ -4.98, -10.08, -18.65 ppm) was the same as the free nucleotide coordinated to Mg^{2+}, but different from the nucleotide in the absence of Mg^{2+}, which was consistent with the previous suggestion that metal B firstly coordinates to the nucleotide in solution and enters the active site during the nucleotide ground binding (Figure 4-7)^2,13,33,34. But this time, when passing through the second desalting column, only the set of intense peaks disappeared (Figure 4-7). The set of weak peaks were retained for protein bound nucleotide (Figure 4-7). The α and γ phosphate were at -9.53 and -5.38 ppm respectively (Figure 4-7). We did not observe the peak for β phosphate probably due to low peak intensity and/or conformational exchange processes that broadened out the peak (Figure 4-7).
Figure 4-7. NMR scans at every step of sample preparation allows assignment of the protein bound triphosphate. A. Sample preparation for $^{31}$P NMR experiments. 3’dAMP is incorporated into sym/subU RNA by RdRp. The blocked 3’H RNA allows addition of the second nucleotide UTP to form the RdRp-sym/subU-UTP ternary complex. A desalting column is used to remove unincorporated free nucleotide. B. The triphosphate of bound and unbound nucleotide can be detected by $^{31}$P NMR. Not shown are the peaks for the phosphate-backbone of the RNA (-1.2 ppm). The spectra for the RdRp complexes (RdRp-sym/subU+3’dATP, RdRp(3’H), RdRp(3’H)+UTP and RdRp(3’H)-UTP (desalted) shown in orange, brown, purple, and green respectively) are compared to the spectra for free UTP in the presence (blue) and absence (red) of Mg$^{2+}$. 

A

5’-GCAUGGGCCC

CCCGGUACG-5’

+ 3’dATP

↓ Desalt

5’-GCAUGGGCCCA$^{3'H}$

CCCGGUACG-5’

+ UTP

↓ Desalt

5’-GCAUGGGCCCAU$^{3pp}$

CCCGGUACG-5’

B
4.3.7 Triphosphate orientation in the ternary complexes is a fidelity checkpoint.

To look at the changes in the triphosphate orientation, we compared the $^{31}$P NMR spectra between cognate ternary complex (RdRp-RNA-UTP) and noncognate ternary complexes (RdRp-RNA-2’dUTP or RdRp-RNA-CTP). The $^{31}$P NMR experiments were designed to minimize signal from free nucleotide by using a short recycle delay. However, this also meant that the amounts of free nucleotide and protein-bound nucleotide were not directly related to the intensities of the NMR peaks. The experimental temperature was also increased up to 303K from 293K in order to avoid overlapping the α phosphate peaks for free nucleotide and protein bound nucleotide. Not surprisingly, there were significant differences between the cognate complex and noncognate complexes (Figure 4-8). The α phosphate for the noncognate ternary complexes (δ -9.84 ppm for protein bound 2’dUTP and -9.87 ppm for protein bound CTP) was less downfield shifted compared to that of the cognate complex (δ -9.36 ppm) (Figure 4-8). These observations suggest that the α phosphate in the noncognate ternary complex was in a different chemical environment than that found in the cognate ternary complex (Figure 4-8). The γ phosphate for the noncognate ternary complexes (δ -5.2, -6.21 ppm for protein bound 2’dUTP and -5.2, -6.16 ppm for protein bound CTP) was also different from that of the cognate ternary complex (δ -5.3 ppm) (Figure 4-8). These observations suggest that the conformation of the γ phosphate in the noncognate ternary complexes was exchanged among multiple conformations and contributed to the destabilization of the ternary complex (Figure 4-8). Interestingly, the conformations of the triphosphate group of noncognate ternary complexes between 2’deoxynucleotide and noncognate nucleobase were very similar as reported by similar chemical shifts, which suggested that WT enzyme utilized partially similar strategies to distinguish between different types of nucleotide. Altogether, these observations suggested that the triphosphate reorientation was an important fidelity checkpoint.
$^{31}$P NMR experiments indicate that the triphosphate conformations of the incoming nucleotide for K167R, R174K, and K359R ternary complexes are different from those of WT enzyme. A, B, and C. Comparison of the conformations of the triphosphate among variants and WT enzyme for cognate UTP (A), 2’deoxynucleotide (B), and noncognate nucleobase (C) ternary complexes respectively. Spectra of K359R, R174K, K167R, and WT ternary complexes are shown in purple, green, red, and blue respectively. Spectra were collected at 303 K with 250 µM RdRp, 500 µM RNA, and 4 mM UTP, 8 mM 2’dUTP or 12 -18 mM CTP.
4.3.8 Substitutions interacting with $\alpha$, $\beta$, or $\gamma$ phosphate contribute to fidelity in different ways.

$^{31}$P NMR experiments were also conducted for R174K, K359R, and K167R variants. Results indicated that the K359R substitution only affected the chemical shift position of the cognate ternary complex such that both $\alpha$ and $\gamma$ phosphates ($\delta$ -9.49, -5.43 ppm) were slightly upfield shifted, compared to WT enzyme (Figure 5-8). For the noncognate ternary complexes, the chemical shift positions in K359R RdRp were highly similar to those in WT enzyme (Figure 4-8). These observations were consistent with previous NMR results that suggested that motif D only fluctuates into the “closed” state to protonate the $\beta$ phosphate moiety of cognate nucleotide. When noncognate nucleotide binds, motif D remains primarily in the “open” state.

For the R174K variant, the chemical environment of the triphosphate in the cognate ternary complex appeared to be significantly different from WT RdRp (Figure 4-8). Specifically, the $\alpha$ phosphate was upfield shifted ($\delta$ -9.74, -9.98 ppm), which suggests that the $\alpha$ phosphate in R174K would be misaligned with the 3’ OH of the primer (Figure 4-8). The $\gamma$ phosphate in the R174K RdRp had two peaks ($\delta$ -5.15, -6.42 ppm), which was different from that of the WT complex ($\delta$ -5.3 ppm) (Figure 4-8). The two peaks in the R174K complex suggested that the triphosphate group was sampling between two conformations. Surprisingly, the conformation of the triphosphate for the R174K variant in the cognate ternary complex was very similar with that in the ternary complex with noncognate nucleobase (Figure 4-8). The similarity between the two ternary complexes was consistent with the kinetic results suggesting that R174K variant favored the “open” state. We also found that the intensities and the width of peaks were different between the ternary complex with cognate nucleotide and with noncognate nucleobase, which may be due to the use of different nucleotide concentrations (i.e. [UTP]=4 mM, [CTP]=12 mM) or due to different intrinsic equilibrium constants between the two conformations (Figure 4-8). Lastly, in
the ternary complex with 2’deoxynucleotide, γ phosphate (δ -5.28, -5.89, -6.32 ppm) had one more peak than WT RdRp (Figure 4-8).

For the K167R variant, there was only one very broad peak for the γ phosphate moiety in the noncognate ternary complexes (Figure 4-8). This broad peak might reflect a faster rate of the nucleotide on/off from the polymerase (i.e. intermediate exchange), which was consistent with the increase in $K_{d,app}$ by ten fold (Table 4-1). This broad peak might also reflect the faster exchange rate among different conformations of triphosphate, which was consistent with the SDKIE data that suggested that the prechemistry conformational change was less rate limiting (Figure 4-8).

4.3.9 The function of motif F is not necessarily coordinated with that of motif D

The substitutions on motif F may or may not affect the conformation of motif D. [Methyl-\textsuperscript{13}C]Met labeled NMR studies were conducted for the R174K and K167R variants. Met354 is a very important reporter for changes in the structure and conformation of motif D. There are chemical shift changes for other reporters including Met6, Met74, and Met225 that accompany by those of the Met354 probe. For the R174K variant, motif D stayed in the “open” state for both cognate ternary complex and noncognate ternary complexes (Figure 4-9). The conformational state of motif D as assessed by the NMR studies suggested that the Lys174 substitution not only disturbed the conformation of the triphosphate group but also perturbed the position of motif D.

The K167R variant had similar NMR spectra as WT enzyme for both cognate ternary complex and noncognate ternary complexes, which suggests that the K167R substitution only affected the conformation of the triphosphate but not the conformation of motif D (Figure 4-9). This finding was consistent with the kinetic data that suggests that the function of motif D was independent of motif F.
Figure 4-9. Functions and conformations of motifs D and F are independent. A, B, and C. Comparisons of [Methyl-\textsuperscript{13}C]Met NMR spectra for K167R ternary complexes (red) with cognate nucleotide UTP (A), 2'dUTP (B), and noncognate nucleobase CTP (C) with those for WT ternary complexes (black). The K167R substitution does not perturb the conformation of motif D with both cognate and noncognate nucleotides. D, E. Comparison of the [Methyl-\textsuperscript{13}C]Met NMR spectra for R174K ternary complexes (green) with cognate nucleotide UTP (D) and noncognate nucleotide (E) with those for WT ternary complexes (black). The R174K variant cannot form the same conformation of motif D as that of WT enzyme. F. Overlapping of cognate (black) and noncognate (green) ternary complexes for R174K variant indicates that the R174K substitution significantly impairs the prechemistry conformational equilibrium such that the “closed” state cannot form even with cognate nucleotide.
4.4 Discussion

In this Chapter, we used biochemical assays and solution-state NMR experiments to gain more information about the role that motif F residues play in nucleotide selection. For these studies, we generated two motif F variants R174K and K167R; Arg174 and Lys167 are proposed to interact with the α and γ phosphate moieties of the incoming nucleotide respectively. Assembly and dissociation assays indicated that the K167R variant had similar binding affinity and dissociation rate for sym/subU RNA as WT enzyme and that the R174K variant had slightly weaker binding affinity and faster dissociation rate for sym/subU RNA compared to WT enzyme (Figure 2, 3 and 4). For the K167R variant, single nucleotide incorporation assays indicated that this variant did not impact the catalytic efficiency of cognate nucleotide incorporation (Table 1). However, the K167R variant had a moderately low fidelity for both sugar and nucleobase selection by enhancing the catalytic efficiency for incorporation of noncognate nucleotides (Table 1). For the R174K variant, single nucleotide incorporation assays indicated that this variant reduced the catalytic efficiency of both cognate and noncognate nucleotide incorporation especially with the drastic reduction of noncognate nucleotide incorporation by 7000 fold (Table 1). Consequently, the R174K variant was the highest fidelity variant of PV RdRp that has been observed so far. AMPCPP assays for the R174K variant suggested that the R174K substitution de-stabilized the closed conformation of the active site (Figure 4-5).

In NMR studies, $^{31}$P NMR experiments for motif F variants and WT enzyme indicated that the orientation of the triphosphate group of incoming nucleotide was a critical fidelity checkpoint. The differences in the $^{31}$P NMR spectra suggested that R174K and K167R substitutions induced very different chemical environments around the phosphates, likely reporting on different orientations of the triphosphate. These studies were consistent with kinetic studies that indicated that the prechemistry conformational change step was disturbed by the
R174K and K167R substitutions (Figure 4-8). \(^1\)H,\(^1\)C [Methyl-\(^1\)C]Met HSQC experiments were also conducted for the motif F variants and were compared to WT enzyme. Previously, we have shown that NMR experiments provide spectral fingerprints for the open and closed states, and these studies indicated that the conformation of motif D is also an important fidelity checkpoint\(^{26, 27, 32}\). Results indicated that the K167R variant had similar spectra to WT enzyme but the R174K variant remained primarily in the “open” state even when cognate nucleotide was bound (Figure 4-9). These findings suggested that the conformational changes and functions of motifs D and F were largely independent, consistent with the R174K/K359L double variant thermodynamic cycle (Figure 4-6).

PV RdRp, like other types of nucleic acid polymerases, follows a two-proton transfer mechanism, where one proton is removed from the primer hydroxyl group and the second proton is used to protonate the pyrophosphate leaving group\(^2, 35\). There is controversy about the identity of the general acid in PV RdRp. Peersen and colleagues have suggested that Arg174 from motif F, located just above the \(\alpha-\beta\) phosphate linkage, would be able to act as the general acid and Lys359 from motif D is not the general acid due to its position distant from the active site (Figure 4-10)\(^{36}\). By contrast, the Cameron and Boehr labs have suggested that motif D undergoes a conformational change that may allow Lys359 to approach the \(\beta\)-phosphate moiety and act as the general acid\(^{24, 26, 27, 32}\). Such a conformational change is strictly dependent on the nature of the nucleotide. Only cognate nucleotide could trigger motif D to shift from “open” to “closed” state\(^{26}\). Kinetic studies also indicated that a Leu substitution at 359 reduces the number of transferred protons from two to one in proton inventory experiments\(^{24}\). Our studies on the R174K variant further support the suggestion that Lys359 is the general acid, and Arg174 plays other important catalytic roles. If Arg174 was the general acid, we might expect that replacement of this residue with Lys would enhance enzyme catalysis and decrease fidelity due to Lys having a much lower pKa than Arg. Instead, we proposed that the major function of Arg174 is to bind and align the \(\alpha\)
phosphate moiety of the incoming nucleotide with the primer 3’ hydroxyl, as suggested by the $^{31}$P NMR experiments. Additionally, the Leu substitution at Arg174 led to an increase in $K_{d,app}$ by 4 fold and a decrease in rate by 455 fold compared to WT enzyme (Table 4-1).

Recently, the Cameron group has proposed a structural model for nucleotide binding to predict enzyme fidelity and assist the design of fidelity variants (Figure 4-10). This model is comprised of two states, those being “occluded” state (which is equivalent to the “open” state in this Chapter) and “competent” state (which is equivalent to the “closed” state in this Chapter) derived from the molecular dynamic simulations of RdRp-RNA binary complexes for WT and a low fidelity variant H273R (Figure 4-10). This model includes many functional residues from conserved motifs and interactions of these residues with an incoming nucleotide (Figure 4-10). The “occluded” state prevents nucleotide binding mainly because the ionic pair of Arg174 and Asp238 occupies and blocks the nucleotide-binding pocket (Figure 4-10). A transition from “occluded” state to “competent” state required a breakage of the ionic interaction between Arg174 and Asp238 and an establishment of new interactions between enzyme and incoming nucleotide (e.g. Arg174 forms a salt bridge interaction with the $\alpha$ phosphate moiety and Asp238 forms a hydrogen bond interaction with Ser288) (Figure 4-10). The Cameron lab suggested that the equilibrium position between the “occluded” state and the “competent” state determines fidelity. For example, if a variant could create more contacts with nucleotide to better stabilize the “competent” state, a low fidelity phenotype would be expected and likewise, interactions that stabilize the “occluded” state would lead to a high fidelity phenotype.

Our studies on K167R and R174K variants nicely fit this model. The Arg substitution on Lys167 not only keeps the positive charge, but also increases the number of proton donors that would further enhance the interaction with the $\gamma$ phosphate moiety and to shift the equilibrium to the “competent” state. There may be even a larger shift to the “competent” state for noncognate nucleotide incorporation, which would lead to a low fidelity phenotype (Figure 4-10).
Additionally, the Arg substitution likely does not change the structure of the active site since this residue is located at the very bottom of the active site and is less likely to disturb other active site residues (Figure 4-10). By contrast, the Lys substitution with a shorter side chain might pull the α phosphate moiety towards itself resulting in a partial loss of contacts between the β phosphate moiety and Lys359 on the other side (Figure 4-10), which explains why an “open” state of motif D was observed in the cognate ternary complex of R174K variant in [Methyl-13C]Met HSQC NMR spectra (Figure 4-9). Additionally, the pulling motion may also change interactions between Lys167 and the γ phosphate moiety and between the α phosphate moiety and the 3’ ribose hydroxyl (Figure 4-10). Altogether, the R174K substitution disturbs multiple interactions between the active site and the triphosphate of the incoming nucleotide, which leads to a reduction in catalysis. Such perturbations might be even more deleterious for noncognate nucleotide incorporation, and thus leading to a high fidelity phenotype.
Figure 4-10. The “occluded” (A) and “competent” states (B) of the active site in PV RdRp. Residues from highly conserved motifs A (green), B (tan), D (blue), and F (red) participate in the regulation of the equilibrium between “occluded” and “competent” states. The figure was reprinted from reference 37. © The American Society for Biochemistry and Molecular Biology.
4.5 References

CHAPTER 5

A NOVEL ANTIVIRAL INHIBITOR TARGETING THE RNA TEMPLATE CHANNEL OF COXSACKIEVIRUS B3

In this Chapter, CVB3 WT RdRp and inhibitor GPC-N114 were obtained from Dr. Frank J. M. van Kuppeveld’s lab in the Department of Medical Microbiology, Nijmegen Center for Molecular Life Sciences & Nijmegen Institute for Infection, Inflammation and Immunity, Radboud University Nijmegen Medical Center, the Netherlands. I was responsible for kinetic studies of the mechanism of action of GPC-N114. Dr. Jamie Arnold helped me with experimental design and related discussions.

5.1 Introduction

Coxsackieviruses and cardioviruses are very common and significant human and animal pathogens. They have brought serious public health issues and great loss to the farming industry. Coxsackieviruses are divided into two major groups. Group A has at least 23 serotypes and mainly infects human mucous membranes and human skin leading to conjunctivitis, hand-foot-and-mouth disease and herpangina\(^1\). Group B has only 6 serotypes but infects multiple human organs including heart, pleura, pancreas and liver. Group B viruses can cause viral meningitis, viral myocarditis, pleurodynia, pancreatitis and hepatitis, which is unlike hepatitis by hepatotropic virus\(^5\). Recently, insulin-dependent diabetes mellitus type 1 was also found to be associated with coxsackievirus B infection\(^8\). These viruses are able to infect humans of all age groups but cause more acute and severe syndromes in neonates and children. In some Asian countries, frequent family outbreaks and community epidemics by coxsackieviruses have been reported. For
example, according to the China Post in May 2015, an outbreak of a coxsackievirus A infection in Taiwan caused more than 600 patients to seek emergency treatment in hospitals within one week. There is also a potential for global epidemics due to global transportation and import of food contaminated with those viruses. In contrast to human infections by coxsackieviruses, cardiovirus A mainly infects animals and has only two serotypes of encephalomyocarditis virus (EMCV)\textsuperscript{10,11}. Unfortunately, we are not sufficiently prepared for outbreaks of coxsackieviruses and cardioviruses despite great efforts towards the development of antiviral drugs and vaccines.

Inhibition of host or viral proteins required for viral growth may suppress replication of coxsackieviruses and cardioviruses. Coxsackieviruses belong to the genus of enteroviruses. Both enteroviruses and cardioviruses are positive sense, single stranded RNA viruses in the family of \textit{Picornaviridae}. Their genomes are usually around 7 to 8 k base pairs (bp). Once a virus enters into a host cell, cellular ribosomes bind to the internal ribosomal entry site (IRES) at the 5’ end of the viral genome to initiate the production of the viral polyprotein\textsuperscript{12,13}. Following translation, the polyprotein is proteolytically processed into structural proteins (VP1 to VP4) and nonstructural proteins (2A, 2B, 2C\textsuperscript{NTPase}, 3A, 3B (VPg, viral protein genome), 3C\textsuperscript{pro}, and 3D\textsuperscript{pol} (RNA-dependent RNA polymerase (RdRp)), and their precursors)\textsuperscript{12,13}. The virus efficiently utilizes those viral proteins and borrows other host factors to complete subsequent genome replication, virion assembly, and maturation\textsuperscript{12,13}. For example, during genome replication, the RdRp assembles with other viral proteins (e.g. 3CD\textsuperscript{14}) and host proteins (e.g. poly(rC)-binding protein\textsuperscript{14}) to form a RNA replication complexes, which are located at cellular membranes in the host cell. The replication complex then initiates RNA synthesis by VPg uridylylation and carries out RNA elongation along the template\textsuperscript{15-18}.

Many laboratories have discovered inhibitors that target either host proteins or virally encoded proteins that may help treat coxsackievirus and cardiovirus infections. Here, I use antiviral studies of coxsackievirus B3 (CVB3) as an example. For the inhibitors against host
proteins, inhibitory effects of cellular reactions (i.e. histone acetylation\textsuperscript{19}, ubiquitin proteasome pathway\textsuperscript{20}) on CVB3 replication have been evaluated. However, inhibition of these host reactions may lead to serious cytotoxicity. Inhibitors against viral proteins include those that target capsid protein (i.e. pleconaril and derivatives\textsuperscript{21}), 2C (i.e. TBZE-029\textsuperscript{22}, guanidine hydrochloride\textsuperscript{23}, and MRL-1237\textsuperscript{24}), 3A (i.e. TTP-8307\textsuperscript{25}, enviroxime-like compound\textsuperscript{26}, and itraconazole\textsuperscript{27}), 3C (i.e. AG7088-like compound\textsuperscript{28} and benzserazide\textsuperscript{29}), and the RNA-dependent RNA polymerase (RdRp) (i.e. ribavirin\textsuperscript{30}, 5’ substituted cytosine analogs\textsuperscript{31}, and amiloride\textsuperscript{32-34}). Among those inhibitors, the RdRp inhibitors are the most attractive, since RdRp performs the core function in genome replication and is the most conserved protein among RNA viruses. EMCV inhibitors are much less studied and only PTU-23 has been reported to target RdRp, but has an unknown inhibitory mechanism\textsuperscript{35}.

RdRp inhibitors can be classified into two categories, those being nucleoside analogs and non-nucleoside inhibitors (NNI). Ribavirin and 5’ substituted cytosine derivatives belong to natural nucleoside analogs that can be incorporated into the viral genome by the RdRp to eventually cause lethal mutagenesis (i.e. due to the accumulation of an excessive amount of mutations in the viral genomes). Amiloride belongs to the NNI class and has broad-spectrum antiviral activity against enteroviruses, coronaviruses, hepatitis C viruses, and the human immunodeficiency viruses (HIV) through various mechanisms\textsuperscript{32-34, 36-39}. One of its inhibitory mechanisms is to act as a competitive inhibitor against nucleotide and catalytic metal ion binding to CVB3 RdRp\textsuperscript{34}. The potency of amiloride is mild with an EC\textsubscript{50} value at the micromolar level in cell culture studies\textsuperscript{32}. Except amiloride, mechanisms of action of other inhibitors against CVB3 RdRp remain poorly understood.

NNIs may have advantages over that of nucleoside analogs in terms of specificity and diversity of antiviral drugs. Although nucleoside analogs usually have a broad viral spectrum, they may also be utilized by host polymerases due to structural and functional similarities among
nucleic acid polymerases, and lead to side effects. NNIs may have higher specificity because they do not generally target the active site, but target other sites of the viral polymerases. These non-active-site binding sites usually differ significantly from cellular polymerases, thus increasing specificity. Additionally, the different mechanisms of action of NNIs expand the diversity of antiviral drugs, which has important clinical implications. For example, treatment of patients with a particular drug may fail due to resistance against that particular drug, but the viral disease may be treated with an NNI with a completely different molecular mechanism. Therefore, the discovery and study of new NNIs are of great significance.

Recently, a novel NNI, GPC-N114 (2,2’-[(4-chloro-1,2-phenylene)bis(oxy)]bis(5-nitro-benzonitrile)), was discovered with broad-spectrum antiviral activities against enteroviruses and cardiовiruses\(^4^0\). Cell-based studies indicated that GPC-N114 is a potent inhibitor with an EC\(_{50}\) value at or close to the submicromolar level for enteroviruses and cardiовiruses in cytopathic reduction assays\(^4^0\). GPC-N114 inhibited viral multiplication and RNA synthesis less efficiently than guanidine hydrochloride for CVB3 and but as efficiently as dipyridamole for EMCV in kinetic studies of viral growth and subreplicon luciferase assays\(^4^0\). The discovery of resistant mutants were consistent with interactions between RdRp and GPC-N114 in EMCV\(^4^0\). Crystal structures of the CVB3 RdRp-GPC-N114 complexes suggested that GPC-N114 is a template channel inhibitor that prevents RNA from binding to the RdRp\(^4^0\).

Here, I used biochemical methods in an attempt to attain a more detailed mechanism of action of GPC-N114 on CVB3 RdRp. Fluorescent polarization studies indicated that CVB3 RdRp had similar RNA binding affinity as the PV and FMDV RdRps but the IC\(_{50}\) value of GPC-N114 against RNA binding was unable to be detected due to the intrinsic fluorescence of this inhibitor. Homopolymeric elongation assays showed that GPCN-114 inhibited the activity of CVB3 RdRp in a concentration dependent manner. But the same inhibitory effects of GPCN-114 were not
observed in sym/subU RNA assays and template switching assays. These observations suggested that reaction conditions and setup in these assays need to be re-evaluated.

5.2 Methods

5.2.1 Materials

\[\gamma^{32}\text{P}]ATP \text{ and } [\alpha^{32}\text{P}]UTP (>7000\text{Ci/mmol}) \text{ were from VWR-MP Biomedical.} \]

Nucleoside 5’-triphosphates and 2’-deoxynucleoside 5’-triphosphates (all NTPs were ultrapure solutions) were from GE Healthcare. Oligonucleotide dT\textsubscript{15} was from Integrated DNA Technologies, Inc. All other RNA oligonucleotides were from Dharmac\textsuperscript{on} Research, Inc. (Boulder, CO). T4 polynucleotide kinase was from New England Biolabs, Inc. All other reagents were of the highest grade available from Sigma or Fisher.

5.2.2 Purification, 5’-\textsuperscript{32}P end labeling, and annealing of sym/subU RNA

For kinetic studies, RNA oligonucleotide sym/subU was gel purified, deprotected, radioactively labeled, and reannealed as described in a previous chapter (see Methods in Chapter 2).

5.2.3 CVB3 RdRp assay

Reactions contained 50 mM Tris 7.0, 10 mM KCl, and 0.8 mM MgCl\textsubscript{2}, and were conducted at 30°C for the entire reaction time course. CVB3 WT was diluted in enzyme dilution buffer (5 mM HEPES 7.0, 150 mM NaCl, and 50% glycerol) immediately before use and was
added to one tenth of the total reaction volume. GPC-N114, dissolved in DMSO, was added to one twentieth of the total reaction volume. Final concentrations of CVB3 WT RdRp, sym/subU RNA or sym/subU RNA modified with 6-carboxyfluorescein at the 5’ end of the oligonucleotide (6-FAM-sym/subU RNA), nucleotides, and GPC-N114 are indicated in the appropriate figure legend. Reactions were quenched by the addition of EDTA to a final concentration of 25 mM at each time point. Products were preheated at 65 °C for 5 min and were analyzed via a highly crosslinked denaturing PAGE gel at constant 90 watts for 2.5 h. The gel was exposed to screen, visualized by PhosphorImager, and quantified by ImageQuant software (Molecular Dynamics). The data was fit into different functions by KaleidoGraph software (Synergy).

5.2.4 Homopolymeric template elongation assay

Reaction conditions in the homopolymeric template elongation assay were the same as the CVB3 RdRp assay as described in section 5.2.3. The primer/template system was replaced with dT15/poly(rA). Reactions contained 20 nM CVB3 WT RdRp, 0~80 µM GPC-N114, 10 µM α 32P-labeled UTP* (2 µCi), 2 µM dT15, and 350 nM poly(rA) for UTP incorporation. For order-of-addition reactions, reactions contained 20 nM CVB3 WT RdRp, 80 nM dT15, 14 nM poly(rA), 10 µM α 32P-labeled UTP* (2 µCi), and 0~100 µM GPC-N114. Reactions were quenched by the addition of 100 mM EDTA at each time point. Samples with 5 µL were loaded onto DE81 filter paper discs and were dried completely. The unincorporated UTP was removed from the discs by shaking in 250 mL of 5% sodium dibasic phosphate for 10 min three times. The discs were dried with absolute ethanol. The corresponding radioactivity was read-out by a liquid scintillation counter in 5 mL Ecoscint scintillation fluid (National Diagnostics).
5.2.5 Template switching assay

Reaction conditions were the same as the CVB3 RdRp assay as described in section 5.2.3. 5 µM CVB3 WT RdRp and 10 µM dT15/rA30 were incubated in 30°C for 3 min to allow formation of CVB3-RNA complexes. The complexes then diluted by 10 fold with final concentrations of CVB3 WT RdRp, dT15/rA30, and UTP to be 0.5, 1 and 500 µM (2 µCi) respectively. Final concentrations of heparin, rA30, and GPC-N114 were indicated in the corresponding figure legend. Reactions were quenched by the addition of 100 mM EDTA. Products were quantitated by a liquid scintillation counter as described in section 5.2.4.

5.2.6 Fluorescence polarization experiments

Reaction conditions were the same as the CVB3 RdRp assay except that salt concentrations were fixed to 10 mM KCl and 15 mM NaCl. Reactions were conducted at room temperature. Different concentrations of enzyme were added immediately from ice, and mini-polarization values were measured after 30s or every minute from the beginning of the kinetic measurement. 6-FAM-sym/subU RNA was fixed to 4 nM. GPC-N114 in 100% DMSO was added to one twentieth of the total reaction volume. Final concentrations of RdRp and GPC-N114 are indicated in the appropriate figure legend.

5.3 Results

5.3.1 6-FAM-sym/subU RNA allows nucleotide incorporation by CVB3 WT RdRp.

Cell culture studies indicated that benzonitrile-containing inhibitor GPC-N114 suppresses CVB3 replication via the inhibition of RNA synthesis. Crystal structure analysis of the CVB3
RdRp complexed with GPC-N114 suggested that GPC-N114 occupies the RNA template-binding channel of RdRp to prevent the native template from binding. To confirm this hypothesis, fluorescence polarization (FP) was applied to investigate how the RNA binding affinity of the RdRp is affected by GPC-N114. FP is a direct and real time measurement of the molecular binding event, and depends on how fast the fluorescently labeled molecule is tumbling in solution. Here, nucleic acid sym/subU RNA was fluorescently labeled by the attachment of a fluorophore to the 5’ end to form 6-FAM-sym/subU RNA. The fast rotating free and slowly rotating bound 6-FAM-sym/subU RNA molecules give small and large component polarization values respectively. Therefore, FP was used to monitor the binding affinity between the 6-FAM-sym/subU RNA and the RdRp in the presence and absence of GPC-N114.

One concern with these assays is that the fluorophore modification may prevent 6-FAM-sym/subU RNA from binding to the RdRp. As such, I conducted nucleotide incorporation assay for both CVB3 and PV WT RdRps using the 6-FAM-sym/subU RNA. Catalytic activities of CVB3 WT RdRps stored in the buffer with 50% glycerol (shown as CVB3 in all the following figure) or without glycerol (shown as CVB3NG in the following figure) were also compared in this assay. The results indicated that 6-FAM-sym/subU RNA did not affect the nucleotide incorporation rate and that the storage buffer with 50% glycerol gave rise to a higher enzyme activity (Figure 5-1). In the following assays, CVB3 WT RdRp was kept and diluted in enzyme dilution buffer containing 5 mM HEPES 7.0, 50% glycerol, and 150 mM NaCl.
Figure 5-1. 6-FAM-sym/subU RNA can be utilized for nucleotide incorporation by CVB3 WT RdRp. Reactions contained 50 mM Tris 7.0, 10 mM KCl, and 0.8 mM MgCl$_2$. Either radioactive sym/subU RNA (A) or fluorescent 6-FAM-sym/subU RNA (B) was incubated with CVB3 RdRp for 90 s at which time, reactions were initiated by the addition of ATP. The reaction was quenched by addition of EDTA to a final concentration of 25 mM. Final concentrations of RNA, RdRp, and ATP were 1, 1, and 500 µM respectively. PV RdRp was used as a control. CVB3 WT RdRps stored in 50% glycerol (lane CVB3) or without glycerol (lane CVB3NG) were also compared for enzyme activity.
5.3.2 GPC-N114 contributes to fluorescence and prevents quantitative inhibition studies.

To determine the dissociation constant (K_D) for 6-FAM-sym/subU RNA binding to CVB3 WT RdRp, both kinetic and titration measurements of nucleic acid binding were performed. It should be noted that in kinetic measurements, the fluorescence intensity and the mini-polarization of the free RNA stayed constant for at least 20 min, which suggested that 6-FAM-sym/subU RNA was sufficiently stable for subsequent titration experiments (Figure 5-2).

The 6-FAM-sym/subU RNA was titrated with CVB3 WT RdRp at various concentrations. Reactions contained 50 mM Tris 7.0, 0.8 mM MgCl_2, 10 mM KCl, 15 mM NaCl, 4 nM 6-FAM-sym/subU RNA, and CVB3 WT RdRp. CVB3 WT RdRp was added immediately from ice, and the mini-polarization value was measured after 30 s. The K_D value for CVB3 WT RdRp was 1.3 ± 0.3 µM (Figure 5-2). However, binding saturation was not achieved likely due to enzyme precipitation. The K_D value for RNA binding to PV WT RdRp (in the same enterovirus genus) and foot-and-mouth disease virus (FMDV) WT RdRp (in aphthovirus genus but from the same family of Picornaviridae) were also measured. It might be predicted that the K_D values for the three polymerases would be similar because of structural and functional similarities of the RdRps. Indeed, the K_D for CVB3 WT RdRp was comparable with PV (1.2 µM ± 0.3) and FMDV (0.7 µM ± 0.2) WT RdRps (Table 5-1).
Figure 5-2. The dissociation constant of 6-FAM-sym/subU RNA to CVB3 WT RdRp is at the micromolar level. A. Kinetic measurements of 4 nM 6-FAM-sym/subU RNA were performed in a reaction buffer containing 50 mM Tris 7.0, 10 mM KCl, and 0.8 mM MgCl₂. B. 6-FAM-sym/sub RNA binding to CVB3 WT RdRp was measured under the same condition as before (see Figure 5-1) except that the salt concentration was fixed to 10 mM KCl and 15 mM NaCl. The concentration of CVB3 WT RdRp used to titrate RNA varied from 10 to 1700 nM. Measured mini-polarization values are plotted as a function of CVB3 WT RdRp concentration. The solid line represents the fit of the data to a hyperbola function. The $K_D$ value equals $1.3 \times 10^3 \pm 0.3 \mu$M. Experiments were conducted in triplicate.
Table 5-1. Comparison of the dissociation constants for 6-FAM-sym/subU RNA binding to CVB3, PV, and FMDV WT RdRps revealed that 6-FAM-sym/subU RNA has similar binding affinity for all RdRps. The reaction buffer was the same as before (see Figure 5-1) except that the salt concentration was fixed to 10 mM KCl and 15 mM NaCl.

<table>
<thead>
<tr>
<th>RdRp</th>
<th>$K_D$ (µM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CVB3</td>
<td>$1.3 \pm 0.3 \times 10^3$</td>
</tr>
<tr>
<td>PV</td>
<td>$0.7 \pm 0.2 \times 10^3$</td>
</tr>
<tr>
<td>FMDV</td>
<td>$1.2 \pm 0.3 \times 10^3$</td>
</tr>
</tbody>
</table>
If the GPC-N114 compound acted as an RdRp inhibitor by binding to the RNA template channel as suggested by the van Kuppeveld lab, GPC-N114 should disturb the interaction between CVB3 WT RdRp and 6-FAM-sym/subU RNA in such a way that the presence of the compound would result in a reduction of the mini-polarization value in a concentration dependent manner. To test this hypothesis, the fluorescence of GPC-N114 in the absence or the presence of 6-FAM-sym/subU RNA or in the presence of both 6-FAM-sym/subU RNA and CVB3 WT RdRp was measured. The mini-polarization values were measured either after GPC-N114 was added or after GPC-N114 was incubated for 1 min followed by the addition of CVB3 WT RdRp (Table 5-2). I should mention that solvent DMSO was used to dissolve GPC-N114, but DMSO did not disturb the mini-polarization value for 6-FAM-sym/subU RNA under the reaction conditions (data not shown). The results indicated that GPC-N114 contributed to fluorescence in all situations (Table 5-2). Surprisingly, the mini-polarization values increased from 183.2 to 303.4 when RdRp and GPC-N114 were added and GPC-N114 concentration increased from 1 to 100 µM, which suggested that GPC-N114 itself contributed to the signal (Table 5-2). The fluorescence intensity also increased by two fold when GPC-N114 was added up to 100 µM (Table 5-2). These observations indicated that the intrinsic fluorescence of GPC-N114 prevents any reliable IC$_{50}$ measurement.
Table 5-2. Comparison of the effects of GPC-N114 on fluorescent intensity and mini-polarization in three different reaction conditions. Reaction buffer was the same as before (see Figure 5-1) except that salt concentration was fixed to 10 mM KCl and 15 mM NaCl. Both intensity and mini-polarization (mP) were increased with increasing concentration of GPC-N114 in the absence of RNA, or in the presence of RNA, or both RNA and CVB3 WT RdRp. The GPC-N114 concentration ranged from 0.1 to 100 µM. Final concentrations of 6-FAM-sym/subU and CVB3 WT RdRp were fixed to 4 nM and 1.3 µM respectively.

<table>
<thead>
<tr>
<th>Rxn</th>
<th>Sample</th>
<th>Intensity</th>
<th>mP</th>
</tr>
</thead>
<tbody>
<tr>
<td>GPC-N114</td>
<td>1µM GPC-N114</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>10µM GPC-N114</td>
<td>1.252</td>
<td>713.4</td>
</tr>
<tr>
<td></td>
<td>100µM GPC-N114</td>
<td>17.48</td>
<td>522.4</td>
</tr>
<tr>
<td>RNA➔GPC-N114</td>
<td>RNA</td>
<td>16.37</td>
<td>40.9</td>
</tr>
<tr>
<td></td>
<td>RNA+1µM GPC-N114</td>
<td>15.08</td>
<td>50.0</td>
</tr>
<tr>
<td></td>
<td>RNA+10µM GPC-N114</td>
<td>17.44</td>
<td>108.0</td>
</tr>
<tr>
<td></td>
<td>RNA+100µM GPC-N114</td>
<td>33.97</td>
<td>312.7</td>
</tr>
<tr>
<td>RNA➔GPC-N114</td>
<td>RNA+water+CVB3</td>
<td>26.54</td>
<td>171.8</td>
</tr>
<tr>
<td>RNA➔CVB3</td>
<td>RNA+DMSO+CVB3</td>
<td>29.55</td>
<td>182.5</td>
</tr>
<tr>
<td></td>
<td>RNA+0.1µM GPC-N114+CVB3</td>
<td>29.82</td>
<td>183.2</td>
</tr>
<tr>
<td></td>
<td>RNA+1µM GPC-N114+CVB3</td>
<td>29.96</td>
<td>181.7</td>
</tr>
<tr>
<td></td>
<td>RNA+10µM GPC-N114+CVB3</td>
<td>33.11</td>
<td>209.6</td>
</tr>
<tr>
<td></td>
<td>RNA+100µM GPC-N114+CVB3</td>
<td>51.09</td>
<td>303.4</td>
</tr>
</tbody>
</table>
5.3.3 GPC-N114 competes against template binding in the homopolymeric template assay.

Considering the challenges with the fluorescence polarization assays, we sought an alternative method for probing the molecular mechanism of GPC-N114. As such, I assayed the effect that GPC-N114 has on UMP incorporation into the homopolymeric A template. Reactions contained 20 nM CVB3 WT RdRp, 2 µM dT15 annealed to 350 nM poly(rA), 10 µM UTP*, and GPC-N114 (0 to 80 µM). Reactions were initiated by the addition of UTP* (Figure 5-3). The rate of UTP incorporations dropped by three-fold when GPC-N114 was increased to 80 µM (Figure 5-3).

To test whether the rate reduction in UMP incorporation was due to GPC-N114 perturbing RNA binding, order-of-addition assays were conducted. In the order-of-addition assays, either the primer/template or GPC-N114 with a specific concentration was first added prior to the addition of CVB3 WT RdRp (Figure 5-4). After incubating for 1 min, the other substrate was added. Reactions were then initiated by the addition of UTP. When GPC-N114 was first incubated with CVB3 WT RdRp, the enzyme activity dropped by three fold (Figure 5-4). However, there was little drop in activity when the primer/template was bound to CVB3 WT RdRp first (Figure 5-4). These observations suggested that GPC-N114 was able to efficiently block the RNA binding channel only when GPC-N114 was able to access the RdRp before RNA binding. In contrast, the binding of GPC-N114 may be too weak to outcompete any RNA already bound to the RdRp. These results are also consistent with the previous crystal structure analysis demonstrating that GPC-N114 binds to the template channel of CVB3 WT RdRp.

Based on the crystal structure of the RdRp-GPCN114 complex, we would predict that GPC-N114 does not interact with the binding channel of incoming UTP. To test this hypothesis, experiments were compared between incubating UTP before and after the addition of GPC-N114. Surprisingly, there was little enzyme inhibition by GPC-N114 when UTP was incubated in the
reaction mix (Table 5-3). This finding suggested that nucleotide indirectly affects the binding affinity of GPC-N114 to the polymerase. Further experiments needed to be conducted to reveal a more detailed mechanism.
Figure 5-3. GPC-N114 inhibits UTP incorporation into the homopolymeric template. Reactions contained 50 mM Tris 7.0, 10 mM KCl, and 0.8 mM MgCl₂. Final concentrations of CVB3 WT RdRp, poly(rA)/dT15, and UTP* were 20 nM, 350 nM/2 µM, and 10 µM (2 µCi) respectively. Concentration of GPC-N114 ranged from 0 to 80 µM. CVB3 WT RdRp was first incubated with RNA and GPC-N114 for 2 min, at which time UTP was added to initiate the reaction. Reactions were quenched by the addition of 50 mM EDTA. A. Reaction scheme for UTP incorporation into the homopolymeric template. B. GPC-N114 inhibited UTP incorporation in a concentration dependent manner. The solid lines represent fits of the data to linear curves. C. The initial velocities obtained from panel B are plotted as a function of GPC-N114 concentration.
Table 5-3. Differences in GPC-N114 inhibition when UTP initiated the reaction or was preincubated with the enzyme prior to initiation suggested that UTP disturbs GPC-N114 binding to CVB3 WT RdRp although GPC-N114 does not occupy the nucleotide binding pocket as shown in the crystal structures of CVB3 WT RdRp-GPC-N114 complexes.

<table>
<thead>
<tr>
<th>GPC-N114</th>
<th>Activity (cpm/s)</th>
<th>UTP initiation</th>
<th>UTP incubation</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>2.1 ± 0.8</td>
<td>1.3 ± 0.3</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>1.6 ± 0.4</td>
<td>1.9 ± 0.1</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>1.8 ± 0.4</td>
<td>1.9 ± 0.1</td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>1.0 ± 0.4</td>
<td>1.7 ± 0.3</td>
<td></td>
</tr>
<tr>
<td>80</td>
<td>0.4 ± 0.1</td>
<td>1.0 ± 0.1</td>
<td></td>
</tr>
</tbody>
</table>
Figure 5-4. Results of order-of-addition reactions indicated that GPC-N114 inhibits the reaction more efficiently when it was incubated with CVB3 WT RdRp before the addition of RNA. Reactions contained 50 mM Tris 7.0, 10 mM KCl, and 0.8 mM MgCl₂. Final concentrations of CVB3 WT RdRp, poly(rA)/dT15, and UTP* were 20 nM, 14 nM/80 nM, and 10 µM (2 µCi) respectively. The concentration of GPC-N114 ranged from 0 to 100 µM. CVB3 WT RdRp was either incubated with GPC-N114 before the addition of RNA (A) or was incubated with RNA prior to the addition of GPC-N114 (B). Reactions were initiated by the addition of UTP and were quenched by the addition of 50 mM EDTA. C. Percentages of activity relative to the control are plotted as a function of the concentration of GPC-N114.
5.3.4 GPC-N114 has no inhibitory effects on the binding of sym/subU RNA to CVB3 WT RdRp.

We were also interested in the effects of GPC-N114 on sym/subU RNA binding. The sym/subU RNA has been widely used in kinetic studies of RdRps from positive strand RNA viruses, especially in the model system of PV RdRp\(^{41-44}\). Interactions between the sym/subU RNA and RdRp mimic the chain elongation complex of RdRp \textit{in vivo}\(^{45}\). Additionally, results associated with the kinetic studies of sym/subU RNA are consistent with many biological studies\(^{45}\). To look at whether GPC-N114 inhibited sym/subU RNA binding to CVB3 WT RdRp, I conducted four different assays at 30 °C. Reactions contained 50 mM Tris 7.0, 10 mM KCl, 0.8 mM MgCl\(_2\), 1 µM CVB3 WT RdRp, 1 µM sym/subU, 500 µM NTPs/ATP, and GPC-N114 from 1 to 100 µM. The four order-of-addition reactions included (1) addition of CVB3 WT RdRp into ATP, sym/subU, and GPC-N114, (2) addition of CVB3 WT RdRp into NTPs and GPC-N114 prior to reaction initiation with sym/subU RNA, (3) addition of CVB3 WT RdRp into sym/subU RNA and GPC-N114 prior to initiation with ATP, (4) addition of CVB3 WT RdRp into GPC-N114 prior to addition of sym/subU RNA and NTPs (Figure 5-5). In contrast to the previous homopolymeric template assays, little to no inhibition was observed in any of the four reactions even when GPC-N114 was added up to a concentration of 100 µM. The absence of any significant enzyme inhibition suggested that there were either inappropriate reaction conditions for observing GPC-N114 inhibition, or there were structural and mechanistic differences between the homopolymeric RNA and the sym/subU RNA (Figure 5-5).
Figure 5-5. GPC-N114 is not competitive against sym/subU RNA. Reactions contained 50 mM Tris 7.0, 10 mM KCl, and 0.8 mM MgCl₂. Final concentrations of CVB3 WT RdRp, sym/subU RNA, and ATP/NTPs were 1, 1, and 500 µM respectively. The concentration of GPC-N114 ranged from 0 to 100 µM. Reactions were conducted by adding CVB3 WT RdRp to initiate reactions (A, B), or by incubating CVB3 WT RdRp with sym/subU RNA and GPC-N114 followed by initiation with ATP addition (C, D), or by incubating CVB3 WT RdRp with NTPs and GPC-N114 followed by initiation with sym/subU RNA addition (E, F) or by incubating CVB3 WT RdRp with GPC-N114 followed by initiation with addition of sym/subU RNA and NTPs (G, H).
5.3.5 GPC-N114 has no effects on template switching.

To test whether GPC-N114 inhibits RNA template switching, a template switching assay in the presence or absence of 100 µM GPC-N114 was conducted. In this assay, 5 µM CVB3 WT RdRp was pre-assembled with 10 µM dT15/rA30 in the presence of 500 µM UTP for 3 min. Then, the enzyme complexes were diluted by ten fold (0.5 µM CVB3 WT RdRp and 1 µM dT15/rA30) in the presence of 500 µM UTP, 10 µM heparin (to trap free CVB3 WT RdRp), 0/50 µM rA30 acceptor, and 0/100 µM GPC-N114. The addition of an excess amount of rA30 acceptor should drive template switching to yield more UTP incorporation than the control (Figure 5-6). If GPC-N114 inhibited template switching, the UTP incorporation with excess rA30 should return to the control level in the presence of GPC-N114. No obvious inhibition was observed even when GPC-N114 was at a concentration of 100 µM (Figure 5-6).
Figure 5-6. GPC-N114 does not show inhibition of RNA template switching. A. Reaction scheme for RNA template switching. Reactions contained 50 mM Tris 7.0, 10 mM KCl, and 0.8 mM MgCl$_2$. CVB3 WT RdRp (5 µM) and dT15/rA30 (10 µM) were incubated in the presence of UTP* (500 µM, 2 µCi) for 3 min. The formed complexes were diluted by 10 fold in reaction buffer containing 500 µM UTP, 10 µM heparin, and with (solid square) or without (circle) 50 µM rA30 in the absence (B) or presence (C) of 100 µM GPC-N114.
5.4 Discussion

Recently, a novel antiviral inhibitor GPC-N114 with broad-spectrum activity against multiple members of enteroviruses and cardioviruses was discovered\textsuperscript{40}. It was suggested that GPC-N114 inhibits RNA replication by blocking the template-binding channel of the RdRp\textsuperscript{40}. To test the hypothesis, FP experiments were conducted by titrating the CVB3 WT RdRp-RNA complexes with GPC-N114. Unfortunately, the intrinsic fluorescence of GPC-N114 also contributed to the mini-polarization value, preventing any accurate measurement of the IC\textsubscript{50} value for this inhibitor (Table 5-2). Further studies indicated that both 6-FAM-sym/subU RNA and GPC-N114 were excited at 490 nm (which is the theoretical excitation light for fluorescein). GPC-N114 likely increased mini-polarization values and interfered with the results. In the future, a new fluorophore, whose excitation wavelength does not overlap with that of FAM, is required for these measurements. For example, Alexa Fluor 594 (a fluorophore that has an excitation maximum at 590 nm and emission maximum at 617 nm) is an ideal fluorophore according to emission tests that showed that GPC-N114 barely has emission when it is excited at 594 nm.

According to previous cell-based studies, the EC\textsubscript{50} was 0.15 µM. However, 100 µM GPC-N114 was required to inhibit only 70% elongation in the homopolymeric template assay (Figure 5-3). The significant difference in the effective concentration of GPC-N114 between cellular and kinetic studies may be due to the following reasons. Firstly, in terms of length and sequence of RNA, the poly(rA)/dT15 template/primer system is artificial and very different from the viral genome that can be relatively easily blocked by GPC-N114. Secondly, the physiological conditions (such as pH, ionic strength and protein composition) in a cell may facilitate binding of GPC-N114 to RdRp more than the reaction conditions I used here. Thus, a lower amount of GPC-N114 may be needed to inhibit RNA synthesis in a cell. Additionally, besides the RdRp, the viral genome replication \textit{in vivo} relies on other viral proteins and host factors to form the replication
machinery, whose structure and environment may allow GPC-N114 to bind to the template channel of the RdRp more easily than the single RdRp in vitro. GPC-N114 in the cell may also undergo some enzymatic modifications prior to binding to the active site of RdRp. However, the active form of GPC-N114 with greater affinity cannot be achieved in vitro. Lastly, the inhibition of GPC-N114 might be magnified via interfering with other RdRp directed events such as Vpg uridylylation, template translocation, and template recombination during genome replication in a cell. In these events, GPC-N114 might present distinct inhibitory effects due to differences in template identity and affinity. We have shown that GPC-N114 is able to prevent primer elongation but has no effects on template switching. It is also interesting to know whether GPC-N114 interferes with the other RdRp events.

Surprisingly, the inhibitory effects of GPC-N114 were different in the assays with the poly(rA)/dT15 template/primer system compared to those with sym/subU RNA (Figure 5-4 and 5-5). The sym/subU RNA is widely used in kinetic experiments of viral RdRps to mimic the elongation complex, and results with the sym/subU RNA have been related to what occurs in cellular and animal studies. Considering the low EC_{50} of GPC-N114 against CVB3 replication and the ability of GPC-N114 to block the homopolymeric template binding in vitro (Figure 5-4), I proposed that GPC-N114 is still able to prevent sym/subU RNA binding to RdRp, but these effects are shielded due to impropriate reaction conditions and assay designs. In sym/subU RNA assays, the concentration of CVB3 RdRp (1 µM) is much higher than that in the homopolymeric template assay (20 nM). The high concentration of sym/subU RNA (1 µM) in the reaction system may result in no change in the rate in the presence of GPC-N114. It is possible that the sym/subU RNA has stronger affinity than the homopolymeric template. The higher concentration of the sym/subU RNA drives the equilibrium to RdRp-RNA binary complexes and the dissociation of GPC-N114 from the RdRp. Decreasing concentrations of sym/subU RNA applied in the reaction may help to observe the inhibitory effects of GPC-N114. Additionally, GPC-N114 may be a
slow-binding inhibitor, and may not reach equilibrium in the 2 min incubation time. For example, in the assay, GPC-N114 was preincubated with CVB3 RdRp for 2 min followed by the addition of sym/subU RNA and NTPs to initiate the reaction. No inhibition was observed even when GPC-N114 was added up to 80 µM. If GPC-N114 is a slow-binding inhibitor, it may not be able to fully establish the equilibrium with RdRp within 2 min and thus NTP incorporation will not be inhibited. Lengthening the preincubation time (i.e. titrate time from 2 to 30 min) between CVB3 RdRp and GPC-N114 prior to initiation may be required for maximizing GPC-N114 inhibition. Lastly, incomplete equilibrium between RdRp and GPC-N114 may also occur in the homopolymeric template assay, but the reaction may be more sensitive to inhibition owing to the lower amount of RdRp in the assay.

GPC-N114 is the first reported inhibitor targeting the template channel of RdRps in picornaviruses. Similar inhibitors targeting nucleic acid substrate binding have been discovered for HIV, hepatitis C virus and dengue virus\textsuperscript{46-48}. Besides traditional active site and allosteric binding sites, the template channel is a brand-new target for antiviral drug discovery. The template channel based antiviral drug may benefit patients who cannot be treated with existing therapeutic options, due to its novel mechanism of inhibition.

5.5 References

9. Cna. Taiwan sees sharp increase in enterovirus cases: CDC.


CHAPTER 6
CONCLUSIONS AND FUTURE DIRECTIONS

6.1 Conclusions

Outbreaks of RNA viruses and the emergence of new viruses have been severe issues in many aspects of human activity including health care, the economy, food security, and social stability. A decade ago, an outbreak of the severe acute respiratory syndrome (SARS) coronavirus, which originated from the Guangdong province in China and later spread globally, caused more than 8000 infections and 812 deaths\(^1\). Some of the survivors from this outbreak have also suffered long-term complications, such that many have been forced to quit their jobs. This has led to new social issues. This outbreak also caused great economic losses, estimated around 40 billion dollars\(^2\). More recently, outbreaks of another coronavirus, the Middle East respiratory syndrome (MERS) virus, have brought new public panics. In Saudi Arabia, MERS caused more than 1000 infections and led to more than 35% fatalities\(^3\). In May 2015, MERS infected 186 people and caused 36 deaths in South Korea\(^4\). Besides coronaviruses, many other RNA viruses also cause severe public health issues and economic loss. In 2011, a coxsackievirus A infection sent 3000 patients to the hospital in Japan\(^5\). In 2001, an outbreak of foot-and-mouth disease among farming animals caused agricultural and food-chain losses up to 4.8 billion US dollars in Great Britain\(^6\). The frequent occurrence of these infectious diseases derives from many reasons including, but not limited to, the unknown identity of RNA viruses, ambiguous syndromes, global transportation, unavailability or limited efficacy of antiviral strategies, high mutation rates of RNA viruses and variations in quarantine measures.
Great efforts have been made towards the development of antiviral drugs and viral vaccines. Drug targets include many viral and host proteins. Among these drug targets, the polymerase of RNA viruses may be the most important target, not only because this enzyme synthesizes viral genomes in an infected cell, but also because the high mutation rate of the replication machinery is tightly linked to many critical viral behaviors including pathogenicity, cross-species transmission, cell tropism, clinical syndrome, antibody escape, and drug resistance. It has been suggested that the polymerase fidelity of RNA viruses plays an important role for viral adaptation. A population of RNA viruses can be described as a ‘quasi-species’ in which members have genetically related genomes that differ by one or more mutations. The genetic diversity of the quasispecies allows rapid selection of viral escape mutants to regain fitness, say when the viruses are challenged by the immune system or antiviral drugs. The genetic diversity of the viral population mainly derives from the functions of the viral polymerase. Many studies have indicated that either an increase or a decrease in viral fidelity can lead to viral attenuation in cell cultures or in animal models. The presence of nucleoside analogues like ribavirin can increase the mutation rate, and lead to error catastrophe. Viruses encoding polymerases with higher fidelity are also attenuated, likely due to the reduced genetic diversity within the viral population. These studies indicated that the nucleotide selection mechanism of the RNA-dependent RNA polymerase is closely tied to viral pathogenicity. A more complete understanding of the nucleotide selection mechanism would thus be important for the development of antiviral strategies.

In this dissertation, I studied the impacts of the motif D equilibrium on enzyme fidelity, the effects of the four substitutions found in the Sabin 1 vaccine strain, the contribution of motif F to enzyme fidelity, and the mechanism of action of inhibitor GPC-N114 that targets the RNA template channel.
In Chapter 2, it was shown that the T362I substitution on motif D in PV RdRp, derived from the Sabin 1 vaccine strain, lowers both sugar and nucleobase fidelity by shifting the prechemistry conformational equilibrium to the catalytically competent “closed” state, even when noncognate nucleotide binds. This substitution also affects conformations and interactions of amino acid residues that are distant from motif D. Lastly, PVs encoding T362I RdRp are ribavirin sensitive and moderately attenuated in a mouse model of PV.

In Chapter 3, it was shown that besides the T362I substitution, the D53N, Y73H, and K250E substitutions also synergistically contribute to the fidelity of the Sabin 1 RdRp. These effects result in the Sabin 1 RdRp having the lowest sugar fidelity, but nucleobase fidelity similar to WT enzyme. The maintenance of nucleobase fidelity may be a consequence of the much greater selective pressure for maintaining nucleobase recognition in the setting of viral replication. The coordination among the four Sabin substitutions reflects a communication between motif D and other motifs such as motif A, B, and the N terminus.

In Chapter 4, it was shown that modification of two motif F residues (i.e. R174K and K167R substitutions) that interact with the triphosphate moiety of the incoming nucleotide also changes RdRp fidelity. In particular, the R174K variant increases enzyme fidelity tremendously, likely by shifting the prechemistry conformational equilibrium to the “open” state. The K167R variant moderately decreases enzyme fidelity by shifting the prechemistry conformational equilibrium to the “closed” state. Both R174K and K167R variants generate different conformations of the triphosphate moiety of the incoming nucleotide from WT enzyme. The R174K substitution also changes the conformation of motif D although the K167R substitution does not. These observations indicate that the triphosphate rearrangement is a critical fidelity checkpoint and that motifs D and F are not necessarily dependent on each other. This is also the first time that the triphosphate rearrangement in PV RdRp complexes has been inferred in the solution state.
In Chapter 5, it was shown that a novel template inhibitor GPC-N114 inhibits nucleotide incorporation by RdRp of coxsackievirus B3 when the homopolymeric template is used but not when the sym/subU RNA is used. The difference in inhibitory effects of GPC-N114 between the two templates suggests that reaction conditions may need to be improved in the sym/subU RNA system in order to observe the inhibition. GPC-N114 also does not inhibit RNA template switching. Efforts to measure the IC_{50} value were not successful due to the intrinsic fluorescence of GPC-N114. An alternative fluorophore that can be attached to RNA that avoids overlapping the excitation light spectrum of GPC-N114 is required for further studies.

Altogether, these studies indicate that motifs D and F determine enzyme fidelity by repositioning a general acid and rearranging the triphosphate group of the incoming nucleotide, that motif D communicates with other functional structures including motif A and the N terminus to regulate fidelity and that GPC-N114 competes with RNA template and inhibits subsequent nucleotide incorporation by CVB3 RdRp. These studies provide a more complete mechanism of enzyme fidelity and a mechanism of action of a novel template channel inhibitor. These studies of PV RdRp and CVB3 RdRp likely extend to other RdRps due to structural, mechanistic, and functional similarities, which greatly facilitates development of antiviral strategies against these viruses.

6.2 Future directions

6.2.1 Fidelity opens up a new way to design live attenuated vaccine strains

Live attenuated vaccines are excellent vaccines due to their strong immune responses, long-term protection, and inexpensive cost^{12}. The traditional way to develop live attenuated vaccines is very empirical, and the complex mechanisms of viral attenuation make it hard to
directly apply the principles learned in the development of vaccines against other viruses. Fidelity-based viral attenuation opens up a new way to generate live attenuated vaccine strains. Previous studies indicated that PV with a single high fidelity substitution (i.e. K359R) has protective capability as good as the Sabin 1 vaccine that contains 55 substitutions\textsuperscript{13}. The K359R substitution in PV is also highly conserved among RNA viruses, which suggests that the same substitution may lead to attenuation of other viruses. Although this method is very promising, we still need to understand the complete fidelity mechanism of RdRp on the molecular level. In order to generate fidelity-based vaccine strain, we still need to understand the criteria for RdRp function in vaccine candidates, the safety of vaccine candidates, and the relationship between fold changes in fidelity and viral attenuation.

PV RdRp is an ideal model system and results from PV RdRp likely extend to other RNA viruses due to the high sequence and structure conservation of RdRps. For example, Figure 6-1 is the sequence alignment among different members of the genus of enterovirus (including PV, coxsackievirus B3 (CVB3), enterovirus 71 (EV71), and human rhinovirus serotype 14 (HRV14)) in the family of Picornaviridae and another member of the genus of aphthovirus (including foot-and-mouth disease virus (FMDV)) in the same family, which shows high similarities in the sequence of RdRps. A very high identity percentage (i.e. at least 65%) is observed in the comparison of PV RdRp with the other three enteroviruses. PV RdRp also has an identity percentage of 31% versus FMDV RdRp. Crystal structures of RdRps of all the RNA viruses are also highly overlapping with the root-mean-square-deviation (RMSD) less than 1.2 angstroms (Figure 6-2).

An ideal fidelity-based vaccine strain should replicate normally in limited cells but have selective disadvantages in animal models in order to stimulate strong immune responses, RdRps for vaccine candidates should have no or moderate changes in catalytic efficiency but have changes in fidelity compared to WT RdRp. If the generated fidelity variants follow the criteria,
then the corresponding virus mutants can be utilized for feasibility tests of vaccines in cell cultures and animal models. For example, previous studies have shown that perturbations of motif D by a T362I substitution leads to a decrease in fidelity in PV RdRp\textsuperscript{14}. An advantage of this substitution is that it has no effect on the catalysis of cognate nucleotide\textsuperscript{14}. PV carrying this mutation grows as well as WT PV in standard cell culture without the addition of any challenging agents (such as ribavirin) and shows moderate attenuation in the PV mouse model\textsuperscript{14}. Thr362 is not highly conserved, but residues corresponding to Thr362 in other enteroviruses are all polar, for example they are either threonine in HRV14, cysteine in CVB3 and EV71, and lysine in FMDV (Figure 6-1). The similar polarity may represent similar interactions or functions of that particular residue. Additionally, other residues that are significantly perturbed by the T362I substitution are also highly similar (i.e. similar in polarity or size or capability of hydrogen bond) such as Phe363 and Asn370 (Figure 6-1). Substitutions of residues at position 362, 363 and 370 may result in various fidelity variants in other enteroviruses, which facilitates further vaccine tests.
Figure 6.1. Alignment of PV RdRp sequence with the sequences from CVB3, EV71, HRV14, and FMDV. Residues that are highly conserved across all the RdRps are highlighted with red boxes. Residues that are similar are indicated by pink boxes. Highly conserved motifs A to G are colored in cyan, yellow, purple, green, light blue, light purple and dark blue respectively.
Figure 6-2. Structural superposition of PV RdRp with RdRps from CVB3, EV71, HRV14, and FMDV. PV, CVB3, EV71, HRV and FMDV are colored in light blue, green, tan, orange, and light purple respectively (PDB 3OL6, 3DDK, 3N6L, 1XR5, and 1U09).
6.2.2 Engineering motif F to generate RdRp variants suitable for live attenuated vaccine

It was shown that motif F is an important structure in PV RdRp fidelity (see Chapter 4). It is interesting to study whether engineering motif F is able to generate live attenuated vaccines against virulent PV. The first step is to obtain fidelity variants of RdRp that meet the criteria (as described above), which is that RdRps for vaccine candidates should have no or moderate changes in catalytic efficiency but have changes in fidelity. Indirect perturbations of residues Arg174, Lys167, and Arg163 on motif F may lead to such fidelity variants (Figure 6-1). Previous studies have shown that R174K and K167R RdRps have high and low fidelity respectively by disturbing the triphosphate rearrangement of a substrate nucleotide (see Chapter 4). One concern with designing variants based on modifying motif F is that directly perturbing the interactions between the three residues and the triphosphate moiety of the incoming nucleotide may result in significant catalytic loss just like the R174K substitution that causes catalytic reduction by more than ten fold (see Chapter 4). Here, I prefer to indirectly disturb the interactions with the triphosphate group through substitutions of the residues that stabilize Arg174, Lys167, and Arg163. For Arg174, the backbone of Arg174 is stabilized by interactions with the backbone of Lys159, and the side chain is stabilized by a hydrogen bond interaction with Glu161 (Figure 6-3). Lys159 is further stabilized by interactions with the side chains of Glu161 and Ile176 on the other side (Figure 6-3). Therefore, substitutions on Glu161, Lys159, and Ile176 may result in changes in fidelity. For example, substitution of Lys159 to Leu may disturb Glu161 and Arg174. The K159L substitution may lead to an increase in fidelity just like the R174K substitution but may have a higher catalytic rate than R174K RdRp due to indirect perturbation. For Lys167, the backbone of Lys167 has a hydrogen bond interaction with Ser164 (Figure 6-3). Substitution of Ser164 to other nonpolar residues may result in changes in fidelity. For Arg163, this residue has multiple hydrogen bond interactions with γ phosphate moiety of the incoming nucleotide. In order
to position Arg163 toward the triphosphate group, there is an extensive network of interactions that consists of Glu39 and Ala41 from the loop near the N terminus and Val168 and Lys172 on motif F (Figure 6-3). Specifically, the backbone of Arg163 is hydrogen bonded with Glu39. The side chain of Arg163 is stabilized by hydrophobic interactions with Val168 and Ala41 and two hydrogen bond interactions with Lys172. Again, substitutions of corresponding interactions may change fidelity. For example, substitution of Ala41 to Gly may decrease fidelity by relaxing more space for noncognate nucleotide incorporation. These residues interacting with Arg163, Lys167, and Arg174 are highly similar, which suggests that the substitutions proposed here are applicable to other RdRps (Figure 6-1). For example, Lys159 is highly conserved among CVB3, EV71, HRV14 and FMDV (Figure 6-1).
Figure 6-3. The structure of motif F and the incoming nucleotide in PV RdRp. A and B. Part of motif F is shown in grey with Arg174, Lys167 and Arg163 colored in yellow and Lys159, Glu161, Ser164 and Ile176 colored in magenta. Glu39 and Ala41 are located on the loop near the N terminus and are colored in green. The acceptor base to the incoming nucleotide is a square and colored in blue. The incoming nucleotide is colored by heteroatom (carbon (tan), oxygen (red), and phosphor (orange)). Cyan lines represent possible hydrogen bond interactions. The crystal structure is based on the PV RdRp ternary complex (PDB 3OLB).
6.2.3 Elucidating the roles of motifs E and G on RdRp function

My studies have further established the roles that residues in structural motifs D and F play in the nucleotide selection mechanism of the RdRp, but functions of two other highly conserved motifs E (from Phe363 to Ala380) and G (from Ser113 to Tyr120) are still poorly understood. Motifs E and G are likely involved in nucleotide selection. Crystal structures of PV RdRp ternary complexes provide some insights into functions of motifs E and G. For motif E, this motif is located very near to the 3’ end of primer strand of RNA duplex and provides stabilizing interactions with the backbone of the primer strand (Figure 6-4). Specifically, the key residue is Lys375 on motif E. This residue has multiple water-mediated interactions and one hydrophobic interaction with the phosphodiester bond of the -1 and -2 nucleotides respectively (Figure 6-4). Lys375 is positioned near the primer strand by interactions with Glu396 and Pro393 and a hydrogen bond interaction with adjacent Phe373 (Figure 6-4). It is possible that substitutions of Lys375 distort the backbone of the primer strand, which results in abnormal arrangement of the 3’ hydroxyl group of this strand and impaired coordination geometry with the catalytic metal ion, which would result in changes to catalysis and fidelity of the RdRp. For example, substitution of Lys375 to Arg may lead to a moderate reduction in catalytic rate for cognate nucleotide incorporation by misaligning the 3’ end with the triphosphate of incoming nucleotide. However, the catalytic rate may be further reduced for noncognate nucleotide incorporation, which leads to a high fidelity variant.

For motif G, Thr114 is a key residue to stabilize the backbone of the template strand. Specifically, Thr114 has a hydrogen bond interaction with the phosphodiester bond of the +1 and has contacts with the +2 nucleotides (Figure 6-5). This residue is involved in a network of interactions that consists of Asp111, Ser115, and Lys127 (Figure 6-5). Lys127 is also hydrogen bonded to the backbone of the +1 nucleotide (Figure 6-5). Perturbations of the network by
substitutions may result in the distortion of the backbone of the +1 nucleotide and impairment of Watson-Crick base pairing between the acceptor base of the +1 nucleotide and the base of incoming nucleotide. The mis-positioning of the base pair may further transmit to the ribose ring and/or triphosphate group of the incoming nucleotide. Consequently, an inaccurate conformation of the incoming nucleotide may lead to changes in catalysis and fidelity. For example substitution of Thr114 to Ser may lead to changes in the position of the +1 nucleobase, which may facilitate noncognate Watson-Crick base pairing by providing more flexibility to the backbone of the RNA and eventually lead to a low fidelity variant. Understanding the functions that motifs E and G have will complete our understanding of the nucleotide selection mechanism, and provide insight into the coordination among different conserved motifs.

Substitutions on motifs E and G may also lead to similar effects in other enteroviruses. For example, in motif E, Lys375 that has direct contact with the backbone of primer strand is highly conserved among all the aligned RdRps (Figure 6-1). Residues that are used to stabilize Lys375 are either highly conserved or similar among EV71, CVB3, FMDV, and HRV14 (Figure 6-1). Specifically, Phe373 is highly conserved among EV71, CVB3, FMDV, and HRV14 (Figure 6-1). Pro393 is highly conversed among enteroviruses but is different from FMDV that contains an Ala residue at the equivalent position, which suggests that fidelity modulation via substitutions at Pro393 may only be applied in enteroviruses (Figure 6-1).
Figure 6-4. Interactions of motif E with the 3’ end of primer strand (PDB 3OLB). Motif E is colored in light blue with Phe373 and Lys375 colored in dark green. Pro393 and Glu396 are colored in yellow. The α helix that contains Pro393 and Glu396 and motif C are both colored in grey. Asp238 on motif C is colored in magenta. The base of the -2 nucleotide in the primer strand is colored in blue. The backbone of the -1 and -2 nucleotides in the primer strand is colored by heteroatom (carbon (tan), oxygen (red), and phosphor (orange)). The cyan lines represent all the contacts between atoms.
Figure 6-5. Interactions of motif G with the backbone of template strand (PDB 3OLB). Motif G is colored in blue with Asp111, Thr114, and Ser115 that are colored in yellow. Lys127 is colored in magenta. From top to bottom, the backbone of the +1 (blue), +2 (blue), and -1 nucleotides in the template strand is colored by heteroatom (carbon (tan), oxygen (red), and phosphor (orange)). Water molecules are colored in red. The cyan lines represent possible hydrogen bonds.
6.2.4 Other strategies to change the genetic composition of diverse viral populations

Besides point mutations as I studied in this dissertation, RNA viruses also use other methods to mutate or repair their genomes\textsuperscript{15}. One way is through genome recombination, which is widespread in RNA viruses\textsuperscript{16}. Genome recombination occurs when the RdRp temporarily stops during elongation and switches to another RNA template to recover the elongation. Consequently, the newly synthesized genome is a combination of different parental templates, which allows the acquisition of new sequences and maybe new functions for expression. Understanding the mechanism of genome recombination is important for understanding the generation of adaptive mutants in viral populations. One way for genome reparation is the terminal adenyldyl transferase (TATase) activity of the RdRp\textsuperscript{17}. Positive-stand viral genomes contain poly(A) tail at their 3’ ends, which are important for replication, RNA stability, and translation. But the poly(A) tail is also vulnerable to deletion and degradation, and the shortened poly(A) tail cannot perform its normal function. Viral RdRp adds adenyldyl groups on the truncated poly(A) tail and thus recovers its function. Understanding the mechanism of the TATase activity of RdRp also contributes to the composition of viral populations, which is important for viral fitness. Altogether, uncovering the complete set of mechanisms of RNA mutation and reparation by the RdRp has great significance for viral biology and development of antiviral strategies.

6.3 References


VITA
Xinran Liu

Education

The Pennsylvania State University, University Park, PA
Ph.D. Candidate in Chemistry 2010 - 2015

The Sichuan University, Chengdu, China
B.S. in Chemistry 2006 – 2010

Publications


Presentations

