BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

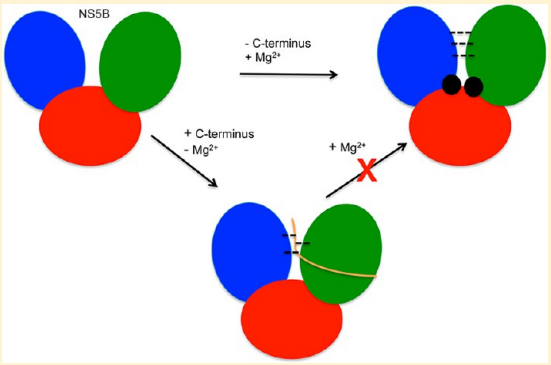
Molecular Simulations Illuminate the Role of Regulatory Components of the RNA Polymerase from the Hepatitis C Virus in Influencing Protein Structure and Dynamics

Brittny C. Davis and Ian F. Thorpe*

Department of Chemistry and Biochemistry, University of Maryland, Baltimore County, Baltimore, Maryland 21250, United States

Supporting Information

ABSTRACT: The RNA polymerase (gene product NS5B) from the hepatitis C virus is responsible for replication of the viral genome and is a validated drug target for new therapeutic agents. NS5B has a structure resembling an open right hand (containing the fingers, palm, and thumb subdomains), a hydrophobic C-terminal region, and two magnesium ions coordinated in the palm domain. Biochemical data suggest that the magnesium ions provide structural stability and are directly involved in catalysis, while the C-terminus plays a regulatory role in NS5B function. Nevertheless, the molecular mechanisms by which these two features regulate polymerase activity remain unclear. To answer this question, we performed molecular dynamics simulations of NS5B variants with different C-terminal lengths in the presence or absence of magnesium ions to determine the impact on enzyme properties. We observed that metal binding increases both the magnitude and the degree of correlated enzyme motions. In contrast, we observed that the C-terminus restricts enzyme dynamics. Under certain conditions, our simulations revealed a fully closed conformation of NS5B that may facilitate *de novo* initiation of RNA replication. This knowledge is important because it fosters the development of a comprehensive description of RNA replication by NS5B and is relevant to understanding the functional properties of a broad class of related RNA polymerases such as 3D-pol from poliovirus. Ultimately, this information may also be pertinent to designing novel NS5B therapeutics.



The hepatitis C virus (HCV) has infected 170 million people worldwide, and approximately 4 million people within the United States.[1,2] There is no cure for HCV infection, and 25% of infected individuals contract chronic liver ailments such as cirrhosis or liver cancer.[1,3] The current standard of care, ribavirin and pegylated interferon, is not effective across the six genotypes of HCV and can have severe side effects.[4] Thus, better treatments are sorely needed. The RNA polymerase from HCV (gene product NS5B) is a validated drug target because it is crucial for replication of the viral genome.[1,5−11] Consequently, it is important to understand the mechanisms by which the activity of this enzyme is modulated.

Among the components of NS5B associated with changes in activity are the C-terminal residues and the presence of magnesium ions. NS5B has the right-handed structural organization into fingers, palm, and thumb domains that is typical of viral RNA polymerases.[8,12] NS5B contains 591 residues; the last 60 (residues 532−591) hydrophobic C-terminal residues are thought to be associated with the membrane of the endoplasmic reticulum (ER) *in vivo* and cause the enzyme to be insoluble *in vitro*.[13,14] Consequently, biochemical studies often make use of NS5B variants in which the C-terminus is truncated.[14,15] This issue also impacts structural studies, with most of the available crystal structures generated for truncated versions of the enzyme. However, NS5B with C-terminal truncations is replication competent and exhibits RNA polymerase activity.[13,14] *In vitro* studies indicate that the presence of the C-terminus decreases enzyme activity.[14] In the absence of membrane association, residues 532−544 wrap around the thumb domain while residues 545−562 occupy the interface between all three domains, interacting closely with several residues from the fingers and thumb (Figure 1).[14,15]

We note that this conformation of the C-terminus of the enzyme occurs in the crystal structures used to initiate our present simulation studies (see Materials and Methods) and may also occur in other *in vitro* situations. Many structural studies and activity assays of NS5B are conducted under conditions where this conformation of the C-terminus is possible. In addition, much of what we currently know about the enzyme's structure and function is derived from such studies. Thus, it is important to understand how this conformation of the C-terminus modulates the structure and function of the enzyme.

Apart from C-terminal residues, NS5B activity is also modulated by the presence of magnesium ions, which are

Received: February 27, 2013
Revised: May 31, 2013
Published: June 5, 2013

   dx.doi.org/10.1021/bi400251g | *Biochemistry* 2013, 52, 4541−4552

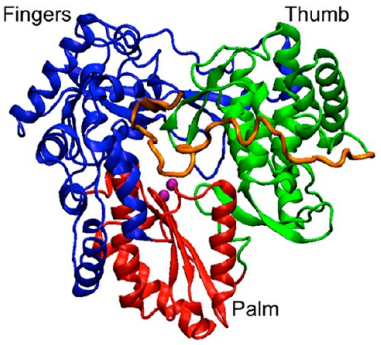


Figure 1. Average structure from simulation of system **4b** in cartoon representation. Fingers, palm, and thumb domains are colored blue, red, and green, respectively. C-Terminal residues 532−563 are colored orange. Magnesium ions are shown in purple space-filling representation. The C-terminal residues lodge between the fingers and thumb domains, preventing the two domains from interacting directly.

required for RNA replication *in vitro*.[16−18] Structural studies indicate two magnesium ions can be coordinated in the palm domain. The metal ions stabilize the enzyme structure and create a favorable geometry of the active site for the initiation of replication.[16,17] The magnesium ions are coordinated by three aspartic acid residues (220, 318, and 319) and are believed to be directly involved in catalysis.[16,17]

NS5B is thought to replicate RNA in two stages: initiation and elongation.[19] These stages are associated with the closed and open conformations of the enzyme, respectively.[20,21] However, while crystal structures of both conformations have been seen in HCV subtype **2a** [Protein Data Bank (PDB) entries 1YUY and 1YV2], only the closed conformation has been seen for subtype **1b**.[22] In the closed conformation, the fingers and thumb domains are rotated toward each other, positioning residues in the template channel for initiation. During elongation, the thumb domain rotates away from the fingers domain in the open conformation, opening the exit channel for nascent RNA at the front of the enzyme.[23] Transitions between the open and closed conformation are thought to be regulated by the Δ1 loop (residues 20−35).[23] Chinnaswamy and associates deleted the tip of the Δ1 loop (residues 26−30) and used electron microscopy to confirm that NS5B remained in the open conformation and did not sample the closed conformation.[23]

To date, crystallographic and biochemical experiments have provided a wealth of information about the structure and activity of NS5B. However, these studies are unable to provide a molecular level understanding of the roles the different components of the enzyme play in modulating activity. The structural intermediates involved in RNA replication have not been fully determined, and much about the process of replication remains unclear, such as the role that the C-terminus or metal ions play in altering the structure and dynamics of the enzyme. Furthermore, the conformational changes necessary for RNA replication suggest that specific motions associated with conformational transitions are required for NS5B to function.

In this article, we employ molecular dynamics (MD) simulations to understand how the C-terminus of the enzyme and the presence of magnesium ions influence the structural and dynamic properties of NS5B. We performed MD simulations of NS5B with different C-terminal lengths (531, 562, and 563 residues), both with and without magnesium ions bound (see Table 1 for a description of each system). We

Table 1. Simulated Enzyme Systems[a]

system	PDB entry	no. of residues	description	no. of atoms simulated (total atoms, waters, ions)
1a	2WHO	531		75448, 22381, 15
1b	2WHO	531	with Mg^{2+} ions	75442, 22377, 19
2a	2HAI	562	L47Q, F101Y, K114N	75687, 22306, 15
2b	2HAI	562	L47Q, F101Y, K114N; with Mg^{2+} ions	75645, 22290, 19
3a	3CO9	562		78190, 23140, 15
3b	3CO9	562	with Mg^{2+} ions	78031, 23085, 19
4a	3HHK	563		82905, 24708, 15
4b	3HHK	563	with Mg^{2+} ions	82875, 24696, 19

[a]Residues after position 531 are considered C-terminal residues (see Figure 1). In any comparative analysis, **1b** refers to the data of the fully closed conformation unless otherwise specified (e.g., intermediate **1b**).

observed that the absence of C-terminal residues in conjunction with the presence of magnesium ions induced the enzyme to exhibit a new conformation that is more closed than that observed in the crystallographic coordinates or in our previous simulation studies.[24] This observation suggests that the crystal structures may not represent the most closed conformation of the enzyme, as originally thought. In our previous studies, we described a "hyper-closed" conformation observed for a structure of NS5B containing an allosteric inhibitor bound to the thumb domain. However, the closed conformation observed here is different from that observed in our previous work.[24] We do not believe that the hyper-closed conformation described for the ligand-bound enzyme allows the enzyme to be active.[24] In contrast, we provide evidence that the fully closed conformation observed in this study is a functionally relevant conformational state. This enzyme conformation may represent the initial state to which substrates bind to begin one round of replication of the linear HCV genome. Our studies suggest that the C-terminal residues prevent the achievement of a completely closed conformation and thus likely inhibit the initiation of replication. The results of this study help us to understand how the activity of NS5B is regulated, information that is relevant to illuminating the function of related viral polymerases such as the 3D-pol from poliovirus and to identifying new ways to inhibit NS5B.

MATERIALS AND METHODS

Structure Preparation and MD Simulation. Crystal structures of genotype **1b** HCV NS5B from the PDB with the following PDB entries were used as the starting coordinates: 2WHO, 2HAI, 3CO9, and 3HHK. 2WHO lacks the C-terminus and contains 531 amino acids. 2HAI and 3CO9 both have 562 residues, and 3HHK has 563. 2HAI also has three point mutations: L47Q, F101Y, and K114R. Each structure contains an allosteric inhibitor bound to a different location of the enzyme. Both 2WHO and 2HAI contain a non-nucleoside inhibitor in the NNI2 thumb binding site at the base of the thumb, ∼30 Å from the active site. 3CO9 and 3HHK contain inhibitors in the NNI3 and NNI4 binding sites, respectively, located in the palm domain adjacent to the active site.

dx.doi.org/10.1021/bi400251g | *Biochemistry* 2013, 52, 4541−4552

Inhibitors were removed from each structure before simulations were begun.

These structures were chosen to represent diverse NS5B constructs and initial coordinates, and we are in the process of simulating these constructs with their inhibitors present to determine the effect of allosteric inhibitors in the various binding sites. In addition, the 2WHO structure contains two manganese ions bound to the divalent ion site and was used for placement of the magnesium ions in each structure. The manganese ions were replaced with magnesium and the metal ions placed in the other constructs via alignment with 2WHO. Simulations were conducted with enzymes in both apo and metal-bound forms.

All structures were prepared for simulation by removing nonprotein atoms from the coordinates and adding hydrogen atoms using GROMACS version 4.0.7.[25] In cases where the crystal structure contained more than one chain in the unit cell, chain A was selected for simulation. Each structure was placed in a truncated octahedral unit cell that was larger than the protein by 12 Å in each dimension, and explicit SPC water[26] molecules were added. To neutralize the system, 15 chloride ions were added to the apoenzyme systems and 19 chloride ions to the enzymes containing magnesium ions. The total number of solvent and solute atoms present in each system is provided in Table 1. The OPLS/all-atom force field[27] was employed to describe inter- and intramolecular interactions.

Minimization of the solvated protein coordinates was conducted for 50000 steps using the steepest descent algorithm and applying periodic boundary conditions. All covalent bonds to hydrogen atoms were constrained using the SHAKE algorithm,[28] and electrostatic interactions were calculated by the particle mesh Ewald method (PME).[29] Ten angstroms was used as the Coulomb cutoff and the nonbonded pair list.

This was followed by equilibration in the *NPT* ensemble for 3−5 ns using a 2 fs time step, during which Parrinello−Rahman coupling was applied to maintain the pressure at 1.01 bar. Once the volume of the unit cell had stabilized, a snapshot of the *NPT* equilibration in which the pressure was closest to 1.01 bar was written out. Further MD simulations were performed using the *NVT* ensemble for 400−700 ns. The v-rescale thermostat was used to maintain a temperature of 300 K. The last 200 ns of equilibrated data was used for data analysis. Snapshots of the MD simulations were collected every 10 ps for data analysis, and VMD was used to view the resulting trajectories.

During the *NVT* simulations of 3CO9, the pressure rose above 1.01 bar. The secondary and tertiary structures of the enzyme were maintained upon visual inspection. Thus, another *NPT* simulation was conducted to re-equilibrate the pressure. *NVT* simulations were then conducted as described above.

Root-Mean-Square Fluctuation. Root-mean-square fluctuations (rmsfs) were calculated for Cα atoms to determine the flexibility of each residue during the equilibrated simulations (eq 1). After each snapshot had been superimposed onto the initial minimized structure, the reference position of each atom ($\tilde{x}_l$) was subtracted from the instantaneous position of the atom in each snapshot of the trajectory $[x_i(t_j)]$

$$\mathrm{rmsf}_i = \sqrt{\frac{\sum_{t_j=}^{T} [x_i(t_j) - \tilde{x}_l]^2}{T/\mathrm{d}t}} \quad (1)$$

where t_j is the time of the jth snapshot, dt is the frequency at which coordinates were written, and T is the total simulation time.

Covariance Analysis. Covariance matrices (eq 2) were calculated using the "covar" utility in GROMACS. The covariance matrix allows the viewing of correlated motions sampled by atoms i and j during the MD trajectory. The normalized covariance between two atoms (C_{ij}) is determined by taking the product of the difference between instantaneous positional coordinate r_i or r_j and average position $\langle r_i \rangle$ or $\langle r_j \rangle$, where r includes the x, y, and z directions:

$$C_{ij} = \langle (r_i - \langle r_i \rangle)(r_j - \langle r_j \rangle) \rangle / \sqrt{\langle (r_i - \langle r_i \rangle)^2 \langle (r_j - \langle r_j \rangle)^2} \quad (2)$$

The covariance matrix is normalized as shown in eq 2 so that atoms with completely positively correlated motion display C_{ij} values of 1, while atoms with completely negatively correlated motions display C_{ij} values of −1. If two atoms are not correlated or move in orthogonal directions, their C_{ij} values will be zero. Cα atoms were selected for this analysis to represent the overall motion of each residue. The covariance matrix was viewed using the "colormap" utility in MATLAB.

Principle Component Analysis (PCA). PCA diagonalizes the covariance matrix (C_{ij}) to generate eigenvalues (λ), using the eigenvector matrix ($\mathbf{V}$) (eq 3).

$$\mathbf{V} \cdot C_{ij} \cdot \mathbf{V}^{\mathbf{T}} = \mathrm{diag}[\lambda_1, \lambda_2, \lambda_3, \ldots] \quad (3)$$

Together, the eigenvalues and eigenvectors comprise the principle components (PCs) of the protein motion. The eigenvectors describe the direction of the motions sampled during the trajectories, while the eigenvalues describe the magnitude of these motions. The PCs are analogous to vibrational modes from a normal mode analysis.[30] Typically, the first PC (λ_1) accounts for most of the motion, while the second (λ_2) accounts for a smaller amount, etc.[31] These modes can be used to identify the inherent fluctuations present in the trajectory.

Projection of Trajectories onto Principle Components. Projections of the trajectories unto the PCs were conducted to visualize the conformational space being sampled by the enzyme during the simulations. These were done using the "anaeig" utility in GROMACS. Protein coordinates from the trajectories, $x(t)$, were compared to a reference structure, $\langle x \rangle$, and then projected onto eigenvectors $\mathbf{V}^{\mathbf{T}}$ from PCA in eq 4.[25] The value of P is zero when the coordinates in $x(t)$ are identical to $\langle x \rangle$.

$$P(t) = \mathbf{V}^{\mathbf{T}}(x(t) - \langle x \rangle) \quad (4)$$

We chose to project our trajectories onto two vectors (PCs 1 and 2) from the simulation of **1b**, because they represent the largest amplitude fluctuations observed in the trajectories. In addition, because **1b** contains magnesium ions but not the C-terminus (see Discussion), we believe these motions are most likely to be functionally relevant.[32,33] These PCs include 400 ns of data representing both the intermediate and fully closed conformations sampled by this enzyme. Root-mean-square inner product (rmsip) values are employed to evaluate the degree of similarity between PC vector spaces obtained from each simulation. rmsip values demonstrate that there is little similarity between the lowest-frequency eigenvectors representing the intrinsic motions from different systems. To perform both the projections and the rmsip calculations, we truncated the systems containing C-terminal residues to 531 residues to match **1** (see Table 1). For the remainder of the paper, use of a

dx.doi.org/10.1021/bi400251g | *Biochemistry* 2013, 52, 4541−4552

single number denotes both trajectories initiated from this structure; i.e., **1** indicates both **1a** and **1b**.

Solvent Accessible Surface Area. The solvent accessible surface area (SASA) was calculated for the residues lining the template channel (14, 15, 93−98, 137−141, 158−162, 168, 224, 225, 269, 282−291, 317, 318, 404, 405, and 444−451) as a measure of the accessibility of the template channel to the environment. These residues were identified using crystal structure 1NB7,[34] which contains a four-nucleotide fragment of RNA template. Any amino acids that were <5 Å from any nucleotide atom in crystal structure 1NB7 were chosen for the SASA calculation during the simulation. The "sas" utility in GROMACS was used for this calculation, and a probe radius of 1.4 Å was employed.

Interdomain Angle. The angle between the domains was determined by calculating the angle between the two vectors connecting the centers of masses of (i) the palm and fingers domains and (ii) the palm and thumb domains. The residues in each domain were selected according to the definition of Lesburg et al. as follows: fingers, 1−188 and 227−287; palm, 189−226 and 288−370; thumb, 371−529.[35]

RESULTS

Table 1 describes the systems that were simulated, while Table 2 describes other relevant enzyme structures that will be discussed throughout this work. The last 200 ns of equilibrated data was used for data analysis of each system. The structures containing C-terminal residues equilibrated around 3 Å root-mean-square deviation (rmsd) from the initial minimized structure, while **1** (which lacks C-terminal residues) equilibrated around 4 Å (see Figure 7). The rmsd plots for systems containing C-terminal residues are provided in SI-1 of the Supporting Information. Data analysis suggests that the presence of the C-terminus suppresses the impact of magnesium ions on the enzyme's properties when both were present. Therefore, we will focus on data for **1**, the structure lacking the C-terminus, to demonstrate the full effects of the magnesium ions, as the effect of metal binding was most pronounced when these systems were compared. In addition, system **1b** revealed a new conformation, which we will term "fully closed".

Table 2. Enzyme Structures Discussed throughout This Work

abbreviation	PDB entry	description
1*	2WHO	crystallographic coordinates
1f	2WHO	average structure from previous simulations of the free enzyme[24]
1v	2WHO	average ligand-bound structure from previous simulations[24]
2*	2HAI	crystallographic coordinates
3*	3CO9	crystallographic coordinates
4*	3HHK	crystallographic coordinates

Impact of the C-Terminus. *C-Terminal Interactions Create a More Rigid Enzyme That Samples Less Conformational Space.* The C-terminus lodges between the fingers and thumb domains, interfering with communication between the two domains (Figure 1). We observed interactions in our simulations consistent with contacts observed by Adachi et al.[14] However, because of the extreme flexibility of the C-terminus, we also observed additional interactions not noted previously. In particular, for **4**, we observed an additional hydrogen bond between the hydroxyl H of Y176 and the hydroxyl O of S563 (3.3 ± 1.57 Å). Overall, the interactions between the C-terminus and the fingers and thumb domains restrict the motion of the entire enzyme.

A comparison of the summed eigenvalues for our simulated structures is given in Table 3 and shows that the presence of the C-terminus decreases the overall enzyme flexibility. The structure of **1b** is also subdivided into separate intermediate and fully closed sections. If one combines these two sections into one unit, **1b** displays the largest amount of conformational sampling of all the simulated systems. While **2** and **3** have the same number of residues, the C-terminus has a different effect on enzyme flexibility in the two cases. **2** is less flexible than **3**, which may result from the mutations in **2** (see Table 1). These mutations exchange residues with smaller side chains for larger side chain residues; this may cause steric hindrance and contribute to the enzymes' decreased flexibility. Structure **4** samples the least conformational space, which is likely due to the hydrogen bond between Y176 and S563 that is noted above imparting additional structural rigidity. Note that **3** has the same sequence as **4** but contains only 562 residues. It is the missing C-terminal serine (S563) that participates in the indicated hydrogen bond in **4**. Thus, **3** serves as a control for the effect of this hydrogen bond in **4**. By inspecting Table 3, one can see that **3** does in fact display greater flexibility than **4**.

Table 3. Summed Eigenvalues (ΣΛ) for Cα Atoms in Each System[a]

system	ΣΛ (Å^2)	15/1593 (%)
1a	6042.13	65.26
intermediate **1b**	7427.09	67.03
fully closed **1b**	7386.18	64.88
total **1b**	15917.98	81.61
2a	6100.25	58.72
2b	6823.20	62.03
3a	7056.11	64.84
3b	8051.91	71.21
4a	5206.16	58.62
4b	3419.11	48.63

[a]For **2**−**4**, the C-terminal residues (532−563) were excluded from the calculation to directly compare those systems to **1**. In general, structures with C-terminal residues display reduced flexibility that we can associate with a reduced level of conformational sampling that is not increased by metal binding. The third column displays the percentage of the summed eigenvalues due to the 1% (15/1593) of the lowest-frequency principal components and illustrates how much of the overall motion is due to these modes.

Projecting the protein coordinates on the two lowest-frequency PCs from **1b** shows that enzymes with C-terminal residues (**2**−**4**) explore less of the conformational space represented in these PCs than either **1a** or **1b** (Figure 2). This is consistent with the decreased enzyme flexibility observed for these simulations (Table 3). This observation demonstrates the ability of the C-terminus to inhibit conformational sampling. The plots also demonstrate that enzymes containing the C-terminus are in conformations very similar to the initial crystal structures and unlike those sampled by **1b**. A table of root-mean-square inner products (rmsips) and the percentage of cosine content for the 1% of the lowest-frequency modes (15/1593) are provided as SI-2 and SI-3 of the Supporting Information, respectively. The rmsip values indicate

dx.doi.org/10.1021/bi400251g | Biochemistry 2013, 52, 4541−4552

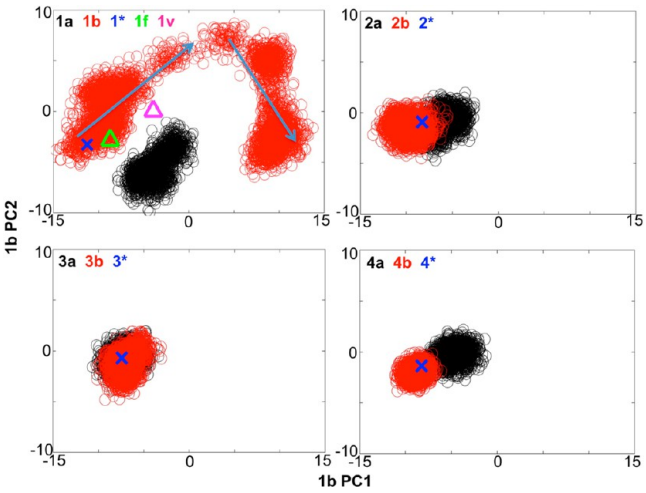


Figure 2. Projections of the protein coordinates from each system onto PCs 1 and 2 from the **1b** trajectory. These PCs were chosen because they represent the most overall motion of all the PCs for each of the trajectories. Each point on the plot represents a single snapshot from the simulation. The black circles represent the apo structures, while the red circles denote the metal-bound structures. The blue X's represent the respective X-ray coordinates. The light blue arrows indicate time progression. The plot shows that **1a** and **1b** explore more of the conformational space represented by these two PCs than the other systems, which contain C-terminal residues. **1b**, which contains magnesium ions, also explores more space than **1a**. The data on the left side of the plot for **1b** corresponds to earlier simulation times while data on the right corresponds to later times. As the simulation proceeds (time progression indicated by arrows), the enzyme becomes more closed. In addition, the positions of the average structures from our previous simulations of 2WHO in a free state (**1f**) and bound to its ligand (**1v**) are shown for comparison to the study presented here. **1f** is much like **1*** and the initial conformations sampled by **1b**. However, **1v** appears to be in a conformation unlike any of the other conformations sampled by **1a**, **1b**, or **1f**.

that the low-frequency modes sampled in each of the trajectories differ substantially, while the low cosine content observed indicates these modes in general do not simply represent random diffusive motions.

The C-Terminus Alters Correlated Motions between Fingers and Thumb Domains. Covariance analysis was used to identify correlated motions occurring in the equilibrated enzyme simulations (see Figure 3). We will focus on the correlations observed for structures **1b**, **2b**, **3b**, and **4b**, as we expect the metal-bound enzymes to be the most functionally relevant. The disorder and overall flexibility of the C-terminus cause it to associate with different parts of the enzyme in each system. This is demonstrated by the varying patterns of correlation experienced by the C-terminus in each system.

The most abundant motions in simulation **1b** are anticorrelated motions between the fingers and thumb domains. These motions make up the bulk of the first PC of the trajectory. Visualization of the first PC shows that the anticorrelated motion between the fingers and thumb domains involves the top of these domains moving toward and away from each other over the course of the simulation. A correlation map of the motions observed during PC1 is shown in SI-4 of the Supporting Information to demonstrate that this PC is largely responsible for this pattern of correlations. Over time, the enzyme structure becomes more closed, as indicated in Figure 2. The closed enzyme conformation likely favors initiation of RNA replication.[23]

Structures **2b** and **3b** display the most extensive pattern of correlated motions overall. In particular, they both exhibit more anticorrelated motion between the fingers and thumb domains than does **1b**. While the two domains still move toward or away from each other, the anitcorrelated motions are between different residues from the fingers and thumb domains compared to the motions observed for **1b**. For **2b**, thumb residues 440−455 (also known as the β-loop) strongly interact with the C-terminus, as do residues from the fingers domain. Thus, the β-loop is forced to move in concert with the fingers domain. This loop has been associated with the correct positioning of the RNA template.[36] If the β-loop is otherwise engaged by interacting with the C-terminus, it may not be available to coordinate the incoming template RNA strand. The first mode for **2b** shows motions throughout the entire enzyme, with most of the motion coming from the early part of the C-terminus (residues 532−544). There is also a shift in the enzyme's conformation, as the bottom halves of the fingers and thumb domains move toward each other.

3b experiences slightly different anticorrelated motions between the fingers domain and the palm and thumb domains compared to those of **2b**, as seen in Figure 3. The lowest-frequency PC for **3b** indicates that the C-terminus is extremely flexible and moves in concert with the fingers domain. The fingers and thumb domains are tightly associated with the latter part of the C-terminus. As the trajectory progresses, the domains become wedged apart even more by a cranking motion of the early C-terminal residues. Overall, this causes the enzyme conformation to become more open.

Finally, **4b** does not display widespread correlated motions and is also the least flexible. This is due to the hydrogen bond between the additional C-terminal residue S563 and residue Y176 noted previously. This additional interaction increases the rigidity of the enzyme, decreasing the magnitude of fluctuations that occur as indicated by the summed eigenvalues (Table 3). Most of the motion represented in the lowest-frequency PC of **4b** occurs in the rear of the fingers domain in residues 140−160 (Δ2 loop). The C-terminus along with the rest of the enzyme is seen to be extremely rigid.

Overall, the motions observed for structures containing C-terminal residues differ greatly from those of **1b**. Although **2b** and **3b** demonstrate anticorrelated motions between the thumb and fingers domains, these motions appear to be governed by the fluctuations of the C-terminus, which interacts with both domains. As previously stated, we expect that the distinct pattern of anticorrelated motions observed between the fingers and thumb domains for **1b** is most likely to represent functional motions because **1b** lacks C-terminal residues and these residues are known to decrease enzyme activity *in vitro*.[14]

The C-Terminus Changes the Electrostatic Potential of the Enzyme Surface. Figure 4 shows the electrostatic potential on the front surface of NS5B for enzymes **1*** (which does not possess C-terminal residues) and **3*** (which does contain a C-terminus). The template and duplex channels have a very positive potential (Figure 4A), while the C-terminus is extremely negative (Figure 4B), causing the two to strongly interact. The negative potential from the C-terminus near the template and duplex channels may create a less favorable environment for RNA template binding and duplex RNA exiting the enzyme (Figure 4C) by destabilizing the negative phosphate groups on the RNA backbone. We note that the

dx.doi.org/10.1021/bi400251g | *Biochemistry* 2013, 52, 4541−4552

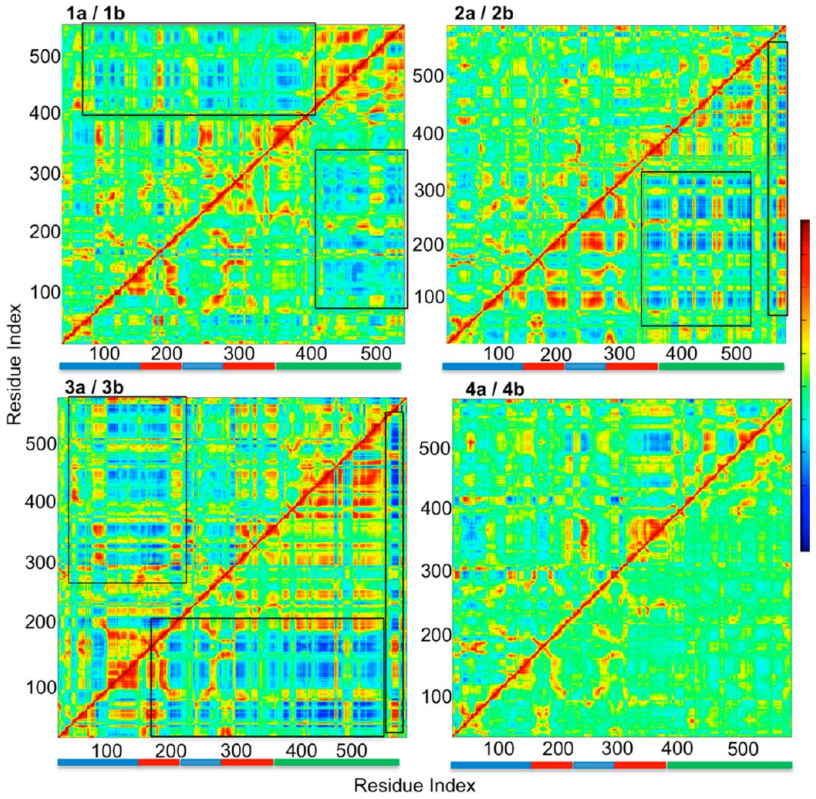


Figure 3. Correlation maps computed for each enzyme simulation. The apo (a) systems are plotted in the top left triangles, while the metal-bound systems (b) are plotted in the bottom right triangles. All the proteins with C-terminal residues display motions that differ from those observed when the C-terminus is absent (**1a** and **1b**). Positive correlations are denoted by red, while anticorrelations (negative correlations) are denoted by blue. As shown by the color scale to the right, computed values have been normalized as described in the text to lie between 1 and −1.

presence of the C-terminus by itself does not preclude template binding, as our simulations suggest the template channel is wide enough to accommodate a template strand even if the C-terminus is present. Consistent with this finding, crystal structure 1NB7 containing a short template strand as well a C-terminal residues demonstrates that both can occupy the template channel simultaneously.[34] However, in 1NB7, the C-terminus interacts directly with the template: we suggest that such interactions may inhibit replication by destabilizing interactions between the template and the remainder of enzyme and preventing the template from reaching the active site. This phenomenon may reflect yet another way in which the C-terminus decreases enzyme activity.

Impact of Metal Binding. *Changes to Structure and Flexibility.* There are immediate structural effects that occur upon metal binding. The catalytic triad (Asp220, -318, and -319) coordinates the magnesium ions within the active site. In the absence of the metal ions, these residues are free to interact with other residues. Figure 5 shows how residues within and around the active site are affected upon metal ion binding. In **1a**, $O^{\delta 1}$ of Asp220 and H^{ε} of Arg222 are observed to rotate toward each other and share a hydrogen bond. Similarly, $O^{\delta 2}$ of Asp319 flips toward H^{α} of Cys366 to form a hydrogen bond. Finally, Asp318 appears to freely sample several conformations rather than interacting with one specific residue. These interactions subsequently alter the dynamics of the enzyme. However, they do not occur in **1b** when the metal ions are present. Projecting trajectory snapshots on principal components (PCs) extracted from the **1b** trajectory (see Figure 2) demonstrates that when metal is bound (**1b**) the enzyme samples conformations distinct from those sampled in the absence of metal (**1a**). In addition, **1b** samples more conformational space along these PCs than **1a**, which is consistent with the increased level of conformational sampling seen for **1b** (see Table 3).

To highlight the differences in flexibility between **1a** and **1b,** an rmsf difference plot comparing the two is shown in Figure 6. Standard errors for these values are provided in SI-5 of the Supporting Information. For **1a**, several regions in the fingers domain exhibit increased flexibility, with limited flexibility in the palm domain and a rather rigid thumb domain. In particular, the Δ2 loop (residues 145−155) appears to be extremely flexible compared to the rest of the enzyme. In contrast, for **1b** we observe moderate flexibility in the fingers domain, a rigid palm domain, and an extremely flexible thumb domain. In comparison to that of **1a**, the Δ2 loop of **1b** is less flexible; this loop (also known as motif F) is thought to be involved in nucleotide triphosphate (NTP) and template binding during replication.[37] Therefore, flexibility in this region may be necessary for the Δ2 loop to function. However, the extreme flexibility seen in **1a** could be detrimental, preventing the loop from being properly positioned for binding to NTPs and template nucleotides. The Δ1 loop (residues 25−35), thought to regulate the opening and closing of the enzyme, is less flexible in **1a** than in **1b**. Flexibility in this region would be necessary for the enzyme to switch between open and closed conformations, as observed in **1b**. Thus, the decreased flexibility of these residues observed in **1a** may indicate this system is less able to convert between open and closed conformations. The

dx.doi.org/10.1021/bi400251g | *Biochemistry* 2013, 52, 4541−4552

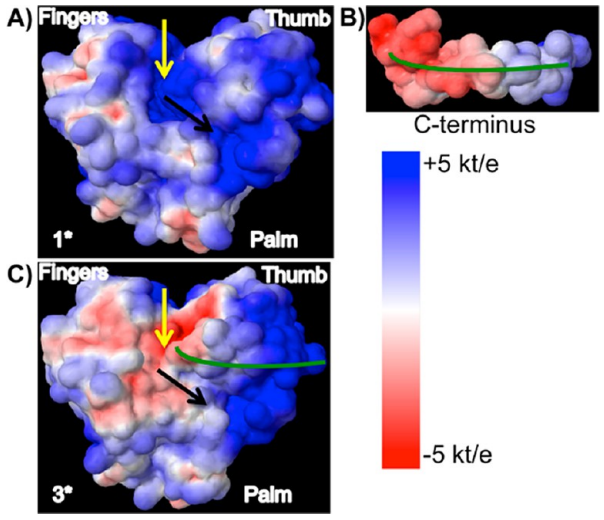


Figure 4. Electrostatic potentials for the front surface of the enzyme for **1*** and **3***, which are representative of the enzyme without and with C-terminal residues, respectively. (A) Without the C-terminus, the front surface of the enzyme displays mostly positive charges and thus a positive electrostatic potential. It is likely that the resulting potential stabilizes the presence of negative charges such as those occurring in the phosphate groups of RNA or incoming nucleotides. (B) The C-terminal residues display a predominantly negative potential, particularly for those residues that are in the center of the cleft. (C) When the C-terminus is present as in **3***, it creates a more negative potential near the RNA template and duplex channels, generating an environment that is less favorable to the presence of negative charge. The template and duplex channels are highlighted by yellow and black arrows, respectively, while the green line is a schematic trace of the C-terminal backbone. The electrostatic potentials were calculated using the PDB2PQR Server[43,44] and the APBS[45] software; default settings were employed in both cases.

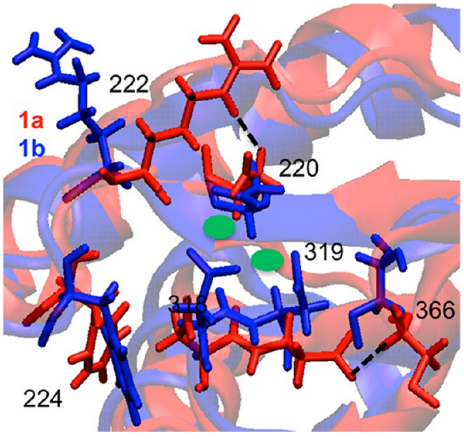


Figure 5. Changes in side chain conformations upon metal binding in the palm domain for **1a** (red) and **1b** (blue). When metal is bound (**1b**), the catalytic triad (Asp220, -318, and -319) coordinate the Mg^{2+} ions, which prevents these residues from interacting with others around them. In the absence of the metal ions (**1a**), Asp319 flips to hydrogen bond with Cys366 while Asp220 and Arg222 rotate to hydrogen bond with each other. These interactions may prevent functionally relevant conformational sampling. Metal ions (green spheres) have been placed where they would appear in the metal-bound structure. Hydrogen bonds are indicated by dashed lines.

rmsf difference plots for systems **2**–**4** are provided in SI-6 of the Supporting Information.

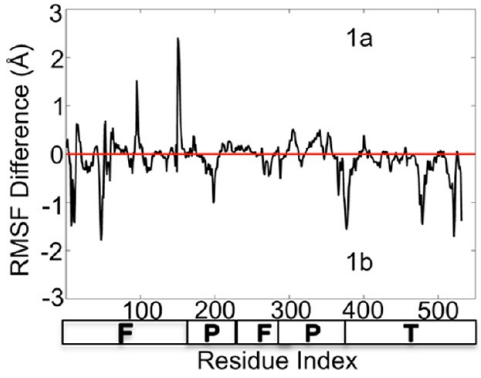


Figure 6. rmsf difference plots for **1a** and fully closed **1b**. The composition of the enzyme domains is shown in the bars underneath the plot (F for fingers, P for palm, and T for thumb). Peaks above the red line (positive) are regions of the enzyme that are more flexible in **1a**, while peaks below the red line (negative) are regions of the enzyme that are more flexible in **1b**. Overall, enzyme flexibility increases with metal binding. Some regions of the fingers domain are more flexible in the absence of metal binding (**1a**), but the motion of these regions differs substantially from that occurring in **1b** (see the text).

Altered Correlated Motions. At first glance, the correlated motions in **1a** and **1b** appear to be very similar, but there is, in fact, a distinct change in correlation upon metal binding (Figure 3). The motions for **1b** were discussed previously. **1a** experiences anticorrelated motions between residues spread over the fingers and palm domains (residues 50–100, 145–170, and 220–360) and the thumb domain (residues 390–520). Visualization of the lowest-frequency PC reveals that the fingers and thumb domains are not moving toward and away from each other as in **1b** but are instead moving up and down in an anticorrelated fashion. This is the dominant enzyme motion seen for **1b**. Overall, the absence of the metal ions decreases the extent of domain communication [with fewer hydrogen bonds between the fingers and thumb domains (see the next section)] and alters the directions of motion, which may weaken the ability of the enzyme to convert between open and closed conformations.

Occurrence of a Fully Closed Conformation. The rmsd plots for **1** (Figure 7) and other analyses suggest major structural changes occur during the course of the simulations.

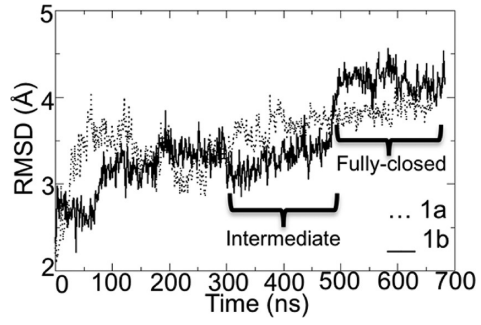


Figure 7. rmsd plots for **1a** and **1b**. For **1a**, the last 200 ns of equilibrated data was used for analysis. For **1b**, 200 ns (residues 291–490) was used to characterize the intermediate structure and 190 ns (residues 501–690) to characterize the fully closed conformation. Snapshots 491–500 were omitted from analysis as this trajectory segment includes the transition from the intermediate to the fully closed conformation.

dx.doi.org/10.1021/bi400251g | *Biochemistry* 2013, 52, 4541–4552

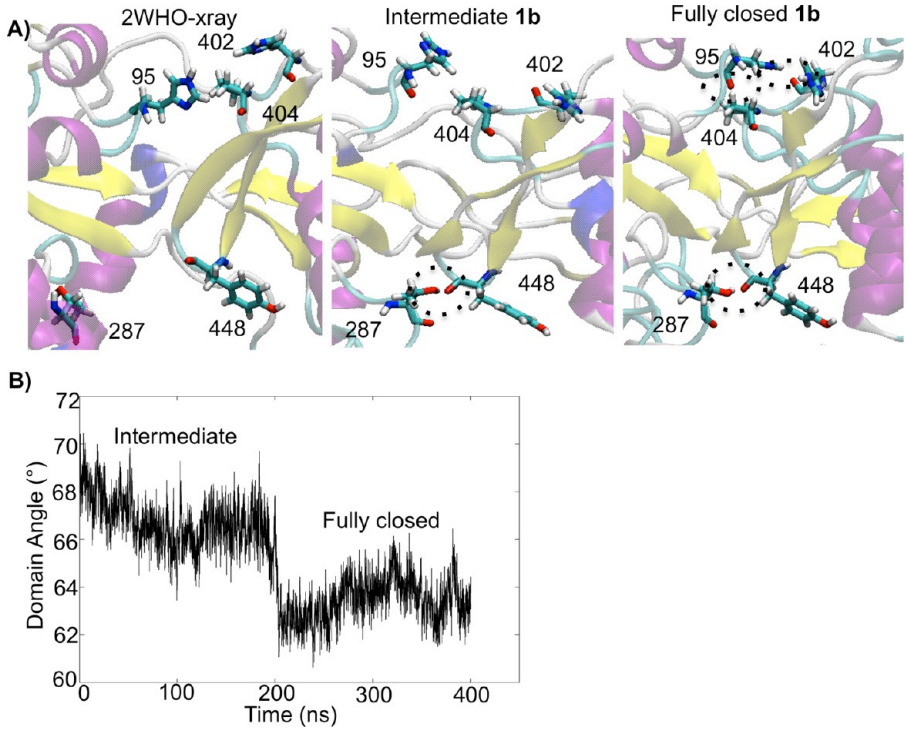


Figure 8. (A) Hydrogen bonds between the fingers and thumb domains for 2WHO (X-ray), intermediate **1b**, and fully closed **1b**. As the enzyme becomes more closed, more hydrogen bonds form between the fingers and thumb domains. Trajectory segments of 290−490 and 500−690 ns were used to calculate the average structures in this figure for intermediate **1b** and fully closed **1b**, respectively. (B) Domain angles for both the intermediate and fully closed structures of **1b**. This plot demonstrates the abrupt transition from the intermediate to fully closed conformation as the fingers and thumb domains come closer together.

We analyzed the final 200 ns for **1a** and the 200 ns equilibrated section labeled "intermediate" and the 190 ns equilibrated section labeled "fully closed" for **1b** to determine the nature of these changes. The average structures for these trajectory segments reveal that **1b** makes a conformational change from the "closed" conformation seen in **1*** to a conformation we term fully closed. In contrast, **1a** does not display this conformational change.

Figure 8A shows magnified views of **1***, intermediate **1b**, and fully closed **1b** at the intersection of the fingers and thumb domains. The residues shown as licorice begin to interact as the enzyme undergoes the transition to the fully closed conformation. Two hydrogen bonds are necessary to stabilize the fully closed conformation: (i) between $H^{\gamma 1}$ of T287 and the carbonyl O of Y448 (thumb β-loop) and (ii) between the carbonyl O of H402 and $H^{\varepsilon 1}$ of H95. The "g_hbond" utility in GROMACS indicates that these hydrogen bonds possess lifetimes of 774 and 583 ps, respectively. The first is located near the lower region of intersection between the fingers and thumb domains, while the other is found at the top of the fingers and thumb domains. There is also a weak electrostatic interaction between the carbonyl O of H95 and $H^{\beta 1}$ of P404. A plot of the hydrogen−acceptor distances and donor−hydrogen−acceptor angles for this interaction is provided in the Supporting Information (SI-7). **1*** has none of these interactions, as none of the distances between these residues are short enough for hydrogen bonding to occur. Intermediate **1b** displays the first hydrogen bond, with the other residues approaching each other more closely compared to those in the X-ray structure. Finally, fully closed **1b** has both hydrogen bonding interactions and additional weak electrostatic interactions. Figure 8B is a plot of the domain angle over time for the intermediate and fully closed structures of **1b**. The plot shows that the domain angle for the intermediate conformation fluctuates around 67°, while the domain angle for the fully closed conformation, which decreases abruptly at 200 ns, fluctuates around 63°. In contrast, **1*** has a domain angle of 71°.

To improve our understanding of which regions of the protein are involved in the transition, we performed covariance analysis on the 400 ns of data for **1b** that included the intermediate and fully closed structures (SI-8 of the Supporting Information), using the transition structure (average structure from 490 to 500 ns) as the reference structure. Regions displaying intense "correlation" as a result of this analysis reveal which parts of the enzyme move the most as a consequence of the structural transition. The correlation map reveals significant changes involving the relative position of the Δ1 loop with respect to the fingers and thumb domains, as well as an altered relative conformation of the fingers and thumb domains. These structural changes correspond to the enzyme conformation becoming more closed as the fingers and thumb domains approach each other and suggest an important role for the Δ1 loop in this process.

Structure **1a** does not experience this conformational change and displays none of the aforementioned interactions. However, the average structure from the last 200 ns of **1a** indicates that the fingers and thumb domains have moved closer together compared to those in **1***, suggesting that the crystallographic coordinates are inherently more open than the enzyme structure in solution.

The closed and open conformations of NS5B are associated with efficient *de novo* initiation and the exit of double-stranded RNA from the duplex channel during elongation, respectively.[20]

dx.doi.org/10.1021/bi400251g | *Biochemistry* 2013, 52, 4541−4552

As previously mentioned, crystallographic coordinates displaying an open conformation have not been reported for HCV genotype **1b**. However, we believe that **1*** is in a conformation more suitable for elongation (i.e., it is more open), while our fully closed structure is more conducive to *de novo* initiation. To test this hypothesis, we probed template and duplex channel widths using the distance between Cα atoms of residues on either side of each channel. For the template channel, we used residues 139 and 405, while for the duplex channel, we used residues 96 and 406. In our previous study, we used residues 14 and 96 to probe the template channel;[24] however, we found that those distances are not representative of the template channel width in these studies because these residues do not remain in the interior of the enzyme as it undergoes a change in conformation.

It is thought that a narrower template channel is necessary for NS5B to identify and interact with the template for replication to begin.[20] Once initiation has taken place, the template and duplex channels become wider to allow the nascent RNA to exit the enzyme. The template channel widths are 20.50, 17.43, and 12.20 Å for **1***, **1a**, and **1b**, respectively, while the corresponding duplex channel widths are 17.40, 13.16, and 10.33 Å, respectively. In comparison, the open conformation of NS5B from HCV genotype **2a** (PDB entry 1YV2) has a template channel width of 20.10 Å and a duplex channel width of 17.30 Å. The agreement between values for 1YV2 and **1*** indicates that **1*** is more structurally similar to an open conformation than the closed conformation commonly suggested. This observation indicates that **1*** can readily accommodate double-stranded RNA but is not in an initiation-ready conformation, as the template channel is too wide.

To more fully examine the generality of the observed correlation between duplex channel width and accommodation of duplex RNA, we also measured the duplex channel of the closely related 3D-pol from poliovirus. Because of the difference in sequence between NS5B and 3D-pol, it was necessary to perform a structural alignment of the two proteins to identify residues 113 and 412 corresponding to NS5B residues 96 and 406, respectively. The alignment was performed between **1*** and the 3D-pol with PDB entry 3OL6[38] (which contains duplex RNA exiting the enzyme) using the "multiseq" utility in VMD.[39] The width of the channel in 3OL6 was found to be 18.6 Å, consistent with that observed in the open conformation present in 1YV2 and **1***. This is additional evidence that the structure of **1*** is consistent with an open conformation.

The C-Terminus Prevents NS5B from Achieving an Initiation-Ready Conformation. As previously discussed, we believe that **1b** is most suitable for *de novo* initiation, as it exhibits a fully closed conformation. Widths of the template and duplex channels indicate that the C-terminus prevents the enzyme from becoming sufficiently closed to initiate replication, as template channel widths for **2b**–**4b** are significantly wider than those seen for **1b**. Figure 9A shows a plot of the template and duplex channel widths for average structures from each protein system as well as for the original crystal structures (**1***–**4***) and X-ray structures of the open and closed conformations of NS5B from genotype **2a** (PDB entries 1YV2 and 1YUY, respectively). The original PDB structures are clustered together and thus share similar conformations. Performing simulations induces the structures to become more closed as indicated by narrower template and duplex channels. However, closure is hindered by the presence of the C-terminal residues lodged between the fingers and thumb domains. In contrast, without interference from the C-terminus, **1b** is able to achieve a fully closed conformation with both channels becoming narrow.

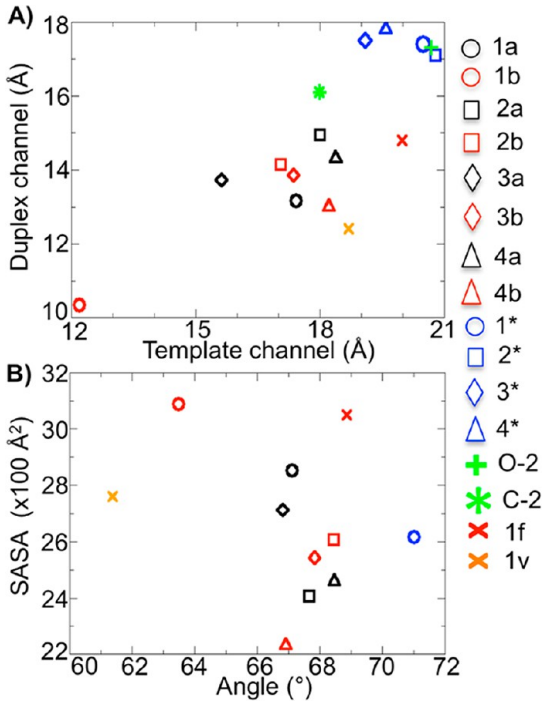


Figure 9. Important structural features vary among the simulations. (A) Plot of RNA template (*x*-axis) vs duplex channel (*y*-axis) widths. Residues 139 and 405 were used to probe the template channel, and residues 96 and 406 were used to probe the duplex channel. X-ray structures of the NS5B from HCV genotype **2a** with PDB entries 1YV2 and 1YUY are represented by O-2 and C-2, respectively; these are thought to epitomize the open and closed conformations, respectively. **1b** has the smallest template and duplex channels, while **1*** and O-2 have the largest channels. Structures containing C-terminal residues have channel widths between those of **1b** and O-2. This observation demonstrates that the C-terminus prevents the enzyme from becoming fully closed. The ligand-bound structure **1v** from our prior studies also does not adopt narrow channel widths. Our **1b** structure is more closed than C-2, which we expect to be more similar to the structures with C-termini as it contains 563 residues. (B) Plot of the angle among the fingers, palm, and thumb domains (*x*-axis) and the SASA of residues lining the template channel (*y*-axis). This plot demonstrates that **1b** is more closed than the other simulated structures in this study, as it has a quite small angle. However, we note that **1v** from our previous studies exhibits an even smaller angle (indicating it is more closed) as well as a smaller SASA value (indicating its template channel is less accessible to RNA). The fact that **1v** displays a decreased SASA despite exhibiting a small angle makes it markedly dissimilar from any of the other enzyme systems and supports our previous assertion that it is not on the pathway of functional conformational transitions.

Figure 9B shows a plot comparing the angle formed from the centers of mass of the fingers, palm, and thumb domains and the SASA for residues lining the template channel. The domain angle represents the degree of closure; i.e., smaller angles indicate a more closed enzyme. **1*** has the largest domain angle (71°), while **1b** has the smallest angle (63°). This change in angle indicates the conformational change from a more open conformation to the fully closed conformation. The structures containing C-terminal residues have values ranging from 66° to 68°, which resemble the domain angle for intermediate **1b**. This

dx.doi.org/10.1021/bi400251g | *Biochemistry* 2013, 52, 4541–4552

plot also shows **1b** has the largest SASA even though it has the smallest angle between domains. This increase in SASA occurs because residues lining the template channel become exposed to the center of the enzyme and increase their accessibility to solvent. This could be a consequence of these residues getting into place for *de novo* initiation in the closed conformation. In addition, **1v** from our previous study, which we described as hyper-closed, has the smallest domain angle but only a moderate SASA value. This hyper-closed conformation may not be amenable to *de novo* initiation, as the increased SASA seen in **1b** may represent an increased capacity for the enzyme to receive and interact with an incoming template. This observation is consistent with the fact that **1v** is derived from a simulation of NS5B bound to an allosteric inhibitor in the thumb domain and thus likely represents an inactive enzyme state. Overall, Figure 9 demonstrates that the original X-ray structures are more open than **1b**, which we believe represents the most functionally relevant state of the enzyme.

DISCUSSION

The C-Terminus Diminishes Enzyme Flexibility and Reduces the Number of Accessible Enzyme Conformations. We have shown that the binding of metal ions, as well as the presence of the enzyme's C-terminus, has a global influence on the structure and dynamics of NS5B. The direct interactions between the C-terminus and the fingers and thumb domains restrict the motion of those domains. This diminishes enzyme flexibility and disrupts communication between the two domains, preventing the enzyme from sampling conformations that are likely to be functionally relevant. We observe that the presence of the C-terminus prevents the enzyme from achieving a fully closed conformation even when metal ions are bound. In addition, the interactions between the C-terminus and the β-loop in the thumb domain (residues 450–455) may prevent this loop from conducting its proposed function of guiding the template to the active site.[36]

We note that the C-terminus of NS5B may adopt a different conformation *in vivo*, where it is expected to be membrane-associated and thus unable to associate with the remainder of the enzyme. Consequently, the conformation of the C-terminus described in these studies is likely to be most pertinent to *in vitro* systems, such as those often used to study the structure and RNA polymerase activity of NS5B. Other conformations of the C-terminus *in vitro* seem unlikely: this region of the enzyme is highly hydrophobic and would not be anticipated to be solvent-exposed in solution. Consequently, it is more probable that the C-terminus would be found in a conformation in which it is sequestered from solvent, consistent with its location in the cleft between the fingers and thumb domains in the available crystal structures. Thus, a truncated enzyme such as **1** is our best model for reproducing the impact of the C-terminus being membrane-associated as it is expected to be *in vivo*.

Experimental studies have shown that the C-terminal residues inhibit NS5B activity, though the specific mechanism by which this occurs is unclear.[13,14] Our simulations are consistent with these data and illustrate possible molecular mechanisms of this inhibitory effect. Moreover, the different enzyme sequences studied in this work, such as 2HAI, illustrate that these observations are robust with respect to sequence variation. This consideration is important given the strong propensity of the HCV genome to mutate because of the error-prone nature of NS5B. Overall, our studies suggest that C-terminal residues are not present between the fingers and thumb domains during replication *in vivo* because they prevent NS5B from achieving a fully closed conformation that would facilitate the initiation of RNA replication.[13−15] Because of the dramatic impact of the C-terminus on enzyme properties, we suggest that those seeking to conduct simulation studies of this enzyme remove C-terminal residues to allow the protein to explore more conformational space and exhibit dynamics that is more likely to be functionally relevant (see below).

Metal Ion Binding Induces Allosteric Effects That Lead to a Newly Observed Fully Closed Conformation. Divalent metal ions such as magnesium are directly involved in catalysis, and previous studies have shown that metal ion binding creates a more structurally stable enzyme.[16,17] In this study, we are able to determine how this occurs at the molecular level. In addition, we show that the metal ions do more than provide stability: they also impact the global structure and dynamics of the enzyme and thus induce long-range, allosteric effects. Overall, metal ion binding increases the flexibility of the enzyme, allowing the enzyme to sample additional conformations, including a newly observed fully closed conformation. We believe this fully closed conformation facilitates *de novo* initiation for two reasons. First, the template channel is narrower, facilitating the interactions required for initiation. Second, there are more residues exposed to the surface of the template channel, putting them in position to receive and guide the template to the active site.

Relationship to Other Studies. In previous studies, we identified a ligand-bound NS5B (**1v**) as having a hyper-closed conformation compared to that of its ligand-free form.[24] However, the fully closed conformation displayed by **1b** in the study presented here differs from the previously observed hyper-closed conformation. The ligand-bound **1v** structure was previously shown to occupy a distinct conformational space from the ligand-free (**1f**) structure. We compare **1v** and **1f** to average structures from our current simulations in Figures 2 and 8 using a number of different metrics. The data show that **1v** from our previous study is unlike any of the PDB structures or our currently simulated systems, supporting our previous suggestion that this ligand-bound structure displays an inactive conformation that is not on the pathway of functional conformational transitions. This observation is consistent with the induced fit model of allosteric inhibition we previously suggested.[24] Such a model would indicate that ligand binding induces a protein conformation that differs from preexisting conformations of the unbound protein.[40,41] We note that the current simulations are much longer and thus are likely to have sampled more of the underlying conformational space than our previous studies. We also note that the prior studies were conducted using the CHARMM 22 force field and CHARMM simulation program, which differs from the OPLS-AA force field and GROMACS simulation program used in the work presented here. While these differences may affect the details of our observations, the fact that we see similar trends in both studies (e.g., similar patterns of correlated motion in **1b**) suggests that our overall conclusions are robust.

Our simulations also share similarities with computational studies performed by Moustafa et al.[42] They performed MD simulations of several picornaviral RNA polymerases (poliovirus, coxsacki virus, foot and mouth disease virus, and bovine viral diarrhea virus), as well as a G64S mutant of the polioviral polymerase. NS5B is structurally and functionally related to these enzymes. Although picornaviral RNA polymerases have ~100 fewer residues than NS5B and lack a C-terminus, they

share the right-handed organization common to many viral polymerases, including the fingers, palm, and thumb domains. Overall, Moustafa et al. observed patterns of correlated motions that appear to be conserved across the picornavirus family and are similar in general to those we observe for metal-bound NS5B, except that we observe more intense correlated motions. Via structural alignment, we determined that the general patterns of flexibility and anticorrelated motions observed between the fingers and thumb domains of NS5B by Moustafa et al. are analogous to those observed in our previous study and are also similar to those seen in **1b**. These observations indicate that RNA-dependent RNA polymerases from the different viruses may experience similar equilibrium motions, suggesting that their dynamics are conserved.[24,42] Thus, these enzymes may function via similar mechanisms and thus be functionally regulated in similar ways. Therefore, the information gained in this work may be useful in understanding the functional properties of a range of viral polymerases. Ultimately, this knowledge may reveal how best to inhibit the function of such polymerases in a therapeutic context.

In addition, our results support the findings of several experimental studies. In particular, a study by Chinnaswamy et al. observed open and closed conformations of NS5B using negative-stain electron microscopy and single-particle reconstruction. Their results suggest that the Δ1 loop regulates the transition from open to closed.[23] We observe the Δ1 loop plays an important role in the transition to the fully closed conformation in the work presented here. Our observations indicate that the Δ1 loop is strongly associated with the thumb domain and suggest that the loop plays a role in regulating the thumb domain's motions. Such a role would be required for it to regulate the transition between open and closed conformations. While PDB structures such as **1***–**4*** are widely believed to be in the closed conformation, their similarities to structures of viral polymerases known to be in the open conformation (such as 1YV2 and 3OL6) suggest they are more appropriately classified as open. Thus, such structures may not be directly applicable to elucidating the process by which RNA replication is initiated, as this process requires a closed enzyme. It may be necessary to first ensure that the closed conformation of the enzyme is attained (e.g., via MD simulations initiated from these structures).

Finally, we compared the structure of the fully closed conformation seen in this work with that of the related RNA polymerase from bacteriophage φ6 (PDB entry 1HI0). This structure represents an initiation complex of the enzyme and is similar to our fully closed conformation in several respects. We identified template and duplex channel residue probes through a structural alignment with NS5B and calculated widths of 14.2 and 13.0 Å for the template and duplex channels, respectively. In addition, we calculated a domain angle of 65.4°. These values are very similar to the corresponding quantities for the fully closed conformation attained by **1b** (see Figure 9A), supporting our assertion that this conformation of NS5B is suitable for the initiation of RNA replication.

■ CONCLUSION

This study reveals that the presence of magnesium ions and C-terminal residues has a significant impact on the dynamics and structure of NS5B. Magnesium ions induce specific structural and dynamic changes that are consistent with an enzyme that is more competent in initiating RNA replication. In contrast, the C-terminus of the enzyme restricts protein dynamics and prevents the sampling of conformations that are likely to facilitate the initiation of replication. This is the first computational study of apo and metal-bound NS5B structures at these time scales and a first step toward a better understanding of the intrinsic functional dynamics of NS5B and other RNA polymerases. This knowledge may ultimately aid efforts to identify inhibitors of these enzymes that can serve as the basis for antiviral therapies.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Supplemental tables and figures. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: ithorpe@umbc.edu. Phone: (410) 455-5728.

Funding

This work was supported by the Department of Chemistry and Biochemistry of the University of Maryland, Baltimore County (UMBC). This work made use of the UMBC High Performance Computing Facility that is supported by the National Science Foundation through the MRI program (Grants CNS-0821258 and CNS-1228778) and the SCREMS program (Grant DMS-0821311), with additional substantial support from UMBC. This study also employed resources from the Extreme Science and Engineering Discovery Environment (XSEDE) that are supported by National Science Foundation Grant OCI-1053575.

Notes

The authors declare no competing financial interest.

■ REFERENCES

(1) Lindenbach, B., and Rice, C. (2005) Unraveling hepatitis C virus replication from genome to function. *Nature 436*, 933–938.

(2) World Health Organization. Hepatitis C Fact Sheet No. 164. http://www.who.int/mediacentre/factsheets/fs164/en/ (revised June 2011).

(3) McHutchison, J., Bartenschlager, R., Patel, K., and Pawlotsky, J. M. (2006) The face of future hepatitis C antiviral drug development: Recent biological and virologic advances and their translation to drug development and clinical practice. *J. Hepatol. 44*, 411–421.

(4) Neuman, A. U., Lam, N. P., Dahari, H., Gretch, D. R., Wiley, T. E., Layden, T. J., and Perelson, A. S. (1998) Hepatitis C Viral Dynamics in Vivo and the Antiviral Efficacy of Interferon-α Therapy. *Science 282*, 103–107.

(5) McKercher, G., Beaulieu, P. L., Lamarre, D., LaPlante, S., Lefebvre, S., Pellerin, C., Thauvette, L., and Kukolj, G. (2004) Specific inhibitors of HCV polymerase identified using an NS5B with lower affinity for template/primer substrate. *Nucleic Acids Res. 32*, 422–431.

(6) Walker, M. P., and Hong, Z. (2002) HCV RNA-dependent RNA polymerase as a target for antiviral development. *Curr. Opin. Pharmacol. 2*, 1–7.

(7) Ortin, J., and Parra, F. (2006) Structure and Function of RNA Replication. *Annu. Rev. Microbiol. 60*, 305–326.

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

(9) Ago, H., Adachi, T., Yoshida, A., Yamamoto, M., Habuka, N., Yatsunami, K., and Miyano, M. (1999) Crystal structure of the RNA-dependent RNA polymerase of hepatitis C virus. *Structure 7*, 1417–1426.

(10) Bressanelli, S., Tomei, L., Roussel, A., Incitti, I., Vitale, R. L., Mathieu, M., De Francesco, R., and Rey, F. (1999) Crystal structure of

the RNA-dependent RNA polymerase of hepatitis C virus. *Proc. Natl. Acad. Sci. U.S.A. 96*, 13034−13039.
(11) Behrens, S., Tomei, L., and De Francesco, R. (1996) Identification and properties of the RNA-dependent RNA polymerase of hepatitis C virus. *EMBO J. 15*, 12−22.
(12) Lescar, J., and Canard, B. (2009) RNA-dependent RNA polymerases from flaviviruses and Picornaviridae. *Curr. Opin. Struct. Biol. 19*, 759−767.
(13) Harrus, D., Ahmed-El-Sayed, N., Simister, P. C., Miller, S., Triconnet, M., Hagedorn, C. H., Mahlas, K., Rey, F., Astier-Gin, T., and Bressanelli, S. (2010) Further insights into the Roles of GTP and the C Terminus of the Hepatitis C Virus Polymerase in the Initiation of RNA Synthesis. *J. Biol. Chem. 285*, 32906−32918.
(14) Adachi, T., Ago, H., Habuka, N., Okuda, K., Komatsu, M., Ikeda, S., and Yatsunami, K. (2002) The essential role of C-terminal residues in regulating the activity of hepatitis C virus RNA-dependent RNA polymerase. *Biochim. Biophys. Acta 1601*, 38−48.
(15) Leveque, V., Johnson, R. B., Parsons, S., Ren, J., Xie, C., Zhang, F., and Wang, Q. M. (2003) Identification of a C-Terminal Regulatory Motif in Hepatitis C Virus RNA-Dependent RNA Polymerase: Structural and Biochemical Analysis. *J. Virol. 77*, 9020−9028.
(16) Benzaghou, I., Bougie, I., and Bisaillon, M. (2004) Effect of Metal Ion Binding on the Structural Stability of the Hepatitis C Virus RNA Polymerase. *J. Biol. Chem. 279*, 49755−49761.
(17) Bougie, I., Charpentier, S., and Bisaillon, M. (2003) Characterization of the Metal Ion Binding Properties of the Hepatitis C Virus RNA Polymerase. *J. Biol. Chem. 278*, 3868−3875.
(18) Bougie, I., and Bisaillon, M. (2009) Metal ion-binding studies highlight important differences between flaviviral RNA polymerases. *Biochim. Biophys. Acta 1794*, 50−60.
(19) Ng, K. K.-S., Arnold, J. J., and Cameron, C. E. (2008) Structure-Function Relationships Among RNA-Dependent RNA Polymerases. *Curr. Top. Microbiol. Immunol. 320*, 137−156.
(20) Chinnaswamy, S., Yarbrough, I., Palaninathan, S., Kumar, C. T., Vijayaraghavan, V., Demeler, B., Lemon, S. M., Sacchettini, J. C., and Kao, C. C. (2008) A locking mechanism regulates RNA synthesis and host protein interaction by the hepatitis C virus polymerase. *J. Biol. Chem. 283*, 20535−20546.
(21) van Dijk, A. A., Makeyev, E. V., and Bamfor, D. H. (2004) Initiation of viral RNA-dependent RNA polymerization. *J. Gen. Virol. 85*, 1077−1093.
(22) Biswal, B. K., Cherney, M. M., Wang, M., Chan, L., Yannopoulos, C. G., Bilimoria, D., Nicolas, O., Bedard, J., and James, M. N. G. (2005) Crystal Structures of the RNA-dependent RNA Polymerase Genotype 2a of Hepatitis C Virus Reveal Two Conformations and Suggest Mechanisms of Inhibition by Non-nucleoside Inhibitors. *J. Biol. Chem. 280*, 18202−18210.
(23) Chinnaswamy, S., Murali, A., Cai, H., Yi, G., Palaninathan, S., and Cheng Kao, C. (2010) Conformations of the monomeric hepatitis C virus RNA-dependent RNA polymerase. *Virus Adapt. Treat. 2*, 21−39.
(24) Davis, B. C., and Thorpe, I. F. (2012) Thumb inhibitor binding eliminates functionally important dynamics in the hepatitis c virus RNA polymerase. *Proteins: Struct., Funct., Bioinf. 81*, 40−52.
(25) van der Spoel, D., Lindahl, E., and Hess, B. (2006) *Gromacs User Manual*.
(26) Berendsen, H. J., Postma, J. P. M., van Gunsteren, W. F., and Hermans, J. (1981) Interaction models for water in relation to protein hydration. *Intermol. Forces*, 331−342.
(27) Jorgensen, W. L., Maxwell, D. S., and Tirado-Rives, J. (1996) Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *J. Am. Chem. Soc. 118*, 11225−11236.
(28) Ryckaert, J.-P., Ciccotti, G., and Berendsen, H. J. (1977) Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-Alkanes. *J. Comput. Phys. 23*, 327−341.
(29) Darden, T., York, D., and Pedersen, L. (1993) Particle mesh Ewald: An Nlog(N) method for Ewald sums in large systems. *J. Chem. Phys. 98*, 10089−10092.
(30) Lange, O. F., and Grubmuller, H. (2006) Can Principle Components Yield a Dimension Reduced Description of Protein Dynamics on Long Time Scales? *J. Phys. Chem. B 110*, 22842−22852.
(31) Kurylowicz, M., Yu, C., and Pomes, R. (2010) Systematic Study of Anharmonic Features in a Principle Component Analysis of Gramicidin A. *Biophys. J. 98*, 386−395.
(32) Ivosev, G., Burton, L., and Bonner, R. (2008) Dimensionality Reduction and Visualization in Principle Component Analysis. *Anal. Chem. 80*, 4933−4944.
(33) Ramanathan, A., and Agarwal, P. K. (2009) Computational Identification of Slow Conformational Fluctuations in Proteins. *J. Phys. Chem. B 113*, 16669−16680.
(34) O'Farrell, D., Trowbridge, R., Rowlands, D., and Jager, J. (2003) Substrate Complexes of Hepatitis C Virus RNA Polymerase (HC-J4): Structural Evidence for Nucleotide Import and De-novo Initiation. *J. Mol. Biol. 326*, 1025−1035.
(35) Lesburg, C. A., Cable, M. B., Ferrari, E., Hong, Z., Mannarino, A. F., and Weber, P. C. (1999) Crystal structure of the RNA-dependent RNA polymerase from hepatitis C virus reveals a fully encircled active site. *Nat. Struct. Biol. 6*, 937−943.
(36) Hong, Z., Cameron, C. E., Walker, M. P., Castro, C., Yao, N., Lau, J. Y. N., and Zhong, W. (2001) A Novel Mechanism to Ensure Terminal Initiation by Hepatitis C Virus NS5B Polymerase. *Virology 285*, 6−11.
(37) Hansen, J. L., Long, A. M., and Schultz, S. C. (1997) Structure of the RNA-dependent RNA polymerase of poliovirus. *Structure 5*, 1109−1122.
(38) Kortus, M. G., Kempf, B. J., Haworth, K. G., Barton, D. J., and Peersen, O. B. (2012) A template RNA entry channel in the fingers domain of the poliovirus polymerase. *J. Mol. Biol. 417*, 263−278.
(39) Roberts, E., Eargle, J., Wright, D., and Luthey-Schulten, Z. (2006) MultiSeq: Unifying sequence and structure data for evolutionary analysis. *BMC Bioinf. 7*, 382−393.
(40) Weikl, T. R., and Von Deuster, C. (2009) Selected-fit versus induced-fit protein binding: Kinetic differences and mutational analysis. *Proteins 75*, 104−110.
(41) Kar, G., Keskin, O., Gursoy, A., and Nussinov, R. (2010) Allostery and population shift in drug discovery. *Curr. Opin. Pharmacol. 10*, 715−722.
(42) Moustafa, I. M., Shen, H., Morton, B., Colina, C., and Cameron, C. E. (2011) Molecular Dynamics Simulations of Viral RNA Polymerases Link Conserved and Correlated Motions of Functional Elements of Fidelity. *J. Mol. Biol. 410*, 159−181.
(43) Dolinsky, T. J., Czodrowski, P., Li, H., Nielsen, J. E., Klebe, G., and Baker, N. A. (2007) PDB2PQR: Expanding and upgrading automated preperation of biomolecular structures for molecular simulations. *Nucleic Acids Res. 35*, W522−W525.
(44) Dolinsky, T. J., Nielsen, J. E., McCammon, J. A., and Baker, N. A. (2004) PDB2PQR: An automated pipeline for the setup, execution, and analysis of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Res. 32*, W665−W667.
(45) Baker, N. A., Joseph, S., Holst, M. J., and McCammon, J. A. (2001) Electrostatics of nanosystems: Application to microtubules and the ribosomes. *Proc. Natl. Acad. Sci. U.S.A. 98*, 10037−10041.