JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

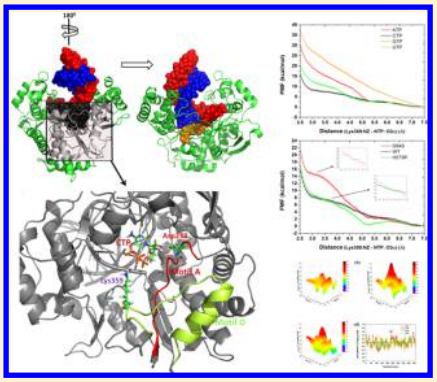
Bridging the Missing Link between Structure and Fidelity of the RNA-Dependent RNA Polymerase from Poliovirus through Free Energy Simulations

Hujun Shen and Guohui Li*

Laboratory of Molecular Modeling and Design, State Key Laboratory of Molecular Reaction Dynamics, Dalian Institute of Chemical Physics, Chinese Academy of Sciences, Dalian 116023, China

Supporting Information

ABSTRACT: RNA-dependent RNA polymerases (RdRps) are enzymes catalyzing RNA replication from a RNA template. Active-site closure in RdRps, normally induced by correct nucleotide triphosphate (NTP) binding, is a prerequisite for the cycle of nucleotide incorporation. So, a complete understanding of polymerase function (in particular polymerase fidelity) of a RdRp requires more complete knowledge of active-site closure in the RdRp. In this work, based on solved crystal structures, we have built different models for the RNA-dependent RNA polymerase from poliovirus (termed PV 3D^{pol}). Through MD simulations and free energy calculations of these PV 3D^{pol} models, we have revealed the dynamic correlation between motif A and motif D and between motif A and incoming NTP, have deepened our understanding of polymerase fidelity from dynamic aspects, and have provided an explanation to the puzzle that arises from different observations based on kinetic studies and structural data.



INTRODUCTION

Poliovirus (PV) has been extensively employed as an important model system for in-depth understandings of polymerase functions, enzyme structures, life cycles, viral pathogenesis, and antiviral drug-resistance of RNA viruses.[1−3] All of these RNA viruses, including hepatitis A virus (HAV), hepatitis C virus (HCV), foot and mouth disease virus (FMDV), and rhinovirus (a common cold virus), impose a great threat to both human and animal health. The RNA dependent RNA polymerase (RdRp) from poliovirus, termed PV 3D^{pol}, functions as an enzyme to replicate the RNA genome of PV during viral life cycles. Crystal structures of PV 3D^{pol} have been recently reported,[4,5] showing that it is structurally similar to other RdRps (see details in Supporting Information). First, similar to other RdRps, the three-dimensional structure of PV 3D^{pol} resembles a cupped right-hand and consists of three structural subdomains: "fingers", "palm", and "thumb", shown in Figure S1 of Supporting Information. Second, although in different RdRps the fingers and palm subdomains can fold in a different manner, the palm subdomain of these RdRps has a similar fold; in particular, the palm subdomain contains highly conserved four structural motifs (A, B, C, and D) forming a center for catalysis (see Figure S2 of Supporting Information).

Kinetic studies[6−11] of the nucleotide incorporation catalyzed by RNA or DNA polymerases suggest that the kinetically observable steps in each nucleotide incorporation cycle should include (1) an incoming nucleotide triphosphate (NTP) binds to a polymerase in complex with nucleic acid primer/template, forming a ternary complex; (2) the ternary complex undergoes a so-called "pre-chemistry conformational change", resulting in the closure of active site; (3) following the active-site closure, a catalytic reaction occurs (termed chemistry step); (4) a postchemistry conformational change occurs, leading to the release of pyrophosphate (PPi) leaving group; (6) the RNA or DNA template/primer (TP) duplex is translocated for the next nucleotide incorporation.

Kinetic and nuclear magnetic resonance (NMR) studies of RNA or DNA polymerases[6,8,10,12−17] suggest that the prechemistry conformational change step of nucleotide incorporation is usually an essential checkpoint for polymerase fidelity since active-site closure is normally a slow and rate-limiting process. Moreover, it has been shown that the active-site closure in PV 3D^{pol} is sensitive to the type of incoming NTP.[17] Unfortunately, the dynamic picture of the active-site closure in PV 3D^{pol} remains obscure.

Experimental studies of PV 3D^{pol} suggest that Lys359 of motif D should serve as a general acid to donate a proton to PPi leaving group.[1,18,19] However, up to now, there is no structure available for PV 3D^{pol} indicating that Lys359 of motif D is poised for catalysis.[5,20] Therefore, in order to clearly comprehend the role of Lys359 in catalysis, it should be necessary to construct a free energy profile to describe the mechanism of active-site closure that is characterized by the motion of Lys359.

Polymerase fidelity is a measure of the ability of a polymerase to discriminate against incorrect nucleotide incorporation. High fidelity indicates a low frequency of errors in RNA replication

Received: July 21, 2014
Published: October 15, 2014

   dx.doi.org/10.1021/ct5006449 | J. Chem. Theory Comput. 2014, 10, 5195−5205

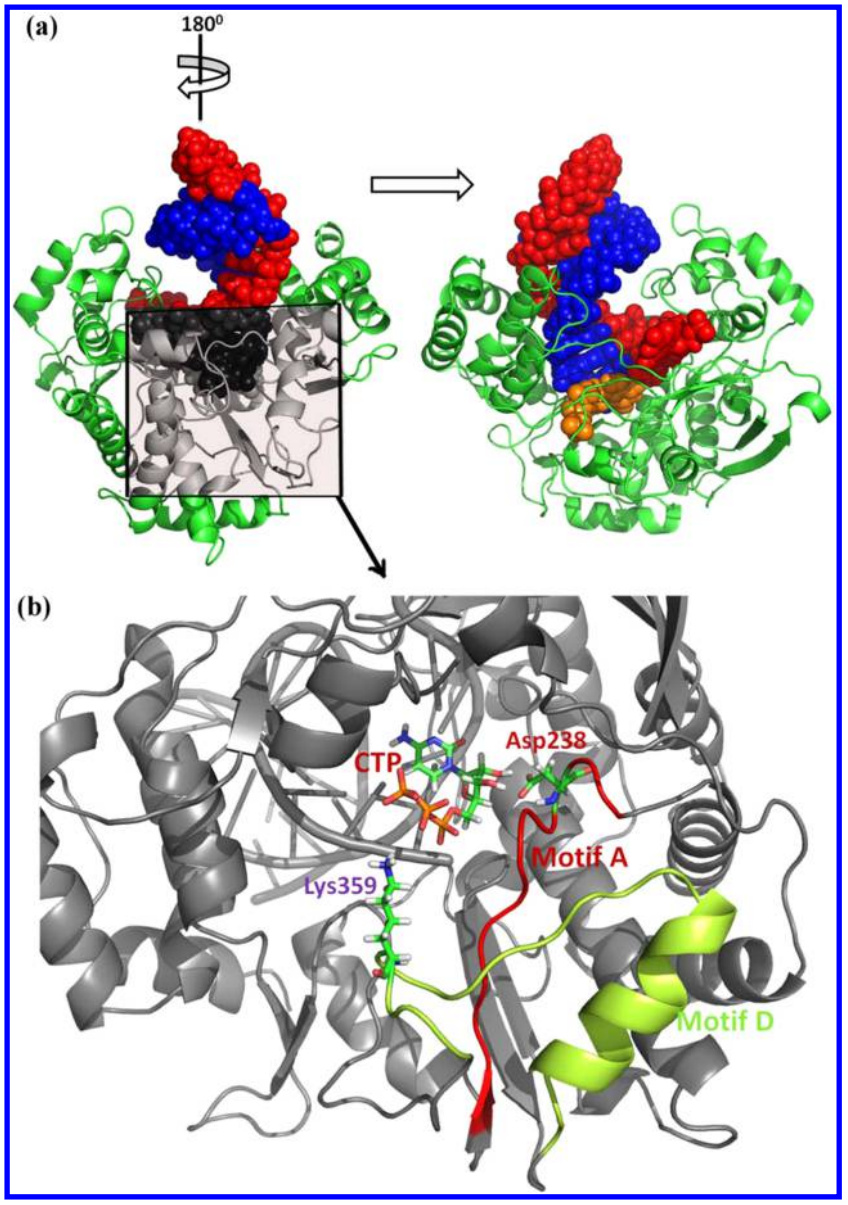


Figure 1. Representation of the ternary complex structure of PV 3D^{pol}. (a) The PV 3D^{pol} ternary complex is composed of polymerase (green cartoon), RNA template (red VDW sphere), RNA primer (blue VDW sphere) and incoming NTP (orange VDW sphere). (b) Representation of important structural elements of PV 3D^{pol}, motif A is represented by red cartoon, motif D by lime cartoon, and Asp238, Lys359, and CTP by colored sticks, respectively.

while low fidelity suggests a high frequency of errors. It has been recognized that viral polymerase fidelity is crucial to the viral population fitness because an increase or a decrease in the viral polymerase fidelity could directly affect virus adaptation and evolution.[1,3] It has been shown that a small change in polymerase fidelity might result in the reduction of viral population fitness.[21] Fidelity differs substantially among different polymerases. For instance, in the case of PV 3D^{pol}, a single mutation can make a significant change in polymerase fidelity: the mutant G64S induces higher polymerase fidelity[22−24] while the mutant H273R leads to lower polymerase fidelity[25,26] as compared to wild type (WT). However, the major obstacle to a complete understanding of polymerase fidelity is the limited information about the dynamic features of the prechemistry conformational change.[27] Although previous MD studies of PV 3D^{pol} have provided valuable information about polymerase dynamics,[28−30] a traditional MD method has its limitation to adequately probe the prechemistry conformational change that usually occurs on microsecond time scales and beyond.

This study continues our previous work[30] by carrying out extensive MD simulations and free energy calculations on different models of PV 3D^{pol}. In this work, different PV 3D^{pol} ternary complexes have been built, including four PV 3D^{pol} ternary complexes in each of which contains wild type PV 3D^{pol}, RNA primer/template, and a correct or incorrect nucleotide triphosphate (such as ATP, CTP, GTP, or UTP), and the ternary complexes of two PV 3D^{pol} mutants (G64S and H273R) in each of which contains a PV 3D^{pol} mutant (G64S or H273R), RNA primer/template, and a correct nucleotide triphosphate (CTP). Based on the computational results obtained from MD simulations and free energy calculations of these ternary complexes (PV 3D^{pol}·RNA·NTP), we gained in-depth understanding of the correlation between polymerase

dx.doi.org/10.1021/ct5006449 | J. Chem. Theory Comput. 2014, 10, 5195−5205

function and dynamics. We discovered that the dynamics of motif D, characterized by the flexibility of motif D and correlated motions of motifs A and D, was strikingly sensitive to the base paring strength between incoming NTP and the nucleotide of RNA template. The knowledge of these dynamic features of functional structural elements in PV 3D^{pol} allows us to present reasonable explanations of why the active-site closure in PV 3D^{pol} would be normally triggered by correct NTP binding, why the catalytic role of Lys359 of motif D cannot be detected by X-ray experiment, and why a single mutation in PV 3D^{pol} can lead to an observable change in polymerase fidelity.

METHODS

Building Models. Based on the crystal structures of PV 3D^{pol} elongation complexes solved by X-ray experiment (PDB codes: 3OL7 and 3OLB),[20] different ternary complex models of PV 3D^{pol} were generated by removing some components from the crystal structures of PV 3D^{pol} except for chain A (corresponds to PV polymerase), chain B (corresponds to RNA template), chain C (corresponds to RNA primer), two coordinated Mg^{2+} ions and crystal water molecules. For each PV 3D^{pol} ternary complex (PV·RNA·NTP), the RNA primer/template was further modified by removing a few unpaired nucleotides as we did in previous work.[30] A representative of these models is presented in Figure 1. Finally, a NTP molecule (ATP, CTP, GTP or UTP) was inserted into the active site of PV 3D^{pol} respectively, forming a base pair with the nucleotide (G) of RNA template, see Figure 2. Thus, ATP, GTP, and UTP are incorrect NTP substrates while CTP represents correct NTP substrate. In the same way, the ternary complexes of two mutant polymerases (G64S and H273R) in complex with RNA primer/template and correct CTP substrate were also built for computational study.

Our PV 3D^{pol} models should be reliable for computational studies due to the following facts: (1) different crystal structures of PV 3D^{pol} bound with RNA and different substrates (such as cognate CTP and its 2′-dCTP, 3′-dCTP and 2′, 3′-dCTP)[20] have been superimposed, showing that there is no significant change in the overall topology of PV 3D^{pol} (RMSD values between these structures are less than 1.0 Å); (2) the crystal structures of PV 3D^{pol} bound with different NTPs[5] suggest that different NTP binding introduces slightly structural difference (RMSD values between these structures are less than 1.0 Å); (3) a single mutation in PV 3D^{pol} introduces a trivial structural change if one compares the crystal structures[22] of the mutant G64S and wild type PV 3D^{pol} (RMSD value between the two structures is less than 0.5 Å).

Molecular Dynamics Simulations. Atomistic MD simulations of all PV 3D^{pol} models were carried out in the AMBER program[31] using AMBER99SB force field[32] with the parmbsc0 force field correction for RNA.[33] Each system was neutralized with a number of sodium ions and then was immersed in a truncated octahedron solvent box filled with TIP3P water molecules,[34] and a distance of at least 15 Å was required between the surface of each PV 3D^{pol} model and the edge of water box. Energy minimization was performed by imposing a strong restraint on each PV 3D^{pol} system and was followed by minimizing whole system for a few thousand steps. NVT simulations were carried out by heating whole system slowly from 100 to 300 K, and the Berendsen thermostat[35] was used to govern the temperature of the whole system. Subsequently, a 100 ps NVT dynamics was performed and was followed by a NPT production run. During the NPT production run, all

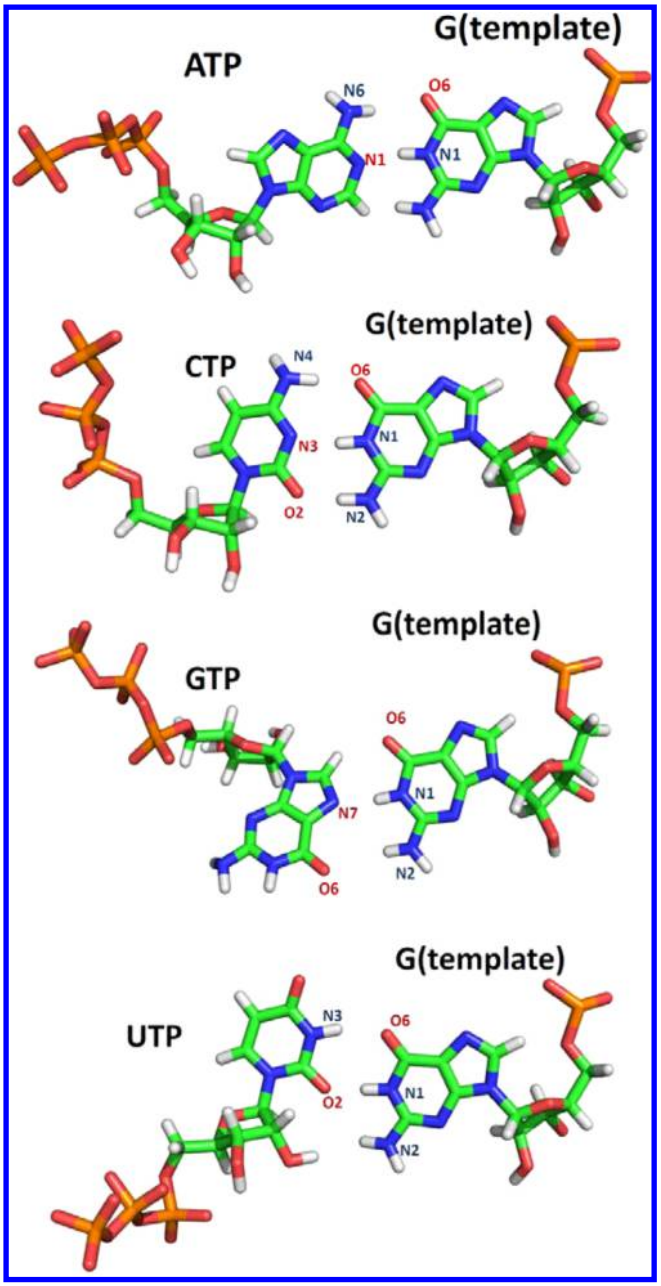


Figure 2. Base pairs in the ternary complexes of wild type PV 3D^{pol}. In each model of the PV 3D^{pol} ternary complexes, an incoming nucleotide triphosphate (NTP), such as ATP, CTP, GTP, or UTP, forms a base pair with an unpaired base (guanine in this case) of RNA template. Thus, CTP is a correct nucleotide triphosphate while other NTPs (ATP, GTP, and UTP) represent incorrect ones.

bonds associated with hydrogen atoms were constrained by employing the SHAKE algorithm[36] such that the integration time step of 2 fs could be used. A cutoff value of 10 Å was set for nonbonded interactions and Particle Mesh Ewald method[37] was employed for treating electrostatic interactions. For each PV 3D^{pol} model, five independent MD simulations were carried out using different velocities that were randomly generated at the beginning of the simulations.

Free Energy Calculations. To construct a free energy profile for the prechemistry conformational change, we have carried out potential of mean force (PMF) calculations on the movement of motif D toward NTP. Our PMF calculation

dx.doi.org/10.1021/ct5006449 | J. Chem. Theory Comput. 2014, 10, 5195−5205

method is similar to a traditional umbrella sampling technique. For each PV 3D^{pol} model, the distance between NZ atom of Lys359 and O3A atom of NTP was considered as the reaction coordinate for all PMF calculations. In each PMF calculation, 50 configurations along the reaction coordinate were generated by pulling NZ atom of Lys359 in the direction of approaching O3α atom of NTP from 7.5 to 2.5 Å at an interval (or window) of 0.1 Å. These configurations served as the starting configurations for subsequent MD simulations using similar simulation protocol described above except a force constant of 100 kcal/mol/Å^2 was applied to restrain the distance between NZ atom of Lys359 and O3A atom of NTP. The length of the simulation (the convergence test has been made, please see Figure S3, Supporting Information), the value of the force constant (100 kcal/mol·Å^2) and 50 windows had been chosen so that the reaction coordinate in adjacent windows could have an adequate overlapping (see Figures S4, Supporting Information). Finally, a weighted histogram analysis method (WHAM)[39,40] was used to generate all free energy profiles.

RESULTS/DISCUSSION

Dynamic Personality of Wild Type PV 3D^{pol} Bound by Different NTP Substrates. For each ternary complex of wild type PV 3D^{pol}, the root-mean-square deviation (RMSD) (crystal structure was considered as the reference) and radii of gyration (RG) values have been calculated for backbone heavy atoms of PV 3D^{pol} from five independent atomistic MD simulations, plotted in Figure S5−S6, Supporting Information. The RMSD curves show the low flexibility of PV 3D^{pol} in different ternary complexes. Meanwhile, the small-amplitude fluctuation in RG curves also demonstrates the high stability of PV 3D^{pol} in different ternary complexes. In other words, it appears that the global dynamics of PV 3D^{pol} in its ternary complex should be insensitive to incorrect NTP binding. As a matter of fact, it has already been discussed in our previous study[30] that the stable ternary complex of PV 3D^{pol} is mainly attributed to the interactions between the polymerase (PV 3D^{pol}) and its RNA template/primer and between the finger tips and the thumb domain.

The root-mean-square fluctuation (RMSF) curves of α carbon atoms in four different ternary complexes of PV 3D^{pol} are plotted in Figure 3a, suggesting that PV 3D^{pol} tends to preserve the similar dynamic pattern among different complexes. In other words, incorrect NTP binding would have a limited impact on the dynamic personality of PV 3D^{pol} as a whole due to its high stability in its ternary complex. Nevertheless, the structural motif D (in particular, the residues around Lys359) was observed to exhibit different behaviors in different PV 3D^{pol} complexes. By subtracting the RMSF values of PV 3D^{pol} in complex with an incorrect NTP substrate from those of PV 3D^{pol} with correct CTP substrate (see Figure 3b), it is clear to see that a mismatching base pair (ATP:G, GTP:G, or UTP:G) would result in a significant decrease in the flexibility of Lys359 as well as its neighboring residues.

According to our previous studies,[28−30] it is known that positively correlated motions exist between motif A and motif D. In this work, it has been found that the positively correlated motions were still preserved between motif A and motif D among PV 3D^{pol} different complexes (see Figure S7, Supporting Information). In fact, the correlated dynamic nature of motifs A and D is mainly attributed to the hydrogen bonding interactions between Thr353 of motif D and Thr235 of motif A and between Thr355 of motif D and Asp233 of motif A (see

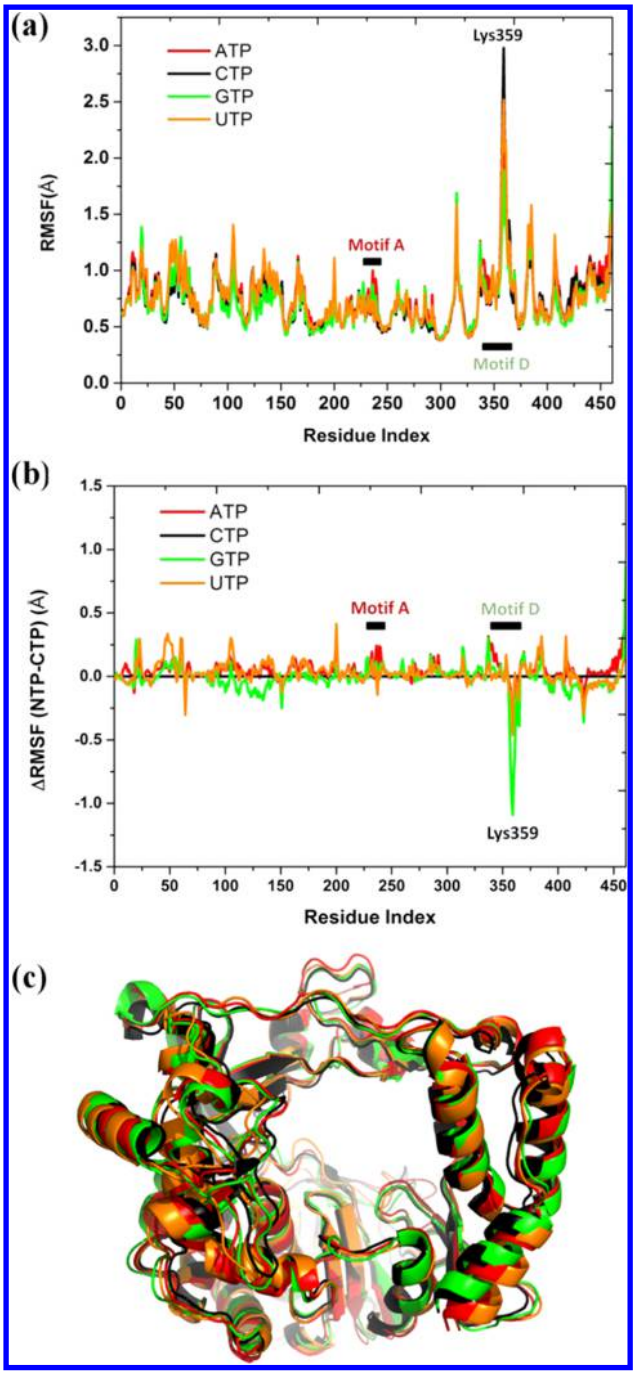


Figure 3. RMSF results for wild type PV 3D^{pol} in different complexes. (a) RMSF curves for α carbon atoms of PV 3D^{pol} in different complexes with ATP (red), CTP (black), GTP (green), and UTP (orange), respectively. For each case, the RMSF values of α carbon atoms were obtained by averaging over five independent MD trajectories. (b) Difference between RMSF results for different PV 3D^{pol} complexes. The RMSF difference (ΔRMSF) is obtained by subtracting the RMSF values for PV 3D^{pol} bound by an incorrect NTP (such as ATP, GTP, or UTP) from those for PV 3D^{pol} with correct CTP binding respectively, such as, ΔRMSF = RMSF(NTP) − RMSF(CTP), where NTP stands for ATP, GTP, or UTP, respectively. (c) Cartoon representation of the average structures of PV 3D^{pol} in different complexes with ATP (red), CTP (black), GTP (green), and UTP (orange) obtained from MD simulations, respectively, and the RMSD values (for backbone atoms) between their structures are less than 1.5 Å.

dx.doi.org/10.1021/ct5006449 | J. Chem. Theory Comput. 2014, 10, 5195−5205

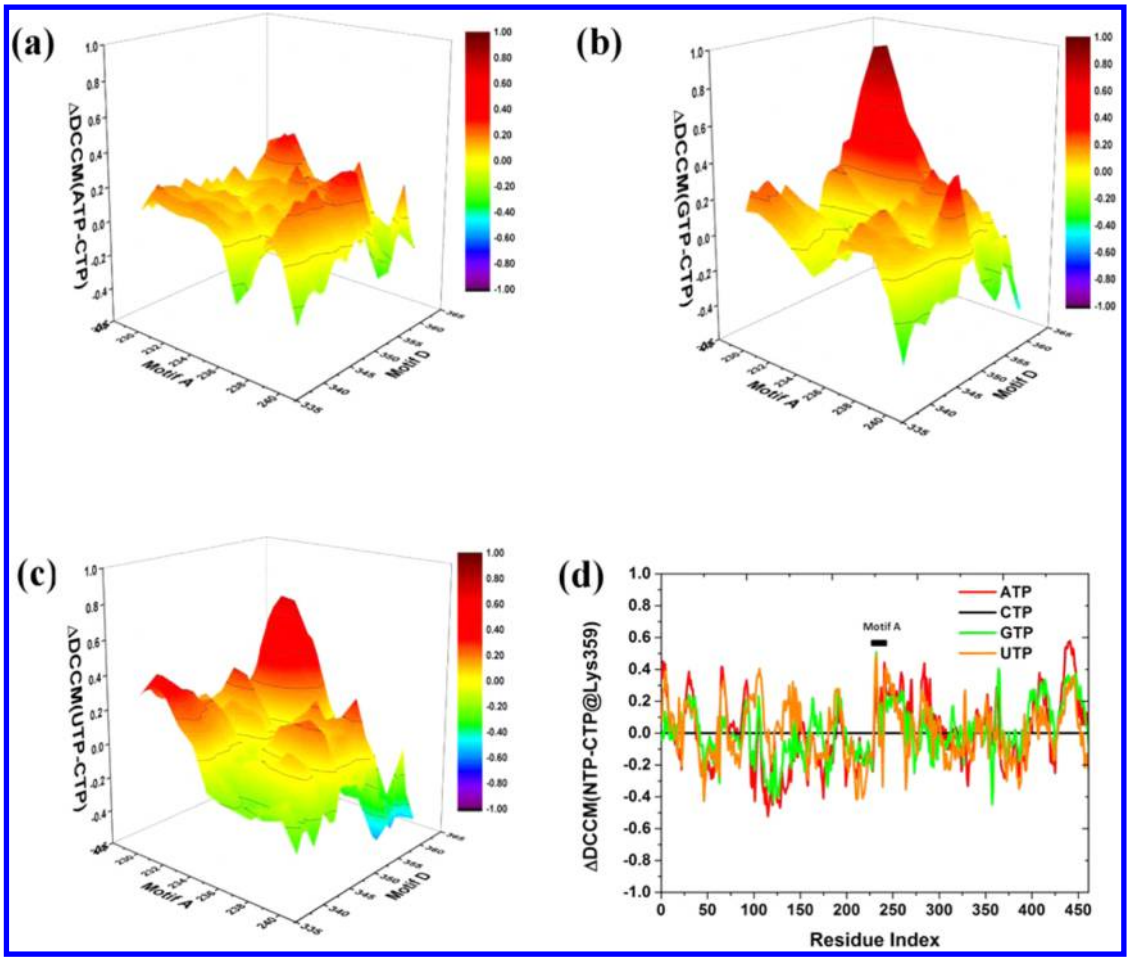


Figure 4. Dynamic cross correlation map (DCCM) results for wild type PV 3D^{pol} in different complexes. In each case, DCCM values of the residues pairs in PV 3D^{pol} were obtained by averaging over five independent MD trajectories. (a) 3D plot of the difference between DCCM results for the PV 3D^{pol} complexes with ATP binding and with CTP binding, such as, ΔDCCM = DCCM(ATP) − DCCM(CTP). (b) 3D plot of the difference between DCCM values for the PV 3D^{pol} complexes with GTP binding and with CTP binding, such as, ΔDCCM = DCCM(GTP) − DCCM(CTP) (c) 3D plot of the difference between DCCM values for the PV 3D^{pol} complexes with UTP binding and with CTP binding, such as, ΔDCCM = DCCM(UTP) − DCCM(CTP). In a−c, DCCM results only show the correlated motions of motifs A and D. (d) Plot of the difference between DCCM results for the complexes with incorrect NTP binding and with CTP binding, and the DCCM results only demonstrate the correlated motions between all residues of polymerase and Lys359. In d, the DCCM difference for ATP, GTP, or UTP binding is plotted by red, green, and orange lines, respectively.

Figure S8, Supporting Information). From Figure 3b, one can see that a mismatching base pair has a limited effect on the flexibility of motif A but it has an observable impact on the flexibility of motif D (Figure 3b). So, it is reasonable to imagine that the positively correlated motions of motif A and motif D would be affected in different PV 3D^{pol} complexes to a certain degree. The comparison of the DCCM results of PV 3D^{pol} with incorrect and correct NTP binding has been made in this work (see Figure 4a-c), suggesting that, in general, the positively correlated motions of motif A and motif D would be strengthened by incorrect NTP binding. In other words, a mismatching base pair in PV 3D^{pol} ternary complex would lead the motions of motif A and motif D to become more correlated in the same direction (see Figure 4d).

In summary, our understanding of the incorrect NTP binding in PV 3D^{pol} ternary complex is that it has (1) a limited impact on the global dynamics of PV 3D^{pol}, (2) a limited influence on the flexibility of motif A, (3) a significant influence on the flexibility of motif D, and (4) a noteworthy effect on the correlated motions of motif A and motif D.

Why Is Active-Site Closure Normally Triggered by Correct NTP Binding? Based on the crystal structures[20] and NMR study[17] of PV 3D^{pol}, it is believed that active-site closure should be normally triggered by correct NTP binding since the nucleotide selection in PV 3D^{pol} should be mainly controlled by base-pairing interactions. In addition, the NMR study[17] of PV 3D^{pol} has suggested that the movement of motif D is critical to active-site closure and is sensitive to incorrect NTP binding.

In this work, the base pairing strength in different PV 3D^{pol} complexes was measured between NTP and Guanine of RNA template (see Figure 2) from MD trajectories, and the results are presented in Table 1. It is not surprising that the matching base pair (CTP:G) has stronger base pairing strength than a mismatching base pair (such as ATP:G, GTP:G, or UTP:G). Moreover, it is known from the table that the base pairing strength for the mismatching base pairs follows this order: GTP:G > ATP:G > UTP:G. There is no doubt that the base pairing strength would affect the stability of a NTP in PV 3D^{pol} complex, and the RMSD values of NTPs measured from MD simulations have confirmed the correlation between base pairing strength and stability (see Figure 5b). Therefore, it is

dx.doi.org/10.1021/ct5006449 | J. Chem. Theory Comput. 2014, 10, 5195−5205

Table 1. Hydrogen Bonding Strength between a NTP and Its Base Pairing Nucleotide (Guanine) of the RNA Template in Wild Type PV 3D^{pol} Complexes[a]

base pair	donor−acceptor	hydrogen bonding occupancy
ATP:G	G(N1)−ATP(N1)	3.40%
	ATP(N6)−G(N6)	1.60%
CTP:G	G(N2)−CTP(O2)	84.15%
	G(N1)−CTP(N3)	72.55%
	CTP(N4)−G(O6)	65.50%
GTP:G	G(N2)−GTP(O6)	47.08%
	G(N1)−GTP(N7)	37.75%
UTP:G	G(O6)−UTP(N3)	0.00%
	UTP(O2)−G(N1)	0.00%

[a]NTP in the wild type PV 3D^{pol} complexes represents ATP, CTP, GTP, or UTP respectively. The names in parentheses are the names of heavy atoms serving as donors or acceptors in the hydrogen bonding interactions. For each case, hydrogen bonding occupancy is calculated by averaging over five independent MD trajectories. All the details about the calculations of the hydrogen bond occupancy are given in Supporting Information.

expectable that, in a PV 3D^{pol} ternary complex, it would become more difficulty for motif A to organize an incorrect NTP for catalysis as compared to a correct NTP.

The RMSF and DCCM results suggest that a mismatching base pair in a PV 3D^{pol} ternary complex would lead to a decrease in the flexibility of motif D and an increase in the strength of the positively correlated motions of motif A and motif D. In brief, the dynamics of motif D is sensitive to the base pairing strength of NTP:G. This can be displayed clearly by the free energy profiles for moving motif D toward active site in different PV 3D^{pol} ternary complexes, presented in Figure 5c. From these free energy profiles, one can see that the possibility of motif D to access the active site is the lowest for the weakest base pair (UTP:G) but is the highest for the strongest base pair (CTP:G). Thus, our understanding of active-site closure is that, due to the weaker hydrogen bonding interactions with the RNA template, an incorrect NTP binding in a PV 3D^{pol} ternary should result in (1) a decrease in the accessibility of motif D to active site and (2) a decrease in the capability of motif A to orientate NTP in a proper position for catalysis as compared to a correct NTP binding, implying that the active-site closure in PV 3D^{pol} should normally triggered by correct NTP binding.

Why Is It Difficult for X-ray Experiment to Detect the Role of Lys359 in Proton Transfer Process? It has been suggested from kinetic and NMR studies[1,17−19] of PV 3D^{pol} that a proton transfers from Lys359 of motif D to pyrophosphate (PPi) leaving group, triggering the translocation of RNA duplex, and the protonated Lys359 of PV 3D^{pol} might play an analogous to His1085 of RNA polymerase II[41,42] in serving as a general acid for the proton transfer process. However, up to now, there are no structural data available for PV 3D^{pol} to demonstrate that Lys359 is positioned in active site for catalysis.[5,20] It has been mentioned previously that to complete our understanding of the role of motif D it is necessary to obtain full knowledge of the dynamics of motif D.[27] Our previous computational study[30] of PV 3D^{pol} suggest that although Arg174 of motif F appears to be positioned more closely to active site as compared to Lys359 of motif D, it is more difficult for Arg174 to donate a proton to a NTP substrate than for Lys359 to do. This study continues our previous work[30] in order to unearth the roles of Lys359 more completely.

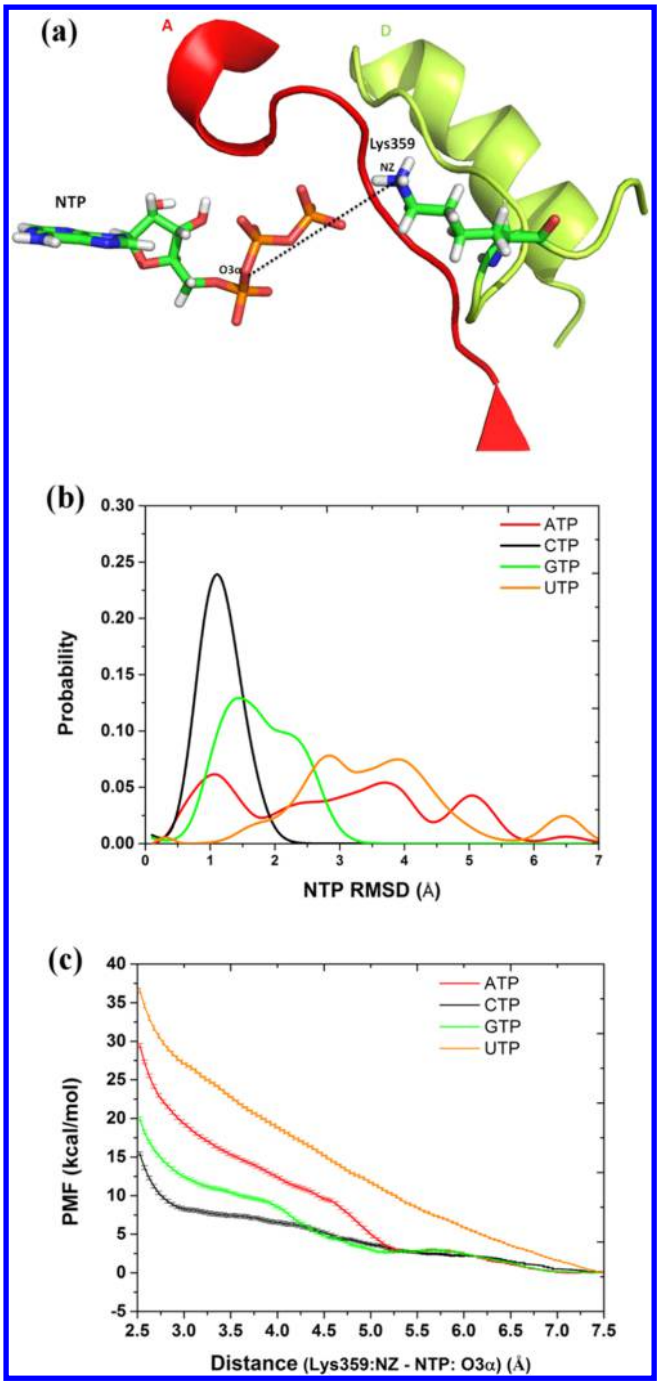


Figure 5. PMF results for moving Lys359 toward the active site in the wild type PV 3D^{pol} complexes with different NTP substrates. (a) Representation of the structural elements of PV 3D^{pol}, motif A is represented by red cartoon, motif D by lime cartoon, Lys359 and NTP by colored sticks, respectively. The distance between NZ atom of Lys359 and O3α atom of NTP is illustrated by a dashing line. (b) Distribution of RMSD values for ATP (red solid line), CTP (black solid line), GTP (green solid line), and UTP (orange solid line) in wild type PV 3D^{pol} complexes. A RMSD value indicates how far a NTP is positioned from the active site. (c) PMF results for moving Lys359 toward ATP (red solid line), CTP (black solid line), GTP (green solid line), and UTP (orange solid line) respectively, and error bars (around 0.2 kcal/mol) are plotted in a horizontal manner.

dx.doi.org/10.1021/ct5006449 | J. Chem. Theory Comput. 2014, 10, 5195−5205

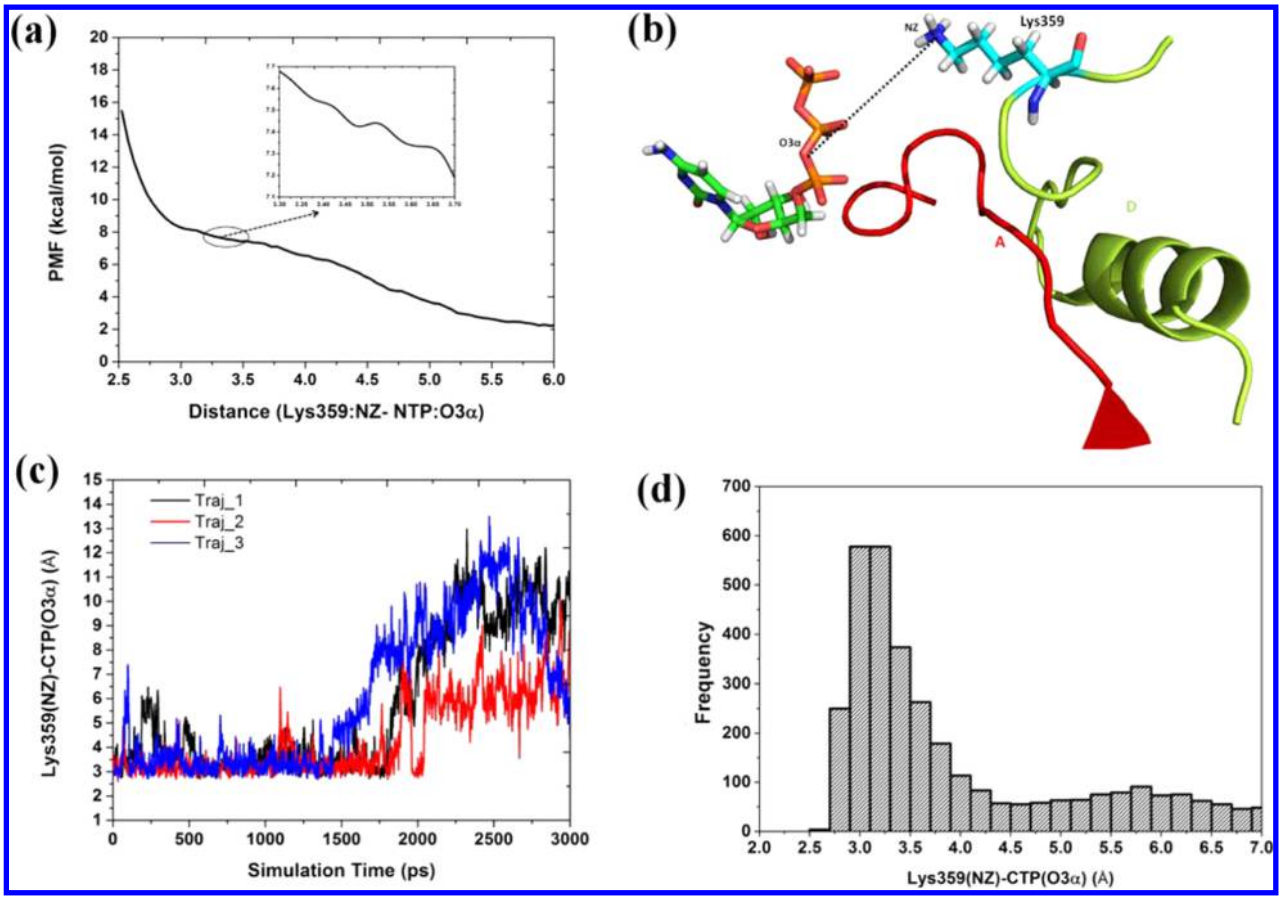


Figure 6. Dynamics of Lys359 in wild type PV 3D^{pol} in complex with CTP binding. (a) PMF result for moving Lys359 toward CTP, and an inset shows the flat region of the PMF curve, illustrating that local energy minima can be found in this region. (b) Representation of the structural elements of PV 3D^{pol}, motif A is represented by red cartoon, motif D by lime cartoon, Lys359 and CTP by colored sticks, respectively. The distance between NZ atom of Lys359 and O3α atom of CTP is indicated by a dashing line. (c) Distances between NZ atom of Lys359 and O3α atom of CTP, obtained from three free MD simulations (without restraints) using different starting structures sampled from free energy simulations. (d) Distribution of the distances between NZ atom of Lys359 and O3α of CTP obtained from three free MD simulations.

From the free energy profile for moving Lys359 toward CTP (see Figure 6a, b), it appears that the free landscape becomes relatively flat as Lys359 approaches to the active site (in the range of distances from 3.0 to 3.5 Å). In addition, some shallow local energy minima exist in this region when one looks at the region more closely (see the inset figure in Figure 6a). This feature would suggest that it might be possible for Lys359 to be positioned in the catalytic center. In order to find out the degree of possibility of Lys359 positioned in the active site for catalysis, we carried out additional MD simulations of wild type PV 3D^{pol} ternary complex with CTP binding. From above MD trajectories used for free energy calculations, we randomly selected a few configurations in which NZ atom of Lys359 and O3α atom of CTP are separated at a distance of around 3.0 Å. Using these structures as a starting conformation, we carried out a few free MD simulations independently (no restraint was imposed on the distance between Lys359 and CTP) by following the MD simulation protocol described in Methods. From these free MD simulation trajectories, the distances between NZ atom of Lys359 and O3α atom of CTP were monitored along the simulation time, suggesting that Lys359 would be able to stay in the active site on a nanosecond time scale (see Figure 6c, d). Shortly, due to the dynamic nature of the movement of Lys359 toward the active site (for instance, the flat energy barrier with some shallow local minima in the range of separations from 3.0 to3.5 Å), it is possible for Lys359 to stay in the active site for a while (nanosecond time scale) to donate a proton to incoming NTP since the time scale of a catalytic reaction usually occurs on the order of picoseconds (ps). From this perspective, our computational results are supportive of the observations made by the kinetic and NMR studies[1,17−19] of PV 3D^{pol}. On the other hand, our free energy calculations have shown that the time scale of Lys359 being positioned in the active site should be the order of nanoseconds, and this time scale might impose a great difficulty on traditional X-ray experiments, explaining why there are no crystal structures currently available for Lys359 being poised for catalysis.

How Does a Single Mutation in PV 3D^{pol} Make an Observable Change in Polymerase Fidelity? RNA polymerase fidelity describes error frequency in viral RNA replication. Compared to replicative DNA-dependent DNA polymerases (DdDps), RNA-dependent RNA polymerases (RdRps) misincorporate at much higher frequencies. High mutation rate in a viral RNA replication would lead a RNA virus population to become an ensemble of RNA viruses, which is called quasispecies.[1] Mutation is an adaptive ability of RNA viruses to survive in the environment. However, it has been found that a slight decrease or a small increase in mutation rate might lead to the reduction of viral population fitness,[21] such that RNA polymerase fidelity is critical to the viral population fitness of RNA viruses. In the case of PV 3D^{pol}, a single

dx.doi.org/10.1021/ct5006449 | J. Chem. Theory Comput. 2014, 10, 5195−5205

mutation can cause an observable change in polymerase fidelity. For instance, it is known that the mutation (G64S) at site 64 could cause polymerase fidelity to increase[13,14] whereas the mutation (H273R) at site 273 would result in a decrease in polymerase fidelity.

To gain more understanding of polymerase fidelity from the aspect of polymerase dynamics, we extended our MD simulations and free energy calculations to explore dynamic properties of two PV mutants (such as G64S and H273R). It has been suggested previously that the movement of motif D should be critical to active-site closure. Thus, the free energy profiles for moving Lys359 toward the active site have been constructed for G64S and H273R polymerases in this work, shown in Figure 7a. From the figure, it is clear to see that it is more difficult for the high fidelity polymerase (G64S) to arrange Lys359 to the active site than for the low fidelity polymerases (H273R and wild type), suggesting that polymerase fidelity should be related to the conformational change of motif D. It has already been discussed that the motion of Lys359 should be correlated with the interactions between motif A and motif D: a positive increase in the correlated motion of motif A and motif D would increase the energy barrier for Lys359 moving toward active site (see Figures 4 and 5b). Simply speaking, the interactions between motif A and motif D would impose a restriction on the freedom of Lys359. In the same way, when the movement of Lys359 becomes more difficulty for the high fidelity G64S polymerase than for the low fidelity H273R polymerase (see Figure 7a), this would indicate that motif A and motif D would become more positively correlated for the high fidelity G64S polymerase than for the low fidelity H273R polymerase. This is confirmed by our results presented in Figure 8. In fact, the mechanistic study of PV G64S mutant polymerase[24] has revealed that the mutant G64S has reduced the equilibrium constant of the prechemistry conformational change step in nucleotide cooperation. Therefore, both computational and mechanistic studies reached the same conclusion: nucleotide incorporation cycle would be more faithful when it is more difficult for Lys359 to achieve the catalytically active conformation.[27]

It has been discussed above that the accessibility of Lys359 to active site is sensitive to the hydrogen bonding strength of NTP:G in PV $3D^{pol}$ complex, so here we have compared the hydrogen bonding strength of the base pair (CTP:G) in the complexes of two mutant polymerases (G64S and H273R) and wild type polymerase. From the results shown in Table 2, one can see that the hydrogen bonding strength is the strongest in H273R and the weakest in G64S among the three polymerases. Similar to the free energy results for the wild type polymerase bound by different NTP substrates, the accessibility of Lys359 to the active site is also correlated with the base paring strength of the base pair (CTP:G) (see Figure 7a and Table 2) in the three cases (wild type, G64S, and H273R), indicating the correlation between the base paring interactions (between NTP and RNA template) and polymerase function. Thus, to understand how a single mutation makes polymerase fidelity change, it is necessary to figure out why or how a single mutation affects the hydrogen bonding interactions between the base pair (CTP:G).

It is known that the hydrogen bonding interactions exist between Gly1 and Gly64 in wild type PV $3D^{pol}$, so it is possible for the mutation (G64S) to affect the hydrogen bonding interactions between the residue pair (at sites 1 and 64). From the results shown in Table 3, one can see that the hydrogen

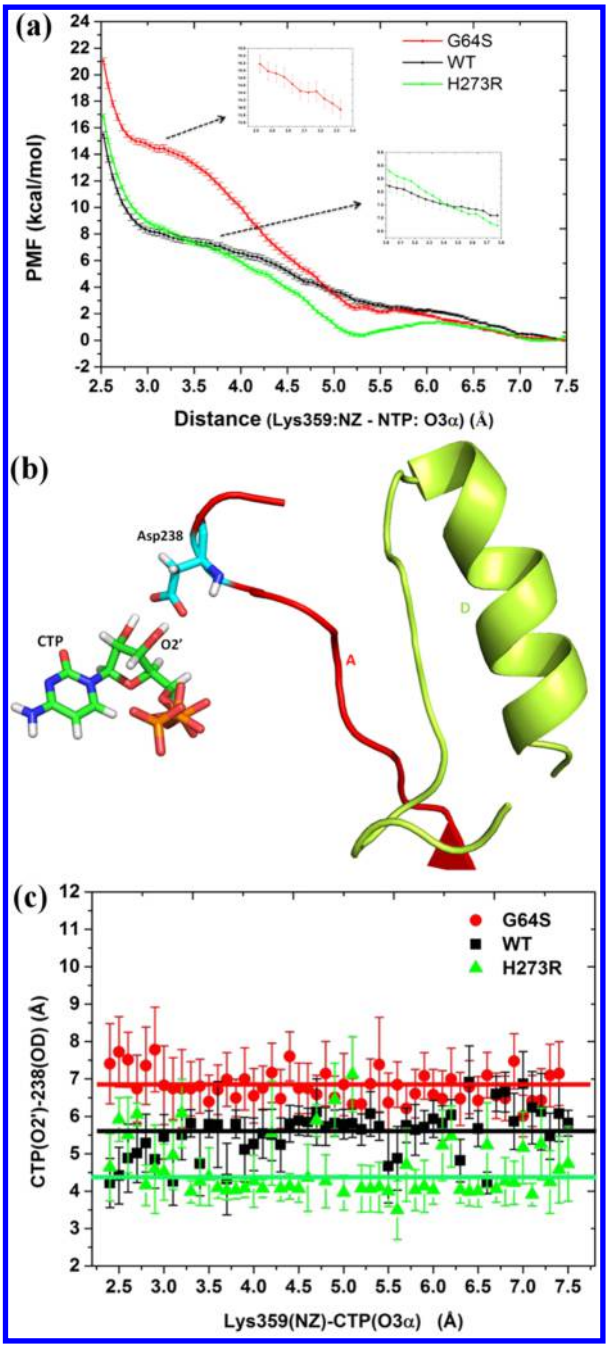


Figure 7. PMF results for moving Lys359 toward the active site in the complexes of different polymerases (wild type, G64S, and H273R) with CTP binding. (a) PMF results for moving Lys359 toward CTP in the complexes of wild type (black solid line), G64S (red solid line), and H273R (green solid line) PV $3D^{pol}$ respectively, error bars (around 0.2 kcal/mol) are plotted in a horizontal manner. (b) Representation of the structural elements of PV $3D^{pol}$, motif A is represented by red cartoon, motif D by lime cartoon, Asp238 and CTP by colored sticks, respectively. (c) Distances between OD atom of Asp238 and O2′ atom of CTP, obtained from the free energy calculations of the complexes of wild type (black solid square), G64S (red solid circle), and H273R (green triangle) PV $3D^{pol}$, respectively when Lys359 and CTP are separated at various distances, and their standard deviations are represented by error bars. In this plot, black, red, and green solid thick lines illustrate the mean distances averaged over all free energy calculations of the complexes of wild type, G64S, and H273R PV $3D^{pol}$, respectively.

dx.doi.org/10.1021/ct5006449 | J. Chem. Theory Comput. 2014, 10, 5195−5205

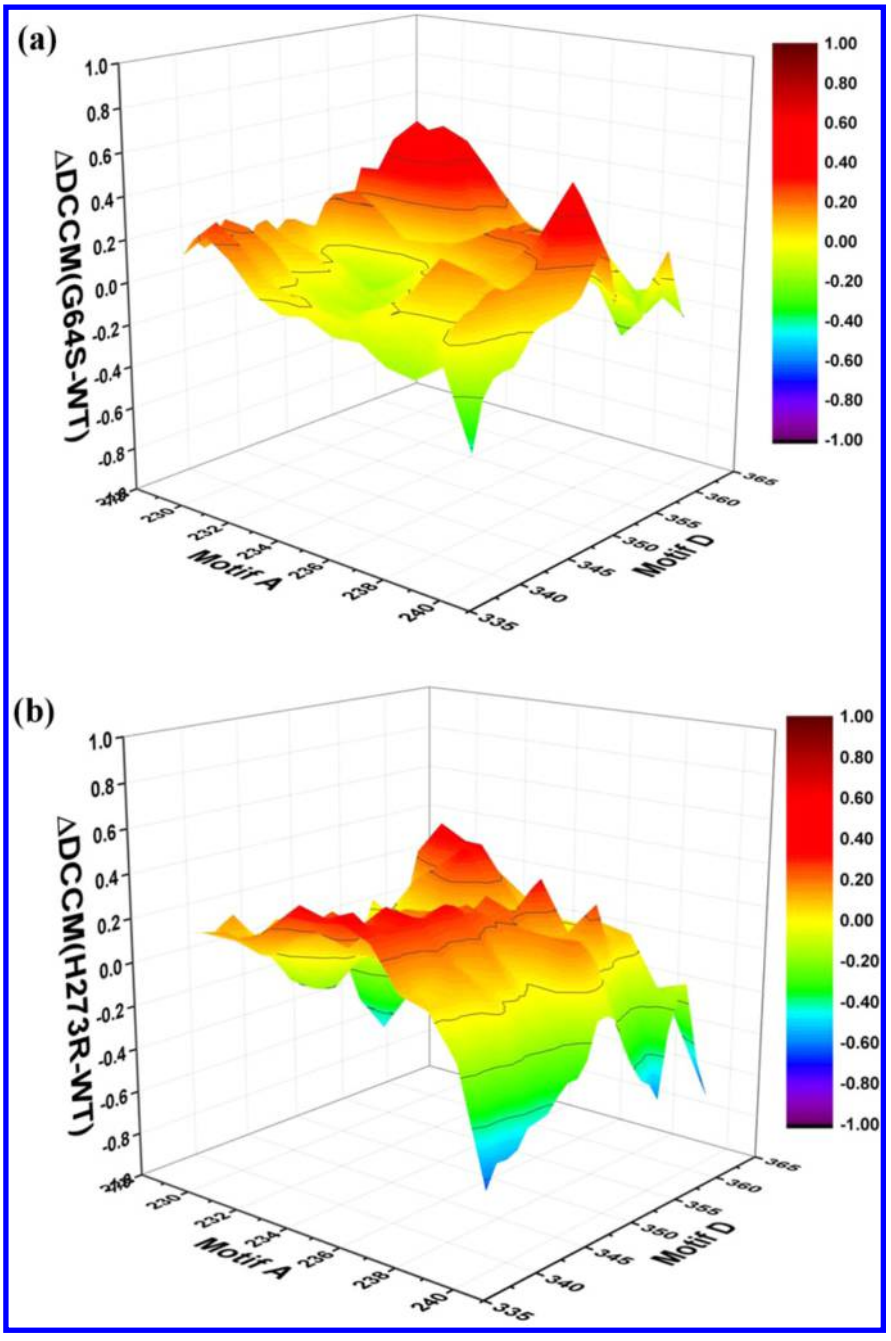


Figure 8. Dynamic cross correlation map (DCCM) results for different polymerases (wild type, G64S, and H273R) with CTP binding. For each case, DCCM results of α carbon atoms were obtained by averaging over five independent MD trajectories. (a) 3D plot of the difference between DCCM results for the complexes of G64S and wild type (WT) PV $3D^{pol}$ with CTP binding, such as, ΔDCCM = DCCM(G64S) − DCCM(WT). (b) 3D plot of the difference between DCCM results for the complexes of H273R and wild type (WT) PV $3D^{pol}$ with CTP binding, such as, ΔDCCM = DCCM(H273R) − DCCM(WT). For each PV $3D^{pol}$ mutant (G64S or H273R), DCCM results only show the correlated motions of motifs A and D.

bonding interactions are reduced between residues 1 and 64 in the high fidelity polymerase (G64S) but are enhanced in the low fidelity polymerase (H273R) as compared to the wild type polymerase. As for the H273R polymerase, due to the mutation from Histidine to Arginine at the residue 273, the hydrogen bonding interactions between residues 273 and 62 would be augmented according to our results shown in Table 4. This might result in strengthening the hydrogen bonding interactions between Gly1 and Gly64 in the case of the H273R polymerase (see Table 3).

The arrangement of Asp238 in PV $3D^{pol}$ should be affected by the change in the hydrogen bonding interactions between residue 1 and residue 64 because the buried N-terminus (Gly1) should be involved in positioning Asp238.[5] From our MD simulations of three polymerases (WT, G64S, and H273R), distances between OD atom of Asp238 and O2′ atom of CTP have been monitored, and it has been observed that Asp238 was differently positioned in different polymerases, shown in Figure 7c. The result suggests that in general the position of Asp238 is located the closest to CTP in the mutant H273R and it is situated the farthest to CTP in the mutant G64S,

dx.doi.org/10.1021/ct5006449 | J. Chem. Theory Comput. 2014, 10, 5195−5205

Table 2. Hydrogen Bonding Strength between the Matching Base Pair (CTP:G) in Three Different PV 3D^{pol} Complexes (Wild Type, G64S, and H273R)[a]

polymerase	donor−acceptor	hydrogen bonding occupancy
wild type	G(N2)−CTP(O2)	84.15%
	G(N1)−CTP(N3)	72.55%
	CTP(N4)−G(O6)	65.50%
G64S	G(N2)−CTP(O2)	79.84%
	G(N1)−CTP(N3)	70.81%
	CTP(N4)−G(O6)	65.56%
H273R	G(N2)−CTP(O2)	84.68%
	G(N1)−CTP(N3)	78.89%
	CTP(N4)−G(O6)	72.46%

[a]The names in parentheses are the names of heavy atoms serving as donors or acceptors in the hydrogen bonding interactions. For each case, hydrogen bonding occupancy is calculated by averaging over five independent MD trajectories. All the details about the calculations of the hydrogen bond occupancy are given in Supporting Information.

Table 3. Hydrogen Bonding Strength between the Residue Pair (Residues 1 and 64) in Three Different PV 3D^{pol} Complexes (Wild Type, G64S, and H273R)[a]

polymerase	donor−acceptor	hydrogen bonding occupancy
wild type	Gly64(N)−Gly1(O)	22.65%
	Gly1(N)−Gly64(O)	14.62%
G64S	Ser64(N)−Gly1(O)	12.40%
	Gly1(N)−Ser64(O)	0.91%
H273R	Gly64(N)−Gly1(O)	31.51%
	Gly1(N)−Gly64(O)	20.59%

[a]The names in parentheses are the names of heavy atoms serving as donors or acceptors in the hydrogen bonding interactions. For each case, hydrogen bonding occupancy is calculated by averaging over five independent MD trajectories. All the details about the calculations of the hydrogen bond occupancy are given in Supporting Information.

Table 4. Hydrogen Bonding Strength between the Residue Pair (Residues 273 and 62) in Three Different PV 3D^{pol} Complexes (Wild Type, G64S, and H273R)[a]

polymerase	donor−acceptor	hydrogen bonding occupancy
wild type	His273(NE2)−Tyr62(OH)	0.03%
G64S	His273(NE2)−Tyr62(OH)	0.11%
H273R	Arg273(NH1)−Tyr62(OH)	14.70%

[a]The names in parentheses are the names of heavy atoms serving as donors or acceptors in the hydrogen bonding interactions. For each case, hydrogen bonding occupancy is calculated by averaging over five independent MD trajectories, and high hydrogen bonding occupancy indicates strong hydrogen bonding strength. All the details about the calculations of the hydrogen bond occupancy are given in Supporting Information.

conforming the structural observation that there is a correlation between the positioning of Asp238 and polymerase function.[5] Meanwhile, this might explain why the hydrogen bonding strength of the base pair (CTP:G) is the strongest in H273R and the weakest in G64S since Asp238 plays a role in stabilizing a NTP in active site. Moreover, it has been mentioned above that, in general, the correlated motions of motif A and motif D become more positive in the high fidelity G64S polymerase and become less positive in the low fidelity H273R polymerase as compared to the wild type polymerase, shown in Figure 8. Thus, it is believed that the positioning of Asp238 is affected not only by the buried N-terminus (Gly1) but also by the interactions between motif A and motif D.

Therefore, according to our computational observations, it becomes clear that a single mutation would cause the following changes in (1) the positioning of Asp238, (2) the strength of the base pair between NTP and RNA template, (3) the correlated motions of motifs A and D, and (4) the accessibility of motif D to the active site. Consequently, the change in the accessibility of motif D to the active site would affect the frequency of misincorporation in PV 3D^{pol} when the prechemistry conformational change is considered as the rate limiting step.

CONCLUSIONS

In this work, we continue our previous work[30] to explore the dynamic properties of PV 3D^{pol} by employing different computational techniques. Through the MD simulations and free energy calculations of different PV 3D^{pol} models, we have made efforts to unearth the dynamic features of the functional structural elements (such as motif A and motif D) in PV 3D^{pol} and to address the inherent correlation between dynamics and function of PV 3D^{pol}. First, our computational results have shown that the overall dynamics of PV 3D^{pol} as a whole was preserved among different ternary complexes but the flexibility of motif D and the correlated dynamics of motifs A and D were both sensitive to the type of the incoming nucleotide triphosphate (NTP). Second, the free energy profiles for moving motif D to access the active site have illustrated that the accessibility of the active site by motif D (in particular Lys359) were strongly correlated with the strength of base paring between the incoming NTP and RNA template, enabling us to explain why the active-site closure in PV 3D^{pol} is normally triggered by a correct NTP binding. Third, our simulation results suggest that the time scale for Lys359 to stay at the active site should be long enough (nanosecond time scale) for the process of proton transfer between Lys359 and NTP but it would be difficult for X-ray experiments to detect it, explaining why no crystal structures are currently available for Lys359 being positioned near the catalytically active site. Finally, we have investigated the dynamic properties of PV 3D^{pol} and its two mutants (high fidelity G64S and low fidelity H273R polymerases) in complex with a correct NTP binding. Our results have shown that a single mutation would observably affect the correlated dynamics of motifs A and D, the base pairing strength between NTP and RNA template, and the accessibility of motif D to the active site, implying the inherent correlation between the frequency of misincorporation in PV 3D^{pol} and the dynamics of motif D and motif A.

ASSOCIATED CONTENT

Supporting Information

Structural description of the RNA-dependent RNA polymerase from poliovirus (PV 3D^{pol}); figures showing the cartoon representation of PV 3D^{pol}, distributions of the distances obtained from free energy calculations of different PV 3D^{pol} complexes, backbone RMSD and radius of gyration (RG) curves of wild type PV 3D^{pol} in different complexes, dynamic cross correlation map (DCCM) results for wild type PV 3D^{pol} in different complexes and representations of structural elements of wild type PV 3D^{pol}. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*E-mail: ghli@dicp.ac.cn.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work is supported by grants from the National Science Foundation of China under Contracts No. 31370714 and No. 91227126, the Ministry of Science and Technology of China (863 project No. 2012AA01A305; 973 project No. 2012CB721002), and "Hundreds Talent Program" of Chinese Academy of Sciences. P. R. China.

REFERENCES

(1) Vignuzzi, M.; Stone, J. K.; Arnold, J. J.; Cameron, C. E.; Andino, R. *Nature* **2006**, *439*, 344−348.

(2) Richards, A. L.; Jackson, W. T. *PLoS Pathog.* **2012**, *8*, e1003046.

(3) Crotty, S.; Maag, D.; Arnold, J. J.; Zhong, W.; Lau, J. Y. N.; Hong, Z.; Andino, R.; Cameron, C. E. *Nat. Med.* **2000**, *6*, 1375−1379.

(4) Hansen, J. L.; Long, A. M.; Schultz, S. C. *Structure* **1997**, *5*, 1109−1122.

(5) Thompson, A. A.; Peersen, O. B. *EMBO J.* **2004**, *23*, 3462−3471.

(6) Kuchta, R. D.; Mizrahi, V.; Benkovic, P.; Benkovic, S. J. *Biochemistry* **1987**, *26*, 8410−8417.

(7) Patel, S. S.; Wong, I.; Johnson, K. A. *Biochemistry* **1991**, *30*, 511−525.

(8) Kati, W. M.; Johnson, K. A.; Jerva, L. F.; Anderson, K. S. *J. Biol. Chem.* **1992**, *267*, 25988−25997.

(9) Capson, T. L.; Peliska, J. A.; Kaboord, B. F.; Frey, M. W.; Lively, D.; Dahlberg, M.; Benkovic, S. J. *Biochemistry* **1992**, *31*, 10984−10994.

(10) Arnold, J. J.; Cameron, C. E. *Biochemistry* **2004**, *42*, 5126−5137.

(11) Arnold, J. J.; Gohara, D. W.; Cameron, C. E. *Biochemistry* **2004**, *42*, 5138−5748.

(12) Zinnen, S.; Hsieh, J. C.; Modrich, P. *J. Biol. Chem.* **1994**, *269*, 24195−24202.

(13) Arora, K.; Beard, W. A.; Wilson, S. H.; Schlick, T. *Biochemistry* **2005**, *44*, 13328−13341.

(14) Rothwell, P. J.; Mitaksov, V.; Waksman, G. *Mol. Cell* **2005**, *19*, 345−355.

(15) Eger, B. T.; Kuchta, R. D.; Carroll, S. S.; Benkovic, P. A.; Dahlberg, M. E.; Joyce, C. M.; Benkovic, S. J. *Biochemistry* **1991**, *30*, 1441−1448.

(16) Castro, C.; Arnold, J. J.; Cameron, C. E. *Virus Res.* **2005**, *107*, 141−149.

(17) Yang, X.; Smidansky, E. D.; Maksimchuk, K. R.; Lum, D.; Welch, J. L.; Arnold, J. J.; Cameron, C. E.; Boehr, D. D. *Structure* **2012**, *20*, 1519−1527.

(18) Castro, C.; Smidansky, E.; Maksimchuk, K. R.; Arnold, J. J.; Korneeva, V. S.; Gotte, M.; Konigsberg, W.; Cameron, C. E. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 4267−4272.

(19) Castro, C.; Smidansky, E. D.; Arnold, J. J.; Maksimchuk, K. R.; Moustafa, I.; Uchida, A.; Götte, M.; Konigsberg, W.; Cameron, C. E. *Nat. Struct. Mol. Biol.* **2009**, *16*, 212−218.

(20) Gong, P.; Peersen, O. B. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 22505−22510.

(21) Bull, J. J.; Sanjuan, R.; Wilke, C. O. *J. Virol.* **2007**, *81*, 2930−2939.

(22) Pfeiffer, J. K.; Kirkegaard, K. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 7289−7294.

(23) Pfeiffer, J. K.; Kirkegaard, K. *PLoS Pathog.* **2005**, *1*, e11.

(24) Arnold, J. J.; Vignuzzi, M.; Stone, J. K.; Andino, R.; Cameron, C. E. *J. Biol. Chem.* **2005**, *280*, 25706−25716.

(25) Korneeva, V. S. Residue Arg-273 as a modulator of the polymerase fidelity. In: Poliovirus RNA-dependent RNA polymerase (in)fidelity: mechanisms consequences and applications. Ph.D. Thesis, The Pennsylvania State University, 2007; pp 155−211.

(26) Smidansky, E. D.; Arnold, J. J.; Cameron, C. E. In *Origin and Evolution of Viruses*, 2nd ed.; Domingo, E., Parrish, C. R., Holland, J. J., Eds.; Elsevier, Ltd.: Amsterdam, 2008; pp135−160.

(27) Cameron, C. E.; Moustafa, I. M.; Arnold, J. J. *Curr. Opin. Struct. Biol.* **2009**, *19*, 768−774.

(28) Moustafa, I. M.; Shen, H.; Morton, B.; Colina, C. M.; Cameron, C. E. *J. Mol. Biol.* **2011**, *410*, 159−181.

(29) Shen, H.; Moustafa, I. M.; Cameron, C. E.; Colina, C. M. *J. Phys. Chem. B* **2012**, *116*, 14515−14524.

(30) Shen, H.; Sun, H.; Li, G. *PLoS Comput. Biol.* **2012**, *8*, e1002851.

(31) Case, D. A.; Cheatham, T. E., III; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. *J. Comput. Chem.* **2005**, *26*, 1668−1688.

(32) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. *Proteins* **2006**, *3*, 712−725.

(33) Perez, A.; Marchan, I.; Svozil, D.; Sponer, J.; Cheatham, T. E., III; Laughton, C. A.; Orozco, M. *Biophys. J.* **2007**, *92*, 3817−3829.

(34) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−935.

(35) Berendsen, H. J.; Postma, J. P. M.; Van Gunsteren, W. F.; Dinola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684−3690.

(36) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327−341.

(37) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089−10092.

(38) Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. *J. Comput. Chem.* **1992**, *13*, 1011−1021.

(39) Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. *J. Comput. Chem.* **1995**, *16*, 1339−1350.

(40) Bartels, C. *Chem. Phys. Lett.* **2000**, *331*, 446−454.

(41) Wang, D.; Bushnell, D. A.; Westover, K. D.; Kaplan, C. D.; Kornberg, R. D. *Cell* **2006**, *127*, 941−954.

(42) Da, L.-T.; Pardo, A. F.; Wang, D.; Huang, X. *PLoS Comput. Biol.* **2013**, *9*, e1003020.