

# Toward the Mechanism of Dynamical Couplings and Translocation in Hepatitis C Virus NS3 Helicase Using Elastic Network Model

Wenjun Zheng,<sup>1\*</sup> Jung-Chi Liao,<sup>2</sup> Bernard R. Brooks,<sup>1</sup> and Sebastian Doniach<sup>3</sup>

<sup>1</sup>Laboratory of Computational Biology, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, Maryland 20892

<sup>2</sup>Department of Bioengineering, Stanford University, Stanford, California 94305

<sup>3</sup>Department of Applied Physics, Stanford University, Stanford, California 94305

**ABSTRACT** Hepatitis C virus NS3 helicase is an enzyme that unwinds double-stranded polynucleotides in an ATP-dependent reaction. It provides a promising target for small molecule therapeutic agents against hepatitis C. Design of such drugs requires a thorough understanding of the dynamical nature of the mechanochemical functioning of the helicase. Despite recent progress, the detailed mechanism of the coupling between ATPase activity and helicase activity remains unclear. Based on an elastic network model (ENM), we apply two computational analysis tools to probe the dynamical mechanism underlying the allosteric coupling between ATP binding and polynucleotide binding in this enzyme. The correlation analysis identifies a network of hot-spot residues that dynamically couple the ATP-binding site and the polynucleotide-binding site. Several of these key residues have been found by mutational experiments as functionally important, while our analysis also reveals previously unexplored hot-spot residues that are potential targets for future mutational studies. The conformational changes between different crystal structures of NS3 helicase are found to be dominated by the lowest frequency mode solved from the ENM. This mode corresponds to a hinge motion of the highly flexible domain 2. This motion simultaneously modulates the opening/closing of the domains 1–2 cleft where ATP binds, and the domains 2–3 cleft where the polynucleotide binds. Additionally, a small twisting motion of domain 1, observed in both mode 1 and the computed ATP binding induced conformational change, fine-tunes the binding affinity of the domains 1–3 interface for the polynucleotide. The combination of these motions facilitates the translocation of a single-stranded polynucleotide in an inchworm-like manner. *Proteins* 2007;67:886–896. © 2007 Wiley-Liss, Inc.†

**Key words:** normal mode analysis; elastic network model; dynamical coupling; helicase; allostery

## INTRODUCTION

Hepatitis C virus (HCV) is a major cause of chronic liver disease. About 3% of worldwide population, or 170 million people, are infected by HCV, and yet no effective treatment is available to date. Much research has focused on the replicative enzymes of HCV as possible targets for effective therapeutic agents.<sup>1</sup> Certain potent inhibitors are under development and test to block the HCV replication activity, including some ATP-binding inhibitors that have been successful drug targets for chronic myeloid leukemia and other serious diseases.<sup>2</sup>

Helicases are enzymes that can unwind double-stranded DNA or RNA in an ATP-dependent reaction. They are required for gene replication, transcription, translation, recombination, and repair.<sup>3–5</sup> In HCV, a non-structural protein NS3 acts as a multifunctional enzyme, where the N-terminal is a serine protease and the C-terminal has RNA helicase activities. The NS3 helicase complex assembles on a 3′ single-stranded DNA tail and moves forward in the 3′ to 5′ direction as it unwinds double-stranded polynucleotides.<sup>6</sup> It constitutes one possible drug target for HCV.<sup>1</sup> To facilitate the assessment of the action of small molecule drugs on the NS3 helicase, we have investigated the dynamics of the mechanochemical functioning of the helicase using an elastic network model (ENM).

The functional mechanism and the corresponding active form of the NS3 helicase remain unclear. It was proposed that NS3 forms an oligomer, which binds the nucleic acid as a string.<sup>7</sup> Further experiments supported this model and showed that multiple NS3 molecules are required for optimal unwinding.<sup>8</sup> Another study suggested the active unwinding form of NS3 is a dimer.<sup>9</sup> A

Grant sponsor: NHLBI, NIH; Grant sponsor: National Center for Biomedical Computing, NIH; Grant number: U54 GM072970.

Wenjun Zheng and Jung-Chi Liao contributed equally to this work.

\*Correspondence to: Wenjun Zheng, Laboratory of Computational Biology, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, MD 20892. E-mail: zhengwj@helix.nih.gov

Received 14 August 2006; Revised 30 October 2006; Accepted 9 November 2006

Published online 20 March 2007 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.21326

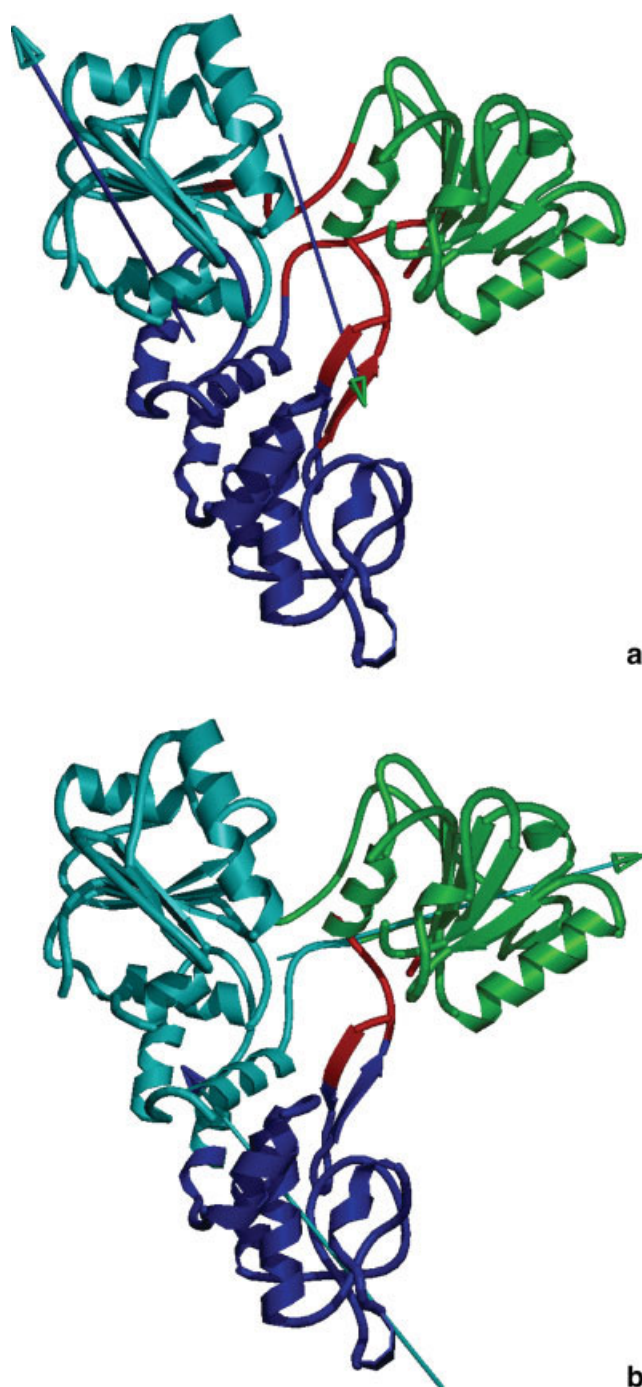


Fig. 1. Dynamical domain partitioning of the conformational changes described by the following normal modes: (a) Mode 1 describes two pairwise interdomain motions between three dynamical domains: domain 1 in cyan (residue range: 190–278, 280–325), domain 2 in green (residue range: 331–414, 418–428, 455–474), and domain 3 in blue (residue range: 435–444, 485–498, 500–620); (b) Mode 2 describes two pairwise interdomain motions between three dynamical domains: domain 1 in blue (residue range: 190–278, 280–325, 480–498, 500–535), domain 2 in cyan (residue range: 326–414, 418–428, 455–479), and domain 3 in green (residue range: 435–449, 536–620). The flexible bending regions are colored in red. The rotation axis for each pairwise interdomain motion is shown as an axis with an arrow head; the color of the axis stem is the same as the domain fixed for the structural alignment, whereas the color of the arrow head is the color of the moving domain.

recent single molecule experiment has revealed that an NS3 monomer is functionally active and the step size of NS3's cyclic movement has been resolved.<sup>10</sup> This result supported an inchworm mechanism for the mechanochemical coupling of this enzyme. Therefore, it is plausible that NS3 is able to unwind nucleic acids in its monomeric form, although they perform optimally as dimer or oligomer.

Several crystal structures of NS3 helicases have been solved.<sup>11–14</sup> In one NS3 structure,<sup>14</sup> the helicase domain is complexed with a single-stranded DNA oligonucleotide. It consists of three structural domains with the oligonucleotide lying in a groove between domains 1, 2 and domain 3 (Fig. 1). Domains 1 and 3 share an extensive interface, whereas domain 2 is flexibly linked to the other two. The NS3 helicase is a member of superfamily II (SF2) of helicases,<sup>15</sup> and the residues that are conserved among SF2 helicases (SF2 motifs) line an interdomain cleft between the domains 1 and 2.<sup>16,17</sup> For these SF2 helicases, motif I is the highly conserved P-loop (Walker A motif) responsible for the ATP binding and catalysis, motif II contains the DExx box (including Walker B motif) responsible for the  $Mg^{2+}$  coordination and positioning of the catalytic water, whereas motif VI contains the conserved Arginine finger from another domain (domain 2 in NS3 helicase) critical for catalysis. The oligonucleotide binds in an orthogonal binding site and contacts relatively few conserved residues. Motif III may be responsible for coordinating the binding of both ATP and the oligonucleotide.

Based on the above structure, a model was proposed by Kim et al. for the translocation of a single-stranded polynucleotide<sup>14</sup>: ATP binding induces a closure between domains 1 and 2, which results in a concomitant movement of the polynucleotide relative to the protein. Later, an inchworm mechanism was proposed for PcrA DNA helicase (it is structurally similar to the NS3 helicase) that supplemented the Kim model with an additional mechanism for duplex destabilization/unwinding that is coupled with the polynucleotide translocation process.<sup>18</sup> Alternative scenarios have been explored as well. For example, a Brownian motor mechanism was proposed where the individual monomer acts as a Brownian ratchet on a rugged energy surface.<sup>19</sup>

To evaluate the propensity of candidate small-molecule drugs to inhibit the helicase action, computational models are needed. However, direct all-atom molecular dynamics simulations are in practice limited in both time scale and system size. Recently, an ENM approach has been proposed to model the interatomic contacts as springs and analyze the protein as a network of coupled harmonic oscillators (see Methods). The normal mode analysis (NMA) of the ENM was shown by various studies<sup>20–22</sup> to give a handful of low-frequency normal modes that are capable of describing various protein conformational changes. This simple model captures the mechanical nature of many biomolecular nanomachines. Therefore the application of the ENM provides a powerful and efficient tool to model the large-scale motions during the functioning of an enzyme.

We have recently developed two ENM-based computational tools to analyze dynamical correlations/couplings within proteins. The first tool,<sup>23</sup> which we refer to as a “correlation analysis,” is based on the computation of pairwise correlation functions that probe significant correlations between spatially distant residues. In this analysis, we introduced a mutational perturbation at every residue and compute the response in a subset of residues that constitute a functional site. This analysis allows us to identify the dynamically important “hot-spot” residues whose perturbations may significantly affect the flexibility of a biologically important functional site. We have previously tested this analysis in a myosin motor by probing the couplings among the nucleotide-binding site, the actin-binding site, and the force-generating converter. A set of hot-spot residues were identified, which are located over several functionally important inter-subdomain joints of myosin motor: many of these residues were previously shown to be functionally relevant by mutational studies.<sup>23</sup>

Our second tool, which we refer to as “deformation analysis,”<sup>22,24</sup> is based on the computation of the global conformational change induced by a local structural deformation at the ATP-binding site in ATP-driven enzymes. This tool has been previously used to study the ATP-binding-induced conformational motions in kinesin and myosin motors.<sup>22,24</sup> This allowed us to predict the collective motions that support the power-stroke scenario for myosin but not for kinesin.<sup>22</sup> It is natural to apply this analysis to other proteins where ATP binding drives large conformational changes. In this paper we apply the deformation analysis to NS3 helicase whose function is also based on ATP binding and hydrolysis. In the cases of kinesin and myosin, crystallographic information was available for the structural deformation at the ATP-binding site during the mechanochemical cycle. However, this information is not available for NS3 helicase because the NS3 structure at the ATP-bound state has not been solved yet. Here we are able to exploit the very close sequence and structure homology in the ATP-binding residues between HCV NS3 helicase and bovine F1 ATPase. We can use this homology and the ATP-bound structure of F1 ATPase to model the putative deformation at the ATP-binding site during the mechanochemical cycle of NS3 helicase.

To validate the applicability of ENM to NS3 helicase, we first compare the ENM-derived normal modes with the crystallographically observed conformational changes between several NS3 helicase structures in Protein Data Bank, and we find modes 1 and 2 adequately account for the observed structural changes. Then we use the above two analysis tools to probe the dynamical mechanism for the couplings between the ATP binding and the polynucleotide binding/translocation, and the results are summarized as follows:

- First, we have identified a small set of dynamically important hot-spot residues, which comprise a sparse network that spans between the ATP-binding site and

the polynucleotide-binding site. Among them are several highly conserved residues of SF2 motifs implicated for helicase functions by other studies including mutational experiments.

- Second, we have found that ATP binding triggers a large closure motion of domain 2 and a small twisting motion of domain 1, which modulates the polynucleotide-binding affinity and drives the translocation of a single-stranded polynucleotide in a way that is consistent with the inchworm mechanism. Furthermore, this conformational change is well described by the normal modes 1 and 2 solved from the ENM, which are the same modes that capture the crystallographically observed conformational changes between several NS3 structures.

## MATERIALS AND METHODS

### Elastic Network Model

Given the  $C_\alpha$  coordinates for a protein’s native structure, we build an ENM by using harmonic potentials with a uniform force constant  $C$  to account for pair-wise contact interactions between all  $C_\alpha$  atoms that are within a cutoff distance ( $R_C = 10 \text{ \AA}$ ). The energy in the elastic network representation of a protein is

$$E_{\text{network}} = \frac{1}{2} \sum_{d_{ij}^0 < R_C} C (d_{ij} - d_{ij}^0)^2, \quad (1)$$

where  $d_{ij}$  is the distance between the  $C_\alpha$  atoms  $i$  and  $j$ , and  $d_{ij}^0$  is the distance between  $C_\alpha$  atoms  $i$  and  $j$  as given in the native structure.

We perform the NMA on the elastic energy in Eq. (1). The eigenvectors of the lowest frequency normal modes are then used to interpret the observed conformational changes between different experimental structures of a protein. The drastic simplification of representing a protein structure by an effective harmonic potential is justified in a study by Tirion,<sup>25</sup> who showed that the use of a uniform spring constant reproduces the slow dynamics computed from the NMA of all-atom potentials. Hinsen further simplified the ENM to the  $C_\alpha$ -only representation.<sup>26</sup>

### Fluctuation-Based Correlation Analysis

Following the formulation in our previous work,<sup>23</sup> we compute the overall mean-square fluctuation of the given subset of residues as follows:

$$F_S = \langle |P_S \bar{r}_S|^2 \rangle \propto \text{Tr}(P_S (H^{-1})_{SS} P_S) = \sum_{1 \leq m < M} \frac{|P_S \bar{v}_S^m|^2}{\omega_m}, \quad (2)$$

where  $\bar{r}_S$  is the structural displacement of the subset residues (or  $C_\alpha$  atoms, same below) away from the coordinates given in the native structure ( $S$  stands for subset);  $P_S$  is the projection matrix that projects out the six

translational and rotational components from  $\vec{r}_S$  (because we are only interested in the internal fluctuations of the subset residues);  $(H^{-1})_{SS}$  is the  $S$  submatrix of the inverse of the Hessian matrix  $H$  of the whole ENM;  $\text{Tr}()$  computes the trace of a matrix;  $\omega_m$  and  $\vec{v}^m$  are the eigenvalue and eigenvector of mode  $m$ , and  $\vec{v}_S^m$  is the  $S$  component of  $\vec{v}^m$ . A cutoff at mode  $M = 3N/10$  is used to compute  $H^{-1} = \sum_{1 \leq m < M} \frac{\vec{v}^m \vec{v}^m}{\omega_m}$  ( $N$  is the number of residues in the protein), and the six zero modes corresponding to rotations and translations are excluded (the nonzero modes start from 1).

In Eq. (2),  $\frac{|P_S \vec{v}_S^m|^2}{\omega_m}$  gives the contribution to  $F_S$  from mode  $m$ , which allows us to identify which modes contribute most to  $F_S$ .

Next we introduce a perturbation at a given position  $i$  and study how  $F_S$  changes in response.

Following our previous paper,<sup>23</sup> we introduce the following energetic perturbation at position  $i$  to mimic the effect of a point mutation (it perturbs by  $\delta C$  the spring constant of the springs that connect the  $C_\alpha$  atom  $i$  to its neighbors within  $R_C = 10 \text{ \AA}$ ):

$$\delta E_i = \delta C \sum_{k: d_{ik}^0 < R_c} \frac{(d_{ik} - d_{ik}^0)^2}{2}, \quad (3)$$

from which the matrix elements of the corresponding Hessian matrix  $\delta H_i$  can be calculated as follows:

$$\delta H_{i,ka,lb} = \frac{\partial^2 \delta E_i}{\partial x_{ka} \partial x_{lb}}, \quad (4)$$

where  $k, l$  are indices for residues and  $a, b = 0, 1, 2$  corresponding to  $x, y, z$  directions.

Then the change of  $F_S$  in response is:

$$\delta F_{S|i} \propto \text{Tr}[P_S \delta(H^{-1})_{SS} P_S] = -\text{Tr}[P_S (H^{-1} \delta H_i H^{-1})_{SS} P_S]. \quad (5)$$

Finally we define the following (normalized) correlation function between subset  $S$  and position  $i$ :

$$\text{CF}_i = \frac{\delta F_{S|i}}{\delta C F_S}. \quad (6)$$

$\text{CF}_i$  assesses the strength of dynamical coupling of residue  $i$  to the fluctuations of residues in the subset  $S$ . Therefore, we can identify the hot-spot residues that are strongly coupled with the residues in subset  $S$  as follows: sort all residues in order of declining  $\text{CF}_i$  and then keep the top 10% of them as the hot-spot residues.

Physically those hot-spot residues are dynamically important, because their contact interactions with neighboring residues may significantly affect the flexibility of the residues in subset  $S$ .

We note that for a given subset  $S$ , its set of hot-spot residues may overlap with  $S$  itself: namely, some hot-spot residues are part of  $S$ , which directly affect the flexibility of  $S$  via local interactions. More interestingly, the other hot-spot residues are spatially distant from  $S$ , which may hint for plausible long-range couplings that can transmit signals between distant sites.

**TABLE I. Alignment of Amino Acid Residues Between the HCV NS3 Helicase and the Bovine F1 ATPase at the ATP-Binding Site**

PDB	Range of residue numbers
1A1V	A207–212, A229–231, A290–293, A322,323, A467
1BMF	F159–164, F186–188, F256–259, F308,309, B373

## Deformation Analysis

The alignment of the ATP-binding residues between the HCV NS3 helicase (PDB: 1A1V) and the bovine F1 ATPase (PDB: 1BMF) is given in Table I.

We consider the following two conformations/states for the ATP-binding residues:

- T: ATP-bound state from bovine F1 ATPase (1BMF: chains B, F).
- F: ATP-free state from HCV NS3 (1A1V: chain A).

To structurally deform the ATP-binding site in the NS3 helicase from F state to T state, we superimpose (by translations and rotations) the T state conformation on top of the F state conformation to minimize the RMSD between them (Fig. 3). Then we extract the structural displacement vector  $\vec{r}_S$  for the subset of ATP-binding residues, and finally compute the induced structural displacement  $\vec{r}_E$  of the rest of the ENM in response to the given deformation  $\vec{r}_S$ <sup>22</sup>:

$$\vec{r}_E = -H_{EE}^{-1} H_{ES} \vec{r}_S, \quad (7)$$

where  $H_{EE}$  and  $H_{ES}$  are two submatrices of the ENM's Hessian matrix  $H = \begin{bmatrix} H_{SS} & H_{SE} \\ H_{ES} & H_{EE} \end{bmatrix}$ .

We can compute the overlap (generalized cosine<sup>22</sup>) between the induced conformational change and each mode. The modes with high overlap value contribute significantly to the induced structural displacement.

## Dynamical Domain Partition

To facilitate visualizing the motions described by a low-frequency mode, we want to decompose a given conformational change (from initial to end conformation) described by a mode into several rigid-body motions: we partition the whole protein structure into dynamical domains, which essentially move (rotate and translate) as rigid bodies.

We use the following procedure for domain partition:

1. We chop the protein into five-residue-long peptides, for each peptide  $i$ , we rotate and translate its initial conformation to superimpose on top of its end conformation to minimize the RMSD between them, and then we record the corresponding rotation matrix  $U_i$  and translation vector  $T_i$ .

- We compute a connectivity matrix  $M$  for all pairs of peptides  $(i, j)$ :  $M(i, j) = 1$  if the initial conformation of peptide  $j$  can be superimposed to its end conformation with  $\text{RMSD} < \text{RMSD}_{\text{cutoff}}$  by using rotation matrix  $U_i$  and translation vector  $T_i$ , and 0 otherwise.
- We generate the initial centroids by recursively removing the peptide  $i_m$  with the maximal connectivity  $C_i = \sum_j M(i, j)$ , together with all peptides  $j$  with  $M(i_m, j) = 1$ . The connectivities are updated after each removal.
- We cluster all  $(U_i, T_i)$  by using the K-means algorithm, where the initial centroids are from step 3.
- Finally we sort all clusters in decreasing size, and each cluster with at least 20 residues (or four peptides) corresponds to a dynamical domain.

$\text{RMSD}_{\text{cutoff}}$  is set by default to 0.5 times the RMSD of the conformational changes. By choosing smaller  $\text{RMSD}_{\text{cutoff}}$ , we can achieve more refined clustering.

The above algorithm is very similar to that used in DynDom developed by Hayward and Berendsen.<sup>27</sup> The main difference is that we only use  $C_\alpha$  coordinates, whereas DynDom uses other backbone atoms as well. This is necessary since the ENM-derived eigenvectors only give the displacements of  $C_\alpha$  atoms.

The idea of domain partition based on a particular (observed or computed) conformational change is complementary to an alternative way of domain partition that assumes certain local structural fragments such as  $\alpha$  helices and  $\beta$  strands as rigid bodies and predicts the motions between them (for example, in Block Normal Mode method<sup>28</sup>). Since  $\alpha$  helices and  $\beta$  strands are often deformable structurally, the rigid body assumption is not always valid. Given the additional information of a conformational change, it is preferable that we identify rigid domains that are consistent with the given conformational change without assuming the local structural fragments to be rigid.

## RESULTS

### Comparison Between Normal Modes and Observed Conformational Changes in NS3 Helicase

Following our previous work,<sup>22</sup> we conduct NMA for the ENM built upon the  $C_\alpha$  coordinates of an HCV NS3 helicase structure (PDB: 1A1V). To visualize the detailed motions, we perform a dynamical domain partition (see Methods) for the conformational motions (with RMSD normalized to 1 Å) as described by the two lowest frequency modes: mode 1 and mode 2 (Fig. 1):

- Mode 1 describes a closure rotation (7.9°) of domain 2 toward domain 3, and a very small twisting motion (0.5°) of domain 1 relative to domain 3 [Fig. 1(a)].
- Mode 2 describes two twisting motions of domain 2 (8.1°) and domain 3 (2.3°) relative to domain 1 [Fig. 1(b)].

Therefore, the dynamical domain partitions of the lowest two modes support the high flexibility of domain 2 when compared with the two other domains. This is con-

**TABLE II. Comparison Between the Normal Modes and the Crystallographically Observed Conformational Changes in the HCV NS3 Helicase**

PDB pair	RMSD (Å)	Mode no.	Overlap
1A1V, 1HEI_A	1.14	1	0.49
		Lowest 10	0.59
1A1V, 1HEI_B	1.76	1	0.44
		2	0.50
		Lowest 10	0.79
1A1V, 8OHM	3.11	1	0.92
1A1V, 1CU1_A	1.74	1	0.95
		Lowest 10	0.76
1A1V, 1CU1_B	1.72	1	0.87
		1	0.84
		Lowest 10	0.90

The overlap values are shown for both the dominant mode(s) and the lowest 10 modes combined. In the latter case, the cumulative

overlap for the lowest 10 modes is computed by  $\sqrt{\sum_{m=1}^{10} I_m^2}$ , where  $I_m$  is the overlap of mode “m.”

sistent with previous studies based on comparisons of different crystal structures, which suggested that domain 2 could undergo rigid-body movements relative to domains 1 and 3.<sup>14</sup>

To quantitatively assess how well the normal modes describe the observed conformational changes in the NS3 helicase, we compare the eigenvectors of the modes with the observed conformational changes from 1A1V to several other NS3 structures by computing the overlaps between them (Table II). We find that all the observed conformational changes are dominated by mode 1 (and together with mode 2 in one case—1HEI\_B). By including all the lowest 10 modes, we can only achieve a slightly higher cumulative overlap than just using the modes 1 and 2 (see Table II). Therefore, the modes 1 and 2 not only qualitatively capture the high flexibility of domain 2 but also quantitatively describe the observed conformational changes. In contrast, the remaining modes contribute very little (overlap < 0.2) individually, although their inclusion may be needed for more accurate descriptions of the observed conformational changes (particularly in case of 1HEI\_A and 1HEI\_B). The above observations validate the use of ENM-based analyses to probe the conformational changes in the NS3 helicase.

### Correlation Analysis

Next we perform a correlation analysis, which perturbs individual residues and computes how that changes the overall fluctuations of a selected subset of residues (see Methods). The goal is to identify those dynamically important hot-spot residues, which are defined as the residues strongly coupled with two functionally important subsets of residues: one comprised of those residues involved in ATP binding/hydrolysis, and the other comprised of the residues involved in polynucleotide binding.



**TABLE III. Hot-Spot Residues Coupled With Two Subsets of Residues: ATP-Binding Site and Polynucleotide-Binding Site**

	Subset of residues	Hot-spot residues (eight clusters)
ATP-binding site	207–212 (motif I)	190, 203, 204, 206–214 (motif I)
	229–231 (motif Ia)	291, 293, 294 (motif II)
	290–293 (motif II)	322–330 (motif III)
	322, 323 (motif III)	412 (polynucleotide binding)
	467 (motif VI)	432 (polynucleotide binding) 450, 456–463 (motif VI) 479–482, 484, 485, 521
Polynucleotide-binding site	230, 232, 254, 255, 269, 271, 272, 275, 298 (domain 1)	293, 294 (motif II)
	369–371, 392, 393, 411, 413 (domain 2)	323–327, 329, 330 (motif III)
	432, 434, 448, 450(interface: domains 2 and 3)	369, 370 (polynucleotide binding)
	501, 502 (interface: domains 1 and 3)	411, 412, 414 (polynucleotide binding)
		430, 432–436 (polynucleotide binding) 446, 448–452, 456–459, 461 (motif VI) 477, 479–482, 484, 485, 489, 490, 493, 521

HCV NS3 helicase and F1 ATPase are both RecA-like ATPases, which contain a highly conserved ATP-binding site.<sup>29</sup> We identify the first subset of ATP-binding residues based on the sequential and structural alignment between the HCV NS3 helicase (PDB: 1A1V) and the bovine F1 ATPase (PDB: 1BMF) (see Methods).<sup>30</sup> It contains 16 residues that are conserved between NS3 and F1 ATPase, which are believed to be involved in ATP binding/hydrolysis (see Table III). The second subset of polynucleotide-binding residues includes 22 residues in contact with DNA in the crystal structure of 1A1V (see Table III), which are obtained from PDBsum.<sup>31</sup>

For the subset of ATP-binding residues, we find that its hot-spot residues, whose perturbations lead to significant changes in the fluctuations of ATP-binding residues, span an extensive network that not only covers the ATP-binding site, but also reaches the distant polynucleotide-binding site [see Table III and Fig. 2(a)]. Among them are several highly conserved residues of SF2 motifs, implicated for helicase functions by other studies (for example: E291 and H293 of motif II, T322-P325 of motif III, Q460-G463 of motif VI). Some conserved polynucleotide-binding residues (D412, V432) are found to be dynamically coupled with the ATP-binding site, supporting their important role in modulating ATPase activity by sensing polynucleotide binding.

For the subset of polynucleotide-binding residues, we find its hot-spot residues also constitute a network that not only overlaps with the polynucleotide-binding site but also reaches the ATP-binding site [see Table III and Fig. 2(b)]. Among them are several highly conserved residues from SF2 motifs that are directly involved in ATP binding, including H293 of motif II, A323-P325 of motif III, Q460 and R461 of motif VI, supporting their critical role in modulating polynucleotide binding.

Therefore, the above two sets of hot-spot residues, although defined for two spatially separated subsets of

residues, share many common residues with each other. These key residues are proposed to mediate dynamical couplings and mutual modulations between the ATP-binding residues and the polynucleotide-binding residues.

Several of these hot-spot residues have been examined by kinetic experiments of mutational species of NS3 helicase (see later). Their coupling behaviors are consistent with the observation from the correlation analysis. Those hot-spot residues that have not yet been experimentally explored suggest new mutational studies. Here we list a few important hot-spot residues in light of mutational studies on NS3 helicase:

1. H293: it is located at the bottom of the interdomain cleft and  $\sim 4$  Å away from V456 and Q460. This histidine was shown to be essential for coupling the ATPase activity to polynucleotide binding: mutations of this histidine in HCV NS3<sup>32</sup> and vaccinia NPH-II helicases<sup>33</sup> result in a functional ATPase with no helicase activity. Therefore, H293 is thought to be a gatekeeper that modulates opening/closing of domain 1 and domain 2 upon ATP binding,<sup>34</sup> consistent with our finding that this residue is a hot-spot residue strongly coupled to the polynucleotide-binding site (Table III).
2. T324: as part of motif III (322–326), this residue connects domains 1 and 2. Mutational studies showed T324A reduces ATPase activity dramatically, which was suggested to result from their roles in modulating the opening and closure of the ATP binding cleft between these two domains.<sup>34</sup> Another mutational study showed that the T<sup>322</sup>AT sequence couples the NTP hydrolysis and the duplex unwinding by the enzyme.<sup>35</sup> A structural comparison between the HCV NS3 helicase and the Rep helicase (from SF1 family) supports the role of motif III in a relay mechanism

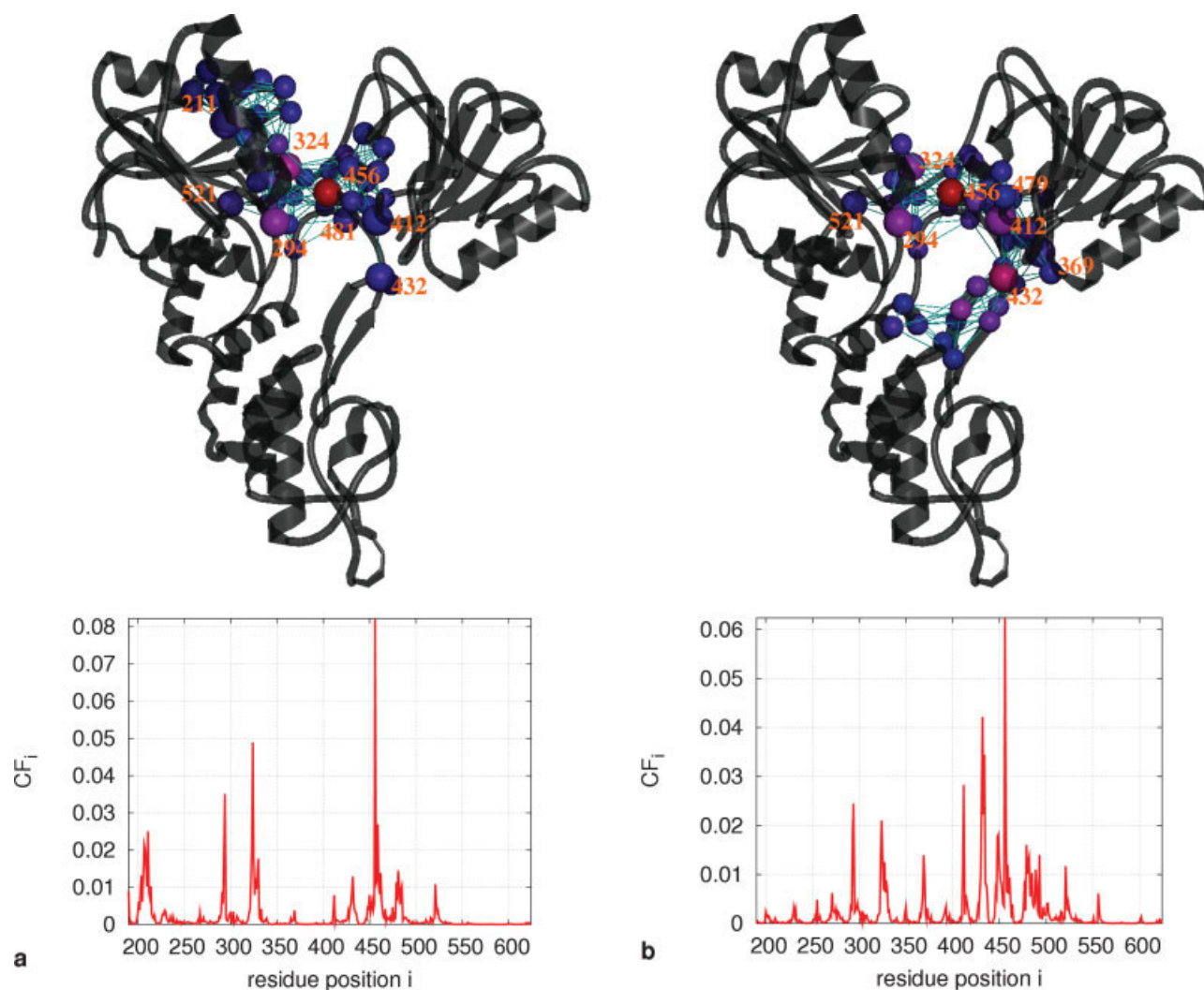


Fig. 2. Dynamically important hot-spot residues for two subsets of residues: (a) ATP-binding residues; (b) DNA-binding residues. In the upper panels, the hot-spot residues are shown as filled circles. Color scheme: red/blue indicates high/low values of correlation functions. Eight representative hot-spot residues are labeled in each panel. In the lower panels,  $CF_i$  (correlation function as defined in Methods) as a function of residue position is shown for each subset.

for transducing the allosteric effects of nucleotide binding and hydrolysis on DNA binding: direct in Rep and indirect in HCV.<sup>36</sup> These results are consistent with our finding of the dynamical importance of T324, which is a hot-spot residue strongly coupled to the polynucleotide-binding site (Table III).

3. V432: this residue is adjacent to the polynucleotide and it is highly conserved among the HCV NS3 sequences. The V432-dU4 base interaction was proposed to induce unstacking of bases and lead to translocation of the polynucleotide in the 5' to 3' direction as domain 2 closes.<sup>14</sup> The correlation analysis shows this residue is a hot-spot residue strongly coupled to the ATP-binding site.
4. Q460, R461: Q460 is located at the bottom of the cleft between domains 1 and 2, opposite from H293. R461 points away from the cleft and is hydrogen bonded to D412 (D412 lines the polynucleotide-binding chan-

nel).<sup>14</sup> A mutant Q460H was shown to possess 34% wild-type NTPase activity but helicase activity detectable only at basal level. Therefore, Q460 is involved in the coupling of ATPase to helicase activity but not directly in ATP hydrolysis.<sup>37</sup> These results are consistent with our finding of the dynamical importance of Q460 and R461: both Q460 and R461 are shown as hot-spot residues strongly coupled to the ATP-binding site, and R461 is also shown as a hot-spot residue strongly coupled to the polynucleotide-binding site.

We then decompose the overall fluctuations of the given subset to individual normal modes to find the modes that contribute substantially (see Methods). We find that mode 1 contributes most for both subsets (its weight is 32.3% for the ATP-binding site, 41.4% for the polynucleotide-binding site), whereas mode 2 ranks second (18% for the ATP-binding site, 22.3% for the polynu-

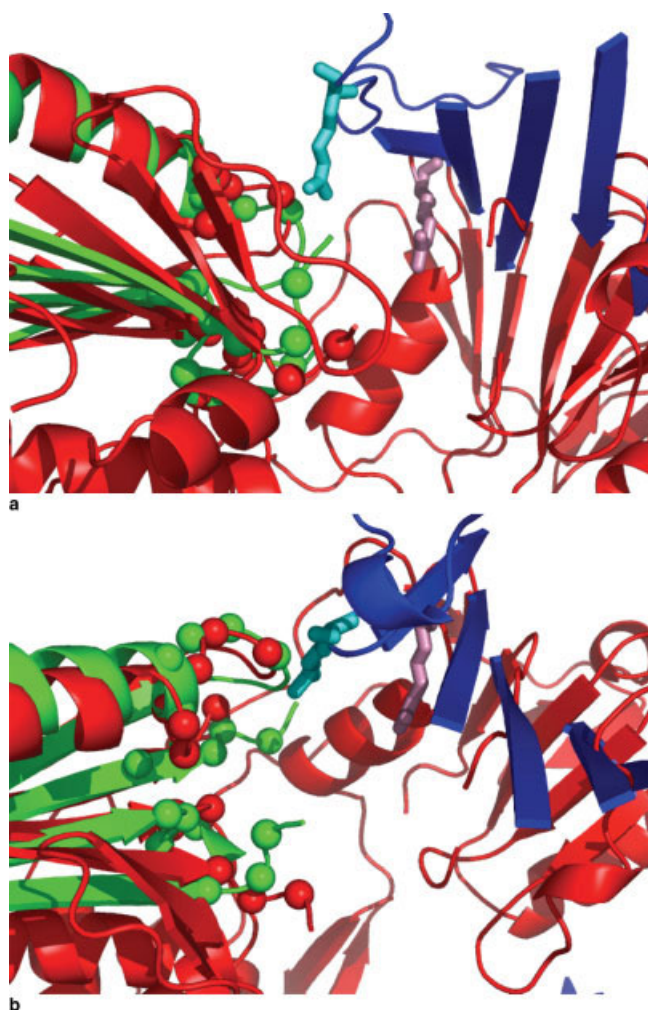


Fig. 3. Structural alignment at the ATP-binding site between the HCV NS3 helicase (PDB: 1A1V) and F1 ATPase (PDB: 1BMF): (a) front view; (b) top view of the ATP-binding site. 1A1V is shown in red; 1BMF's two chains are shown in green (chain B) and blue (chain F). ATP-binding residues are shown as filled circles. Note there is large difference in alignment between R373 of 1BMF (cyan) and R467 of 1A1V (purple).

cleotide-binding site). This result supports the dominant role that mode 1 (and mode 2 to a less extent) plays to modulate the fluctuations at both the ATP-binding site and the polynucleotide-binding site.

### Deformation Analysis

Complementary to the earlier correlation analysis, which analyzes the amplitude of fluctuations at ligand-binding sites, the deformation analysis further probes the directionality of global structural changes triggered by ligand binding by applying local forces to the binding site.

In the crystal structure of NS3 helicase (PDB: 1A1V), the ATP-binding site is empty and thus adopts an open conformation. No crystal structures have been solved for the ATP-bound state of NS3 helicase. To understand the effect of hydrogen bond network formed during ATP

binding in NS3 helicase, we seek for homologous proteins with structurally solved ATP-bound state. As stated earlier, both NS3 helicase and F1 ATPase belong to the large family of RecA-like ATPases, which possess highly conserved structural motifs in the ATP-binding site.<sup>29,38</sup> Furthermore, F1 ATPase is one of the best studied RecA-like ATPases with structures solved in many conformational states.<sup>30,39,40</sup> Assuming the hydrogen-bond network formed by ATP binding is highly similar in the two proteins due to the strong structural homology between the HCV NS3 helicase and the F1 ATPase at the ATP-binding site, we can approximately model the residue positions at the ATP-binding site of NS3 helicase in the ATP-bound (or closed) conformation using the closed conformation of F1 ATPase.

Here we take the closed conformation of the ATP-binding site from an ATP-bound crystal structure of the bovine F1 ATPase (PDB: 1BMF's chains B, F) and structurally superimpose it on top of the "open-state" ATP-binding site of the HCV NS3 helicase (PDB: 1A1V) (see Fig. 3). We find that the most obvious difference is in the position of the Arginine finger, i.e. R373 of F1 ATPase's chain B and R467 of HCV NS3,<sup>41–43</sup> suggesting a significant displacement of this residue as a result of ATP binding. The movement of Arginine finger has been shown in the open-to-closed conformational change of F1 ATPase and of other RecA-like ATPases.<sup>30,43–46</sup>

Then we apply external forces to the ATP-binding residues of NS3 helicase to displace them toward the closed-state coordinates given in 1BMF, and compute the induced conformational change of the whole helicase in response (see Methods). As shown in our previous study of myosin,<sup>22</sup> this calculation allows us to approximately predict the large-scale interdomain motions triggered by a given local structural deformation, which are often well described by a handful of lowest frequency modes.

The induced conformational change (Fig. 4) is then analyzed by dynamical domain partition (Methods). We find that it consists of two rigid-body motions between three dynamical domains, which roughly coincide with the three domains (domains 1, 2, 3) used in NS3 literature. Domain 2 closes toward domain 3 by a rotation of 25.4°. There is also a smaller twisting motion of domain 1 by 11.1° relative to domain 3. As a result of the aforementioned two motions, domain 2 grips tighter on polynucleotide and simultaneously pushes it toward domains 1 and 3; meanwhile, by swinging away from the polynucleotide, domain 1 weakens its polynucleotide-binding affinity, which allows polynucleotide to slide through the cleft between domains 1 and 3. These motions are consistent with the key propositions of the Kim model.<sup>14</sup>

The induced conformational change is found to be dominated by mode 1 (overlap = 0.89), which describes a similar combination of two rotations: a large closure rotation of domain 2 toward domain 3, accompanied by a small twisting motion of part of domain 1 relative to domain 3 [see Fig. 1(a)]. This is consistent with the dominance of mode 1.



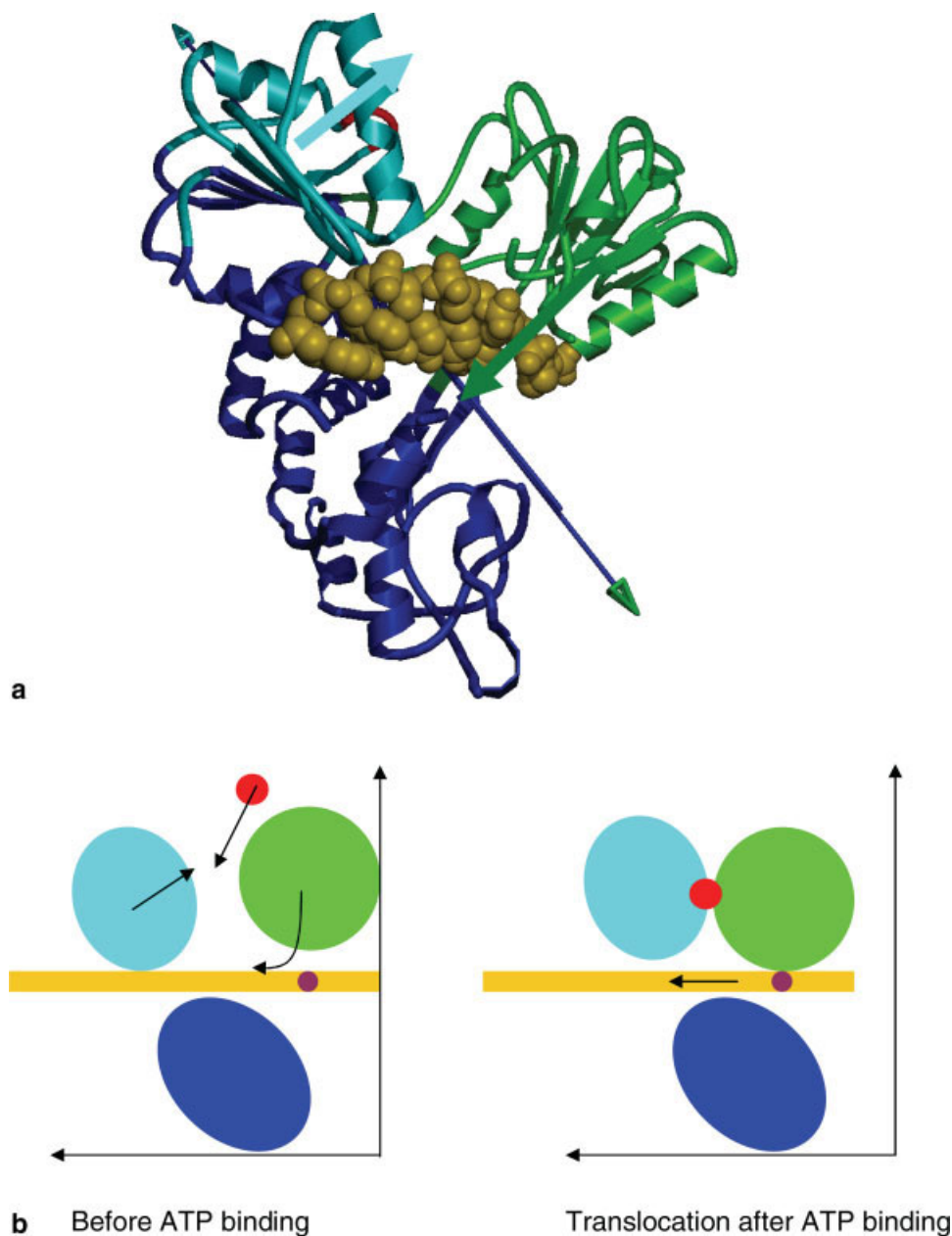


Fig. 4. ATP-binding induced conformational change. (a) Dynamical domain partition: it consists of two rigid-body motions between three dynamical domains: domain 1 in cyan (residue range: 190–194, 200–204, 210–269, 281–285), domain 2 in green (residue range: 321–414, 418–430, 432–434, 450–484), and domain 3 in blue (residue range: 195–199, 270–278, 280, 286–320, 435–449, 485–498, 500–620). The flexible bending regions are colored in red. The bound DNA is shown as CPK in gold. The two long arrows represent rotation axes (see Fig. 1 legend). The two short arrows point to directions of movements of domains 1 (cyan: away from DNA) and 2 (green: toward DNA). (b) Schematic cartoon showing the conformations before (left) and after ATP binding (right): the color scheme is the same as in (a), the motions are shown by arrows. The red dot represents the incoming ATP and the DNA is colored by gold.

## DISCUSSION

In this study we have investigated the dynamical mechanisms in HCV NS3 helicase based on a simple elastic model. By comparing the ENM-derived normal modes with the observed conformational changes between different crystal structures of NS3 helicase, we

have found the lowest two modes to dominate the observed changes, thus validating the application of ENM in NS3 helicase. Using our ENM-based correlation analysis we have identified dynamically important hot-spot residues, which are consistent with published mutational studies.<sup>32–37</sup> Furthermore this analysis predicts several hot-spot residues whose mutational effects have

not yet been explored experimentally. This provides suggestions for future experiments. Our deformation analysis provides a dynamic model for the ATP-binding driven translocation of single-stranded polynucleotide in NS3 helicase, which is consistent with the inchworm mechanism observed by the Bustamante laboratory.<sup>10</sup>

The identification of mode 1 as most dominant by both analyses, together with its dominance in the observed conformational changes between different NS3 helicase structures, all support the functional importance of the opening/closing motion of domain 2 as described by this mode. This large-scale hinge motion is unique in that it simultaneously modulates the opening/closing of the domains 1–2 cleft where the ATP-binding site is, and the domains 2–3 cleft where the single-stranded polynucleotide binds. Furthermore, this motion requires low energy cost, because domain 2 is flexibly linked to the rest of enzyme. As a result of this closure motion, we propose that the domain 2 grips tighter on polynucleotide and pushes it toward domains 1 and 3 to drive its translocation. Additionally, the small twisting motion of domain 1 induced by ATP binding may fine-tune the polynucleotide-binding affinity of the domains 1–3 interface in an opposite way than that of the domains 2–3 interface, which may further facilitate the translocation of a single-stranded polynucleotide in an inchworm fashion. Therefore, despite their structural similarity, domain 1 and domain 2 play different roles in the translocation process: the former is an affinity modulator while the latter is both an affinity modulator and a motor that drives the translocation.

Our results support a key proposition of the inchworm scenario<sup>18</sup>: the domains 1 and 2 switch alternatively between tight and weak polynucleotide binding states and move relative to each other. However, the inchworm scenario invokes six conformational states, while our results suggest that the above process could be completed by a one-step power-stroke that is executed by the highly coordinated motions between domains 1 and 2. Therefore two states (prestroke and poststroke) rather than six is sufficient to accommodate the predicted ATP-binding-induced conformational change. Our results provide a detailed dynamical picture for the forward movement (power stroke), which is also a key component of the Brownian motor model.<sup>47</sup> The Brownian motor model, however, proposes that the power stroke occurs after ATP hydrolysis when NS3 helicase assumes the tight polynucleotide binding state along single-stranded polynucleotide, while our results support a power-stroke triggered by ATP binding.

Computer representation of receptor flexibility is a major consideration in the de novo design of small molecules inhibitors.<sup>48</sup> In particular, the ability to represent backbone changes taking place during allosteric motions has so far been unavailable. The demonstration of the potential of the ENM methods presented in this paper provides a promising approach, which will allow representations of backbone conformational changes whose consideration is important for drug-design purposes.

## REFERENCES

1. Frick DN. Helicases as antiviral drug targets. *Drug News Perspective* 2003;16:355–362.
2. Druker BJ, Talpaz M, Resta DJ, Peng B, Buchdunger E, Ford JM, Lydon NB, Kantarjian H, Capdeville R, Ohno-Jones S, Sawyers CL. Efficacy and safety of a specific inhibitor of the BCR-ABL tyrosine kinase in chronic myeloid leukemia. *N Engl J Med* 2001;344:1031–1037.
3. Gorbalenya AE, Koonin EV, Donchenko AP, Blinov VM. Two related superfamilies of putative helicases involved in replication, recombination, repair and expression of DNA and RNA genomes. *Nucleic Acids Res* 1989;17:4713–4730.
4. Lee CG, Hurwitz J. A new RNA helicase isolated from HeLa cells that catalytically translocates in the 3' to 5' direction. *J Biol Chem* 1992;267:4398–4407.
5. Ray BK, Lawson TG, Kramer JC, Cladaras MH, Grifo JA, Abramson RD, Merrick WC, Thach RE. ATP-dependent unwinding of messenger RNA structure by eukaryotic initiation factors. *J Biol Chem* 1985;260:7651–7658.
6. Gwack Y, Kim DW, Han JH, Choe J. DNA helicase activity of the hepatitis C virus nonstructural protein 3. *Eur J Biochem* 1997;250:47–54.
7. Levin MK, Patel SS. The helicase from hepatitis C virus is active as an oligomer. *J Biol Chem* 1999;274:31839–31846.
8. Tackett AJ, Chen Y, Cameron CE, Raney KD. Multiple full-length NS3 molecules are required for optimal unwinding of oligonucleotide DNA *in vitro*. *J Biol Chem* 2005;280:10797–10806.
9. Serebrov V, Pyle AM. Periodic cycles of RNA unwinding and pausing by hepatitis C virus NS3 helicase. *Nature* 2004;430:476–480.
10. Dumont S, Cheng W, Serebrov V, Beran RK, Tinoco I, Pyle AM, Bustamante C. RNA translocation and unwinding mechanism of HCV NS3 helicase and its coordination by ATP. *Nature* 2006;439:105–108.
11. Yao N, Hesson T, Cable M, Hong Z, Kwong AD, Le HV, Weber PC. Structure of the hepatitis C virus RNA helicase domain. *Nat Struct Biol* 1997;4:463–467.
12. Yao N, Reichert P, Taremi SS, Prorise WW, Weber PC. Molecular views of viral polyprotein processing revealed by the crystal structure of the hepatitis C virus bifunctional protease-helicase. *Structure* 1999;7:1353–1363.
13. Cho HS, Ha NC, Kang LW, Chung KM, Back SH, Jang SK, Oh BH. Crystal structure of RNA helicase from genotype 1b hepatitis C virus. A feasible mechanism of unwinding duplex RNA. *J Biol Chem* 1998;273:15045–15052.
14. Kim JL, Morgenstern KA, Griffith JP, Dwyer MD, Thomson JA, Murcko MA, Lin C, Caron PR. Hepatitis C virus NS3 RNA helicase domain with a bound oligonucleotide: the crystal structure provides insights into the mode of unwinding. *Structure* 1998;6:89–100.
15. Gorbalenya AE, Koonin EV, Donchenko AP, Blinov VM. Two related superfamilies of putative helicases involved in replication, recombination, repair and expression of DNA and RNA genomes. *Nucleic Acids Res* 1989;17:4713–4730.
16. Gorbalenya AE, Koonin EV. Helicases: amino acid sequence comparisons and structure-function relationships. *Curr Opin Struct Biol* 1993;3:419–429.
17. Caruthers JM, McKay DB. Helicase structure and mechanism. *Curr Opin Struct Biol* 2002;12:123–133.
18. Velankar SS, Soultanas P, Dillingham MS, Subramanya HS, Wigley DB. Crystal structures of complexes of PcrA DNA helicase with a DNA substrate indicate an inchworm mechanism. *Cell* 1999;97:75–84.
19. Levin MK, Gurjar MM, Patel SS. A Brownian motor mechanism of translocation and strand separation by hepatitis C virus helicase. *Nat Struct Mol Biol* 2005;12:429–435.
20. Atilgan AR, Durell SR, Jernigan RL, Demirel MC, Keskin O, Bahar I. Anisotropy of fluctuation dynamics of proteins with an elastic network model. *Biophys J* 2001;80:505–515.
21. Delarue M, Sanejouand YH. Simplified normal mode analysis of conformational transitions in DNA-dependent polymerases: the elastic network model. *J Mol Biol* 2002;320:1011–1024.
22. Zheng W, Doniach S. A comparative study of motor protein motions using a simple elastic network model. *Proc Natl Acad Sci USA* 2003;100:13253–13258.
23. Zheng W, Brooks BR. Identification of dynamical correlations within the myosin motor domain by the normal mode analysis of an elastic network model. *J Mol Biol* 2005;346:745–759.

24. Zheng W, Brooks BR. Probing the local dynamics of nucleotide-binding pocket coupled to the global dynamics: myosin versus kinesin. *Biophys J* 2005;89:167–178.
25. Tirion MM. Large amplitude elastic motions in proteins from a single-parameter, atomic analysis. *Phys Rev Lett* 1996;77:1905–1908.
26. Hinsen K. Analysis of domain motions in large proteins. *Proteins* 1998;33:417–429.
27. Hayward S, Berendsen HJC. Systematic analysis of domain motions in proteins from conformational change; New results on citrate synthase and T4 lysozyme. *Proteins* 1998;30:144–154.
28. Li G, Cui Q. A coarse-grained normal mode approach for macromolecules: an efficient implementation and application to Ca(2+)-ATPase. *Biophys J* 2002;83:2457–2474.
29. Bird LE, Subramanya HS, Wigley DB. Helicases: a unifying structural theme? *Curr Opin Struct Biol* 1998;8:14–18.
30. Abrahams JP, Leslie AG, Lutter R, Walker JE. Structure at 2.8 Å resolution of F1-ATPase from bovine heart mitochondria. *Nature* 1994;370:621–628.
31. Laskowski RA, Chistyakov VV, Thornton JM. PDBsum more: new summaries and analyses of the known 3D structures of proteins and nucleic acids. *Nucleic Acids Res* 2005;33:D266–D268.
32. Heilek GM, Peterson MG. A point mutation abolishes the helicase but not the nucleoside triphosphatase activity of hepatitis C virus NS3 protein. *J Virol* 1997;69:4727–4736.
33. Gross CH, Shuman S. Mutational analysis of vaccinia virus nucleoside triphosphate phosphohydrolase II, a DExH box RNA helicase. *J Virol* 1995;69:4727–4736.
34. Tai CL, Pan WC, Liaw SH, Yang UC, Hwang LH, Chen DS. Structure-based mutational analysis of the hepatitis C virus NS3 helicase. *J Virol* 2001;75:8289–8297.
35. Pause A, Sonenberg N. Mutational analysis of a DEAD box RNA helicase: the mammalian translation initiation factor eIF-4A. *EMBO J* 1997;11:2643–2654.
36. Korolev S, Yao N, Lohman TM, Weber PC, Waksman G. Comparisons between the structures of HCV and Rep helicases reveal structural similarities between SF1 and SF2 super-families of helicases. *Protein Sci* 1998;7:605–610.
37. Wardell AD, Errington W, Ciaramella G, Merson J, McGarvey MJ. Characterization and mutational analysis of the helicase and NTPase activities of hepatitis C virus full-length NS3 protein. *J Gen Virol* 1999;80:701–709.
38. Ye J, Osborne AR, Groll M, Rapoport TA. RecA-like motor ATPases—lessons from structures. *Biochim Biophys Acta* 2004;1659:1–18.
39. Menz RI, Walker JE, Leslie AG. Structure of bovine mitochondrial F(1)-ATPase with nucleotide bound to all three catalytic sites: implications for the mechanism of rotary catalysis. *Cell* 2001;106:331–341.
40. Wang H, Oster G. Energy transduction in the F1 motor of ATP synthase. *Nature* 1998;396:279–282.
41. Nadanaciva S, Weber J, Wilke-Mounts S, Senior AE. Importance of F1-ATPase residue alpha-Arg-376 for catalytic transition state stabilization. *Biochemistry* 1999;38:15493–15499.
42. Caruthers JM, McKay DB. Helicase structure and mechanism. *Curr Opin Struct Biol* 2002;12:123–133.
43. Soutanas P, Dillingham MS, Velankar SS, Wigley DB. DNA binding mediates conformational changes and metal ion coordination in the active site of PcrA helicase. *J Mol Biol* 1999;290:137–148.
44. Crampton DJ, Guo S, Johnson DE, Richardson CC. The arginine finger of bacteriophage T7 gene 4 helicase: role in energy coupling. *Proc Natl Acad Sci USA* 2004;101:4373–4378.
45. Putnam CD, Clancy SB, Tsuruta H, Gonzalez S, Wetmur JG, Tainer JA. Structure and mechanism of the RuvB Holliday junction branch migration motor. *J Mol Biol* 2001;311:297–310.
46. Liao JC, Jeong YJ, Kim DE, Patel SS, Oster G. Mechanochemistry of t7 DNA helicase. *J Mol Biol* 2005;350:452–475.
47. Levin MK, Gurjar MM, Patel SS. ATP binding modulates the nucleic acid affinity of hepatitis C virus helicase. *J Biol Chem* 2003;278:23311–23316.
48. Alberts IL, Todorov NP, Dean PM. Receptor flexibility in de novo ligand design and docking. *J Med Chem* 2005;48:6585–6596.