Structure, Vol. 13, 565–577, April, 2005, ©2005 Elsevier Ltd All rights reserved. DOI 10.1016/j.str.2005.01.017

Network of Dynamically Important Residues in the Open/Closed Transition in Polymerases Is Strongly Conserved

Wenjun Zheng,[1,*] Bernard R. Brooks,[1] Sebastian Doniach,[2] and D. Thirumalai[3,*]
[1]Laboratory of Computational Biology
National Heart, Lung, and Blood Institute
National Institutes of Health
Bethesda, Maryland 20892
[2]Departments of Physics and Applied Physics and Laboratory for Advanced Materials
Stanford University
Stanford, California 94305
[3]Institute for Physical Science and Technology
University of Maryland,
College Park, Maryland 20742

Summary

The open/closed transition in polymerases is a crucial event in DNA replication and transcription. We hypothesize that the residues that transmit the signal for the open/closed transition are also strongly conserved. To identify the dynamically relevant residues, we use an elastic network model of polymerases and probe the residue-specific response to a local perturbation. In a variety of DNA/RNA polymerases, a network of residues spanning the fingers and palm domains is involved in the open/closed transition. The similarity in the network of residues responsible for large-scale domain movements supports the notion of a common induced-fit mechanism in the polymerase families for the formation of a closed ternary complex. Multiple sequence alignment shows that many of these residues are also strongly conserved. Residues with the largest sensitivity to local perturbations include those that are not so obviously involved in the polymerase catalysis. Our results suggest that mutations of the mechanical "hot spots" can compromise the efficiency of the enzyme.

Introduction

Polymerases play a central role in copying the genetic information with great fidelity. Although these enzymes are linked by a common function, they have vastly different sequences. Despite the possibility that polymerases must have appeared early in evolution, the lack of sequence similarity (except in the active site) is surprising. Using sequence comparison and crystallographic structures, polymerases have been classified into several different families. The molecular structures of a number of polymerases have been determined (Steitz, 1999). These include the well-characterized structures belonging to DNA polymerase I (or pol I), DNA polymerase α (pol α), reverse transcriptase (RT), and RNA-dependent RNA polymerases. Although the detailed structures of the polymerases in the four families (pol I, pol α, RT, and polymerase family X) vary greatly, they share similar structural features in the U-shaped DNA binding cleft. The common structural feature is described, following the description for the Klenow fragment (Ollis et al., 1985), using the right hand metaphor with palm, fingers, and thumb denoting the various domains. The thumb and fingers domains display considerable variations in structures. However, there is a common structural core in the palm domain of all polymerases except those in pol β family. The common structural core consists of the conserved arganine residues that bind dNTP and are involved in the catalysis of the polymerization reaction.

Based on the structures of the binary polymerase-DNA complexes and the ternary polymerase-DNA-dNTP complex, a model for DNA polymerization has been suggested (Steitz, 1999). In the first step, the duplex DNA binds to the unliganded polymerase triggering the thumb to close around the DNA. Binding dNTP to the binary complex causes rotation of the fingers from the open conformation to a closed state that is primed for catalysis. The closed form is required to get a firm grip by the enzyme on the substrate, while the open form facilitates translocation. It is clear that the open/closed transition in DNA/RNA polymerase, which involves large-scale "mechanical" movement of domains, is functionally significant. The purpose of this study is to first determine the residues that trigger the open/closed transition and then assess the extent to which these residues are conserved. To achieve this, we use normal mode analysis (NMA) of the Elastic Network Model (ENM) to analyze DNA/RNA polymerases.

In recent years NMA analysis of ENM of large protein complexes have been used to obtain insights into ligand-induced conformational transitions (Atilgan et al., 2001; Delarue and Sanejouand, 2002; Kundu and Jernigan, 2004; Tama and Sanejouand, 2001; Zheng and Doniach, 2003). The low-frequency normal modes computed from the ENM have been successful in describing large-amplitude motions of protein complexes. Collective motions in the hinge-bending regions in the palm and connection domains of HIV-1 reverse transcriptase have been studied using a Gaussian network model (Bahar et al., 1999). Delarue and Sanejouand (2002) used ENM to investigate the open/closed transition in DNA-dependent polymerases and found that such transitions are typically well described by the lowest and/or the second lowest frequency normal modes deduced from the structure of the open form. The previous studies have focused only on the overall motions using the ENM of polymerases. A molecular understanding also requires identification of the network of residues that are involved in triggering the domain motions. We hypothesize that if a given normal mode is linked to a specific biological function that requires accurate coordination of interdomain motions, the network of residues that are dynamically important to the normal mode may be evolutionarily strongly conserved. If this hypothesis holds then mutations of the dynamically relevant residues can compromise the efficiency

*Correspondence: thirum@glue.umd.edu (D.T.); zhengwj@helix.nih.gov (W.Z.)

of the enzymes. Linking conservation of the network of residues and its role in function-related dynamics allows for a mapping between biological functions and the corresponding mechanically "hot" residues.

We establish the hypothesized relation between the dynamical importance of a network of residues in conformational transitions and their sequence conservation using four DNA/RNA polymerases as test cases for which structures and mutational data are available. Extensive analysis reveals striking structural analogies among the pol families for the fingers domains (see reviews Steitz, 1999 and Doublie et al., 1999). The open/closed transition of the fingers domain, observed in several families of polymerases upon dNTP binding, is believed to be the rate-limiting conformational change needed for tight dNTP binding and subsequent catalysis (induced fit model; Steitz, 1999). Following the study by Delarue and Sanejouand (2002), we focus on the open/closed transition of the fingers domain in several polymerases to establish our hypothesis. In all polymerases there is a network of residues spanning a large part of the structure that are strongly conserved. Our results show that mutations of these residues would adversely affect the domain motions needed to form a tight complex with the incoming substrate. The alteration in the normal modes, in response to a perturbation, shows that communications between residues that are far apart in sequence and structure is needed to trigger the mechanical open/closed transition in polymerases. The methodology introduced here is relevant for probing allosteric transitions in other biological nanomachines.

Results and Discussion

Our hypothesis is that clusters of residues that are involved in large-scale domain movements are also evolutionarily highly conserved. To test this proposal, we consider the biologically important open/closed transition in four DNA/RNA polymerases. The polymerases in the closed form enable the tight interaction between the enzyme and the template substrate. Thus, it is important to decipher the network of residues that are involved in this complex reorganization and to assess the extent to which they are conserved.

To establish the proposed hypothesis, we first identify the normal mode for each polymerase that describes the domain movement in the open/closed transition. For the modes of interest, we calculate (see Experimental Procedures) residue level response to a local perturbation. The final step involves correlating the effect of perturbation on specific residues to their conservation using multiple sequence alignment.

Identification of the Normal Mode in the Open/Closed Transition

Taq DNA Polymerase I

Waksman and coworkers (Li et al., 1998) solved the structures of two ternary complexes of the large fragment of *Thermus Aquaticus* (Taq) DNA polymerase I bound to a DNA duplex and dNTP. The transition from the open form (Protein Database [PDB] code, 2KTQ) to the closed from (PDB code, 3KTQ) of the enzyme is thought to consist of two successive rotations of helices N, O, O1, and O2. This complex motion results in the closing of the finger-palm crevice (Li et al., 1998). The O helix (residues 660–671) changes its orientation significantly in order to approach the active site (Figure 1B) so that the first single-stranded DNA base of the template can pair with dNTP. This results in the assembly of a productive closed ternary complex poised for catalysis. Residue Y671 on the O helix, which is part of the motif B, is highly conserved in pol I family (Delarue et al., 1990).

The NMA analysis (Experimental Procedures; Table 1) for the open form structure 2KTQ shows that the open/closed transition is dominated by mode 4 (overlap = 0.50; see also Delarue and Sanejouand, 2002) that involves the fingers domain rotating toward the active site. It describes well the measured conformational change (Figure 1A, fourth panel) involving the large-scale rotation of the O helix. The motion of O1 and O2 helix is somewhat overestimated. The thumb movement predicted by mode 4 is not seen in the measured conformational change presumably because it is inhibited in the presence of a bound DNA duplex.

Bacteriophage T7 RNA pol I

The C-terminal two-thirds of T7 RNA polymerase is homologous to the polymerase domain of the pol I DNA polymerase family. With the exception of a few insertions on T7 RNA pol, alignment of the homologous regions of T7 RNA pol with *E. coli* Klenow fragment (KF; Ollis et al., 1985) provides a good fit on the fingers domain (Jeruzalmi and Steitz, 1998). Helix Y of T7 RNA pol (residues 644–661) aligns precisely with helix O1 (residues 770–778) of KF—a motif that is also found in Taq pol I. Among the set of residues conserved in Pol I family is the Y639 in motif B (homologous to Y766 in KF and Y671 in Taq pol I), which is believed to be critical to the regulated stacking of the active site.

The NMA analysis (Table 1) for the open form structure (PDB code, 1ARO; Jeruzalmi and Steitz, 1998) shows that the open/closed transition is dominated by mode 1 (overlap = 0.66) that describes a hinge motion of the fingers domain relative to the rest of the structure (the thumb domain is absent in 1ARO). Mode 1 correctly describes the measured conformational change (Figure 2A, fourth panel). The observed conformational change also includes a translational shift of the upper fingers which moves both ends of helix Y.

Human/Rat DNA pol β

DNA polymerase β (pol β) is a small DNA gap-filling polymerase. Structural alignments with other polymerases revealed a common active-site geometry in the palm domain (Pelletier et al., 1994). However, they differ structurally in the thumb and fingers domains. Structural comparison with the pol I family shows that the N helix (residue 276–288) of the thumb domain is the analog of the O helix of pol I (Li et al., 1998). The common function of the two helices is the interaction with the incoming dNTP in the closed ternary complex. Because motif B does not exist in pol β as in pol I family, it is proposed that the thumb domain, instead of the fingers domain, interacts with the incoming dNTP to form a closed ternary complex (Pelletier et al., 1994).

In the open state (PDB code, 1BPX; Pelletier et al., 1996) to the closed state 1BPY (PDB code, 1BPY; Pel-

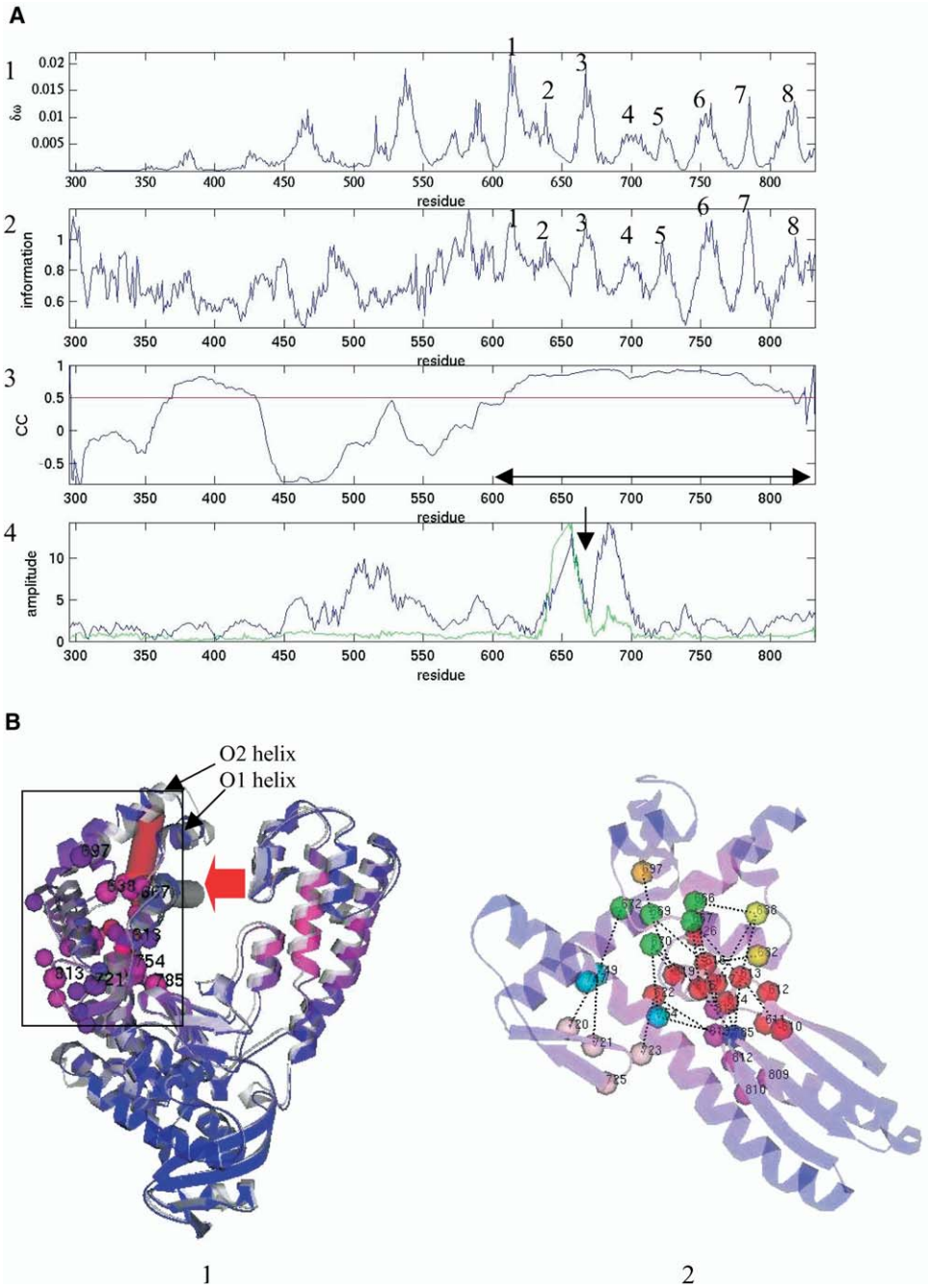

Figure 1. Analysis of the ENM for Taq pol I

(A) Panel 1 shows δω (residue-dependent response to local perturbation) in mode 4. Numbers 1–8 denote clusters of residues with high-δω values centered at positions 613, 638, 667, 697, 721, 754, 785, and 813, respectively. Panel 2 shows the residue information S calculated using multiple sequence alignment with PsiBlast. The highly conserved eight clusters of residues (1–8) also have high-δω values (compare panels 1 and 2). The position-specific crosscorrelation coefficient between δω and S, with the red line representing CC = 0.5 in the third panel, shows significant correlation between residues with high S and δω. The fourth panel shows the amplitude of displacement of mode 4 (blue) and the amplitude of the measured conformational change (green).

(B) Panel 1 shows the structural alignment of 3KTQ (gray) with 2KTQ (color change from blue to red indicates increase in δω). The O helix (residues 660–671) is shown as cylinder, and the O1 helix (residues 674–679) and the O2 helix (residues 686–699) are pointed by arrows. The high-δω residues (Q613, F667, I638, E615, L813, L616, D785, A814, Q754, D610, N666, L670, L622, M747, V810, V669, I614, E721, P812, G809, L723, Y611, R617, F632, V720, E626, G725, L619, G672, F697, S612, and F749, in order of descending δω) are shown, and the eight center residues are numbered. Panel 2 shows the enlarged side view of the hinge region (structure in the box in panel 1), where the high-δω residues are numbered and shown as space filled circles. The colors represent clusters: 1, red; 2, yellow; 3, green; 4, orange; 5, gray; 6, cyan; 7, blue; and 8, magenta. The top-ranking links within 10 Å between clusters are shown as dashed lines (see Experimental Procedures).

Table 1. Summary of the Results for the NMA-Based δω Calculation for Four Polymerases

Name	Conformational Change (PDB)	Dominant Mode	Overlap	Z_w
Taq DNA pol I	2KTQ→3KTQ	4	0.50	0.28
T7 RNA pol I	1ARO→1CEZ	1	0.66	0.42
Human/Rat DNA pol β	1BPX→1BPY	1	0.71	0.04
		3	0.49	0.13
HIV-1 RT	1N5Y→1RTD	4	0.51	0.35

letier et al., 1994) transition, a large change in the position of the thumb domain results in a more closed active site in 1BPY than in 1BPX (Pelletier et al., 1996). The NMA analysis (Table 1) for the open structure shows that the open/closed transition is dominated by mode 1, with an overlap = 0.73 that is identical to that found by Delarue and Sanejouand (2002). This mode describes a hinge motion of the 8-KD/fingers domains relative to the thumb/palm domains with both having similar amplitude (Figure 3A, fourth panel). (The measured conformational change has larger movement in the thumb domain than in the fingers, which is consistent with the 8-KD/fingers being bound to the DNA duplex). The observed movement of the thumb domain that closes on the palm domain is described by the subdominant mode 3 (overlap = 0.49). This mode describes the hinge motions between the thumb and palm domains and between the fingers and palm.

HIV-1 Reverse Transcriptase

HIV-1 reverse transcriptase (RT) is responsible for the catalytic transformation of single-stranded viral RNA into the double-stranded linear DNA. It is a dimer of two related chains, p66 and p51, that have four common domains (fingers, palm, thumb, and connection). We will focus on the polymerase activity of p66, for which multiple structures of different states have been solved. We chose the open binary form (PDB code, 1N5Y; Sarafianos et al., 2002) and the closed ternary form (PDB code, 1RTD; Huang et al., 1998). For analysis in this study, we include p51 in the normal modes calculation but exclude it from the statistical analysis.

In the transition between 1N5Y and 1RTD, the fingers tip closes where the outer part of the fingers domain rotates with respect to the base. Hinge points for the three projecting loops are near residues 21/45, 58/77, and 130/144 (Huang et al., 1998). The NMA analysis (see Experimental Procedures; Table 1) for the open form structure 1N5Y shows that the open/closed transition is dominated by the mode 4 (overlap = 0.51) that describes the rotation of the three projection loops in the fingers. It gives a good description of the measured conformational change (Figure 4A, fourth panel), although the scale of movement in the rest of the structure is overestimated.

Dynamically Important Residues Are Strongly Conserved

Taq DNA Polymerase I

To test the hypothesis that dynamically relevant clusters of residues (ones with high δω values) are also likely to be evolutionarily conserved (with high residue information) we calculated δω (Experimental Procedures) for mode 4. The high-δω residues are distributed over the hinge region covering the fingers, palm, and thumb domains (Figure 1B). The positive Z_w (0.28) (Experimental Procedures) suggests a statistically favorable conservation of the high-δω residues. To obtain a detailed picture of the highly conserved regions, we have examined the correlation between residue information S and δω (Figure 1A, first and second panels). Comparison of the results in the first two panels in Figure 1A shows clearly that clusters of residues with high δω values also have high S values. This correspondence is particularly impressive for the part of sequence between residue 300 and the C-terminal that covers the fingers and palm domains. The position-specific crosscorrelation coefficient CC (see Experimental Procedures) exceeds 0.5 in this region (Figure 1A, third panel). These results support the proposed hypothesis for this polymerase. There are a few clusters of high-δω residues in the thumb domain that do not appear to be conserved. This may be because mode 4 does not provide a correct description of the thumb movement.

Bacteriophage T7 RNA pol I

The high-δω residues for mode 1 are distributed over the hinge region covering the fingers and palm domains (Figure 2B). The positive Z_w (0.42) suggests a statistically favorable conservation of the high-δω residues. Comparison of the results in the first and second panel of Figure 2A shows that clusters of residues with high δω values also have high S values. The correlation is especially sharp for the part of sequence between residue 400 and the C-terminal which covers the fingers and palm domains. The position-specific cross-correlation coefficient CC (Experimental Procedures) exceeds 0.5 in this region (Figure 2A, third panel).

Human/Rat DNA pol β

Because of the importance of two modes (1 and 3), we calculated δω for residues for both modes. For mode 1, the high-δω residues are distributed near the fingers/palm domains while the thumb is essentially distortion-free (Figure 3A). Its small Z_w value (0.04) indicates insignificant conservation of the high-δω residues. For mode 3, the high-δω residues are distributed over both the fingers/palm domains and the thumb domain (Figure 3A), and its Z_w (0.13) indicates favorable conservation of the high-δω residues.

Comparison of S and δω in the first and second panel of Figure 3A shows the following: First, the correlation for mode 1 is relatively poor, which is inferred from the lack of a long contiguous segment of sequence with the position-specific crosscorrelation coefficient CC that exceeds 0.5. Second, for mode 3 we find a much better correlation for the part of sequence between 120 and C-terminal (six clusters are identified). In this case CC exceeds 0.5 in this region (Figure 3A, third panel). For mode 3, there are two clusters with high-δω residues on the 8-KD/fingers domain (centered at 20 and 100) that do not match with conserved residues (Figure 3A). This is consistent with the inaccurate description or overestimation of the fingers movement by mode 3. For this polymerase, we will focus on mode 3 from now on.

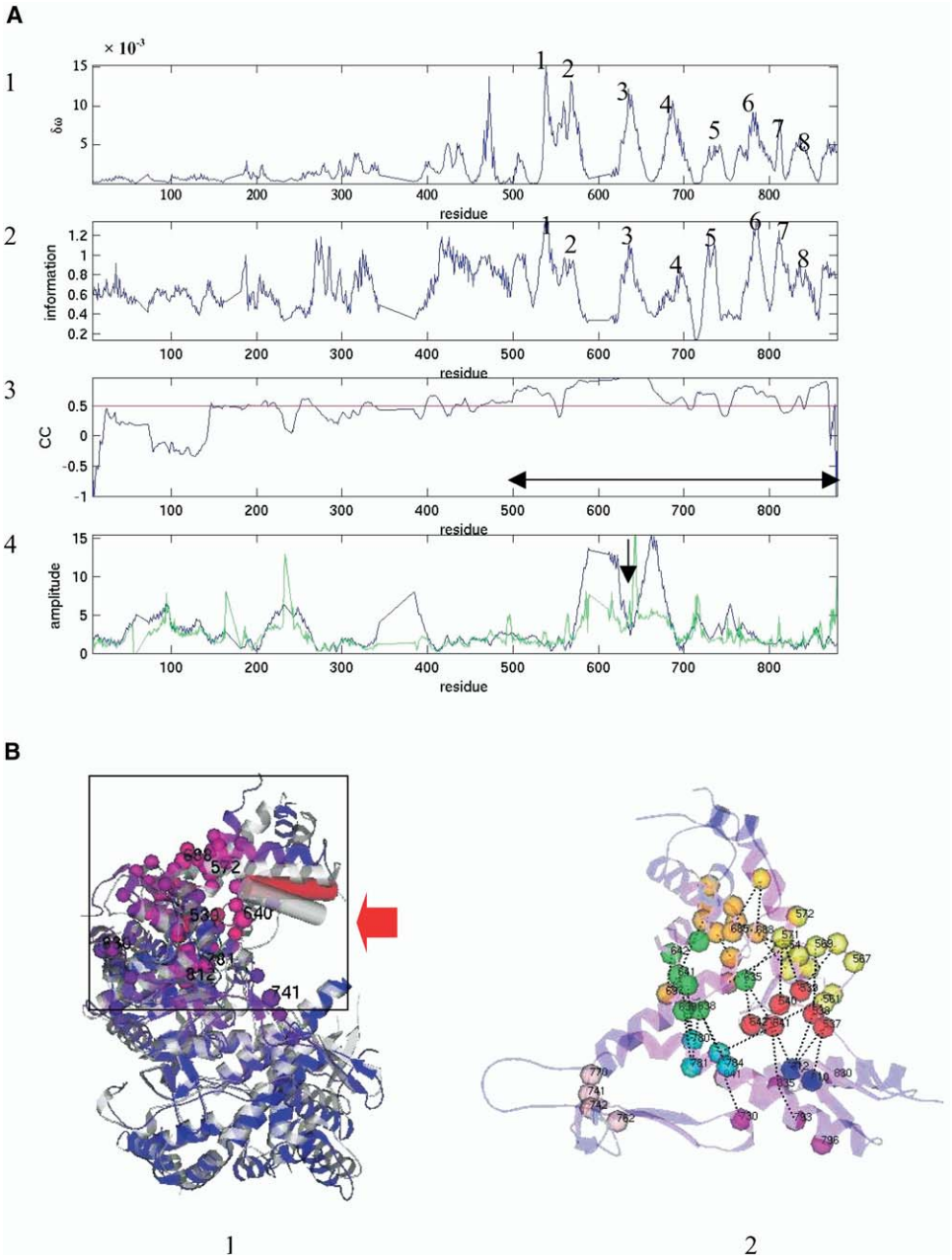

Figure 2. Results for T7 RNA pol

(A) Panel 1 shows δω for mode 1. Numbers 1–8 denote clusters of residues with high δω values centered at positions 539, 572, 640, 688, 741, 781, 812, and 830, respectively. The residue information is shown in panel 2. See caption to Figure 1 for further explanation of the remaining panels.

(B) Panel 1 shows the structural alignment of 1CEZ (gray) with 1ARO (same color code as in Figure 1B). Helix Y is shown as cylinder. The high-δω residues (S539, N781, S541, D812, M635, G640, T688, G572, P780, Y571, A638, Y639, S684, D537, V690, G538, Q568, K741, S641, H784, S686, R557, I810, V841, P730, K642, E830, E683, V567, V687, G542, P742, P563, V554, K577, K793, V685, D770, D569, V796, T835, C540, N697, A691, L680, L561, W682, N762, E553, V783, Q649, A548, and A558, in order of descending δω) are shown, and the eight center residues are numbered. Panel 2 shows the enlarged side view of the hinge region (structure in the box in panel 1) where the high-δω residues are numbered and shown as space filled circles. The colors represent clusters: 1, red; 2, yellow; 3, green; 4, orange; 5, gray; 6, cyan; 7, blue; and 8, magenta. The top-ranking links within 10 Å between clusters are shown as dashed lines (see Experimental Procedures).

HIV-1 Reverse Transcriptase

For mode 4, the high-δω residues are distributed over the hinge region covering the fingers and palm domains (Figure 4B). The positive Z_w (0.35) suggests a statistically favorable conservation of the high-δω residues. The results in the first two panels of Figure 4A show that clusters of residues with high δω values also have high S values. This correspondence is most definitive for the part of sequence between residue 50 and 200 which covers the fingers and palm domains. The position-

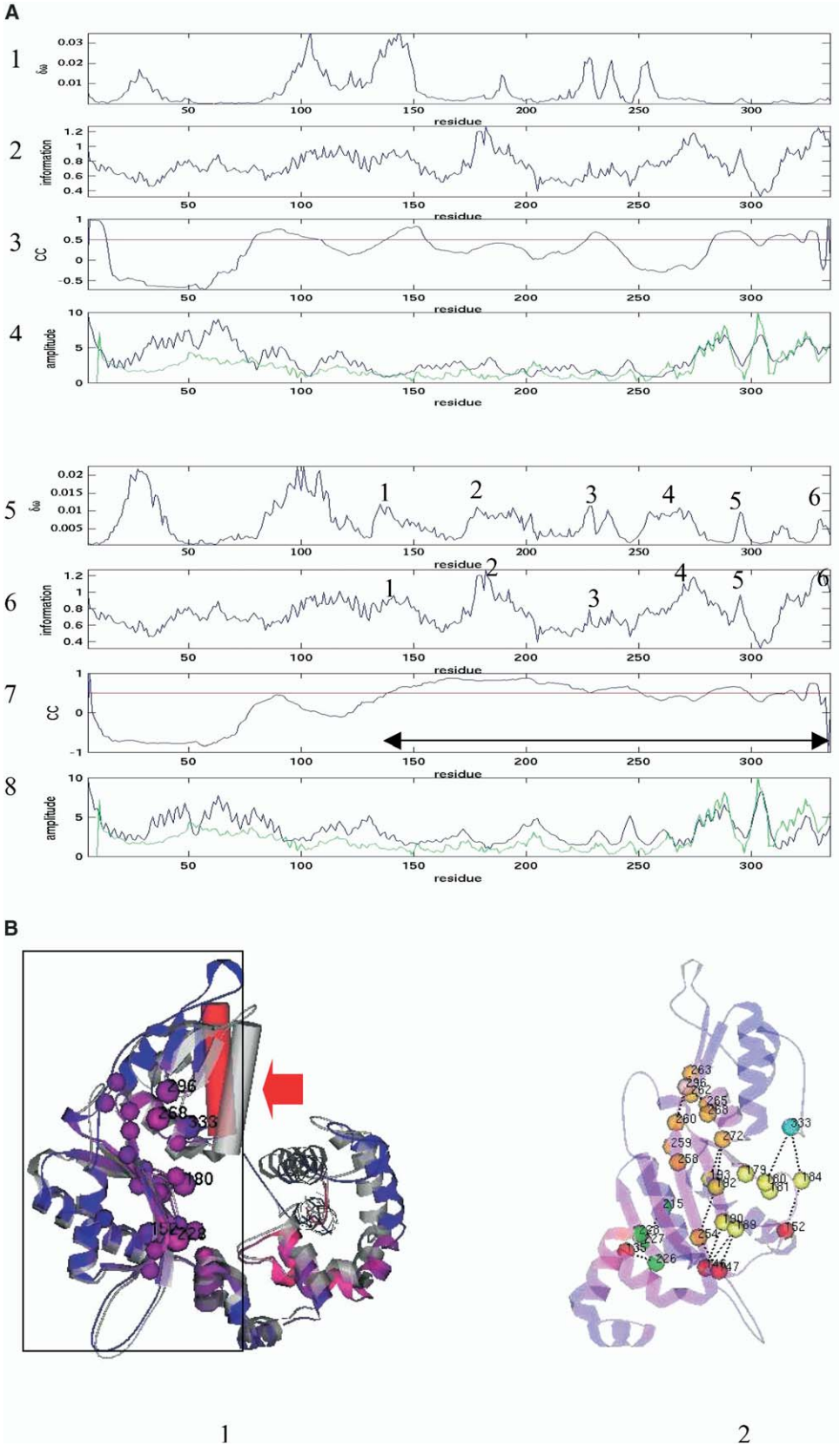

Figure 3. Summary of Findings for pol β

(A) Panel 1 (5) shows δω for mode 1 (3). For mode 3, numbers 1–6 denote clusters of residues with high-δω values centered at positions 152, 180, 228, 268, 296, and 333, respectively. Panel 2 (6) shows the residue information S. For mode 3, the highly conserved six clusters of residues (1–6) also have high δω values (compare panels 5 and 6). See caption to Figure 1 for further explanation of the remaining panels.

specific crosscorrelation coefficient CC exceeds 0.5 in this region (Figure 4A, third panel).

The general finding that the clusters of residues that have high δω-values are, to a large extent, highly conserved for these four polymerases validates our hypothesis. These clusters of residues with high degree of conservation are involved in the large-scale domain movements that facilitate ligand binding. This implies that a network of residues that are far apart in sequence and are spatially distant is conserved for the "mechanical" function of the polymerases.

Identifying the Network of Dynamically Important Residues in the Open/Closed Transition

It is important to identify the set of residues in the signaling process that trigger the open/closed transition. To this end we select residues that have high δω values from the fingers/palm domains (see Experimental Procedures for details). They form a network connected by pairwise links between spatial neighbors that are within 10 Å. Based on the elastic energy stored within each link (see Experimental Procedures), we rank them and focus on those "high-ranking" links (Experimental Procedures). The clusters of such important residues, which are both spatially apart and are also well separated in sequence, are identified as transmitting forces between spatial neighbors through the high-ranking links. This supports the existence of long-range communication in the open/closed transition in polymerases.

Taq DNA Polymerase I

The network of high-δω residues in the hinge region is formed by eight clusters (Figure 1B). The center of this network is formed by Cluster 1, which is linked to five neighboring clusters (2, 3, 5, 7, and 8), and residues in cluster 3 are connected to those in clusters 1, 2, 4, and 6. The residues in these clusters are directly or indirectly linked via the high-ranking links (Figure 1B, second panel). The interactions between residues from different clusters are likely to be important in the modulation of the open/closed conformational change. Perturbations due to ligand (dNTP) binding can induce responses in ligand binding residues (clusters 1, 3, and 5), which can then be transmitted, via these links, to the entire hinge region (especially nonligand binding residues in clusters 4, 6, and 8). Such a process triggers the functionally important hinge motion, eventually resulting in the open/closed transition.

A detailed discussion of the residues in these clusters follows:

1. Residues in cluster 1 (D610, Y611, S612, Q613, I614, E615, L616, R617, L619, L622, and E626) and 2 (F632 and I638) are on the fingers/palm interface (Li et al., 1998). Some of these are involved in dNTP binding and catalysis (conserved motif A: 605–617). The mutant Taq pol containing hydrophilic substitution I614K has significantly reduced fidelity and has a high propensity to extend mispairs (Patel et al., 2001). Residues L619, L622, F632, and I638 in Pol I family are likely to be involved in the stabilization of the fingers-palm interface and modulation of the open/closed conformational change.
2. Residues in cluster 3 (N666, F667, V669, L670, and G672) are located in the fingers domain that covers the O helix (660–671). Among these, F667 and G672 are involved in DNA and dNTP binding. Mutations of the O helix residues (A661, I665, and F667) show that interactions within O helix residues can contribute to fidelity. Finally, V669 and L670 are conserved in Pol I family.
3. Residues in clusters 4 (F697), 5 (V720, E721, L723, and G725), and 6 (M747, F749, and Q754) are on the fingers domain. Some of these (M747 and Q754) are involved in DNA binding. F697, V720, E721, L723, and F749 are conserved in Pol I family.
4. Cluster 7 (D785), which is located in the palm domain is involved in DNA binding and catalysis. It is one of the two active site carboxylates for catalysis.
5. Residues in cluster 8 (G809, V810, P812, L813, and A814) are in the palm domain. V810 and L813 are conserved in Pol I family for their hydrophobicity.

In summary, residues in the eight clusters with high-δω (Figure 1A) are highly conserved. The conservation of some of these is for functional reasons including interactions with DNA/dNTP. However, there are a number of strongly conserved high-δω residues that do not directly interact with DNA/dNTP. We associate their conservation with their dynamical importance to the normal mode 4, which is responsible for a fingers-closing conformational change needed for the formation of a tightly closed ternary complex of polymerase. These residues are good targets for mutagenesis study to further examine how they affect the function, possibly by altering the rate of the open/closed transition.

The proposed link between the conserved residues involved in the mechanical open/closed transition and the polymerase function has some experimental support. Residues I614 (cluster 1) and F667 (cluster 3), which have been shown by mutational studies to be critical for high-fidelity DNA polymerization, strongly support the connection between the fingers-closing normal mode and the fidelity of DNA synthesis (Delarue and Sanejouand, 2002). The critical role of residues in cluster 3 that serves as a hinge region to the fingers rotation motion is also evident from the plot of the displacement amplitude of mode 4. It corresponds to the deep valley between two sharp peaks corresponding

(B) Panel 1 shows the structural alignment of 1BPY (gray) with 1BPX (same color code as in Figure 1B). Helix N is shown as cylinder. The high-δω residues (L228, I260, G268, H135, Y296, D226, S180, F272, D192, T227, G179, G184, K262, R258, D263, R152, G189, D190, Y265, V193, F146, V215, R254, L259, F181, E147, and R333, in order of descending δω) are shown, and the six center residues are numbered. Panel 2 shows the enlarged side view of the hinge region (structure in the box in panel 1), where the high-δω residues are numbered and shown as space filled circles. The colors represent clusters: 1, red; 2, yellow; 3, green; 4, orange; 5, pink; 6, cyan. The top-ranking links within 10 Å between clusters are shown as dashed lines (see Experimental Procedures)

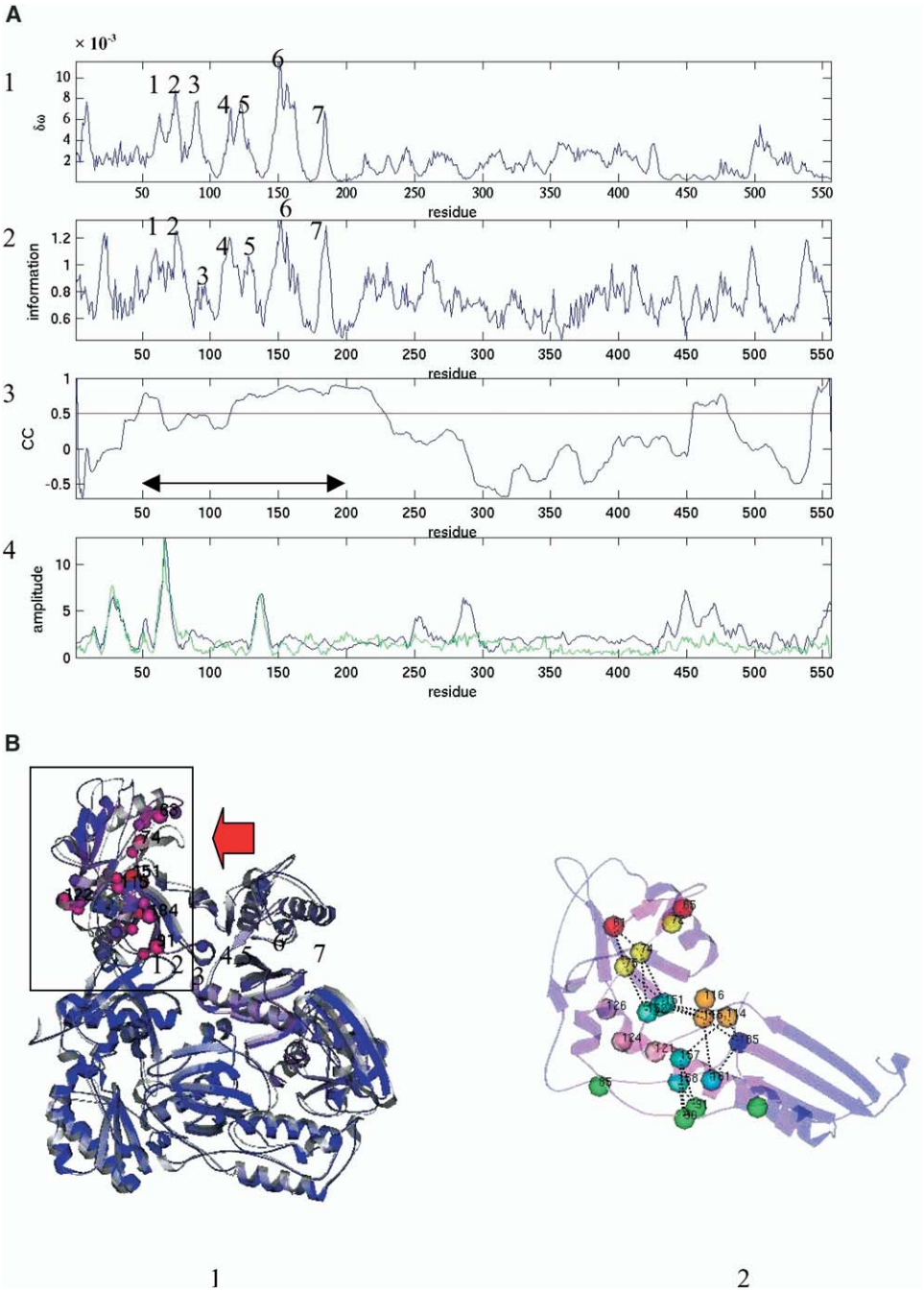

Figure 4. Analysis of the ENM for HIV-1 RT

(A) Panel 1 shows δω for mode 4. Numbers 1–7 denote clusters of residues with high-δω values centered at positions 65, 74, 91, 116, 126, 157, and 185, respectively. Panel 2 shows the residue information S. The highly conserved seven clusters of residues (1–7) also have high δω values (compare panels 1 and 2). See caption to Figure 1 for further explanation of the remaining panels.

(B) Panel 1 shows the structural alignment of 1RTD (gray) with 1N5Y (same color code as in Figure 1B). The closing of the outer part of the fingers (the loop 58–77 in particular) toward the palm is shown. The high-δω residues (P157, L74, G152, Q151, F116, V75, Q91, A114, K126, F124, R72, K65, Q85, Y115, D121, V90, F61, D185, A158, Q161, P150, and H96, in order of descending δω) are shown, and the seven center residues are numbered. Panel 2 shows the enlarged side view of the hinge region (structure in the box in panel 1) where the high-δω residues are numbered and shown as space filled circles. The colors represent clusters: 1, red; 2, yellow; 3, green; 4, orange; 5, pink; 6, cyan; 7, blue. The top-ranking links within 10 Å between clusters are shown as dashed lines (see Experimental Procedures).

Table 2. List of Crystal Structures of DNA/RNA Polymerases Used in the Present Study

Proteins (Family Name)	PDB (State)
Thermus aquaticus (Taq) DNA pol I (pol A/pol I)	2KTQ (open ternary: DNA) 3KTQ (closed ternary: ddCTP+DNA)
Phage T7 RNA pol (RNA pol monosubunit)	1ARO (binary: T7 lysozyme) 1CEZ (open binary: promoter)
Human/Rat DNA pol β (pol X/pol β)	1BPX (binary: DNA with gap) 1BPY (ternary: ddCTP+DNA)
HIV-1 reverse transcriptase (RT)	1N5Y (binary) 1RTD (closed ternary)

to the tip regions of the fingers (Figure 1A, pointed by an arrow) that undertakes the largest motion.

Bacteriophage T7 RNA pol I

Just as before, the network of high-δω residues in the hinge region is formed by eight clusters (Figure 2B). The center of the network is formed by cluster 1, which is linked to five neighboring clusters (2, 3, 6, 7, and 8). Residues in cluster 3 are connected to those in 1, 2, 4, and 6. The residues in these clusters are directly or indirectly linked by the high-ranking links except for cluster 5 (Figure 2B, second panel). The interactions between residues from different clusters are important in the modulation of the open/closed conformational change.

The structural homolog of the fingers and palm domains between T7 RNA pol and DNA pol I allows for a mapping between these clusters.

1. Residues in cluster 1 (D537, G538, S539, C540, S541, G542, and A548) are on the palm domain. Among these the two residues important in catalysis, D537 and S539, are identical to Taq pol's D610 and S612. Therefore, these map onto cluster 1 of Taq pol. In addition, G538, C540, S541, and G542 are also conserved although they are not directly involved in the active site. Interaction between residues in clusters 2, 3, 6, and 7 and those in cluster 1 are mediated by C540, S541, and G542.
2. Residues in cluster 2 (E553, V554, R557, L561, P563, V567, Q568, D569, Y571, G572, and K577) are on the fingers domain. The active site residue D569 is identical to Taq pol's D637. Therefore, it maps onto cluster 2 of Taq pol. Others such as V554, L561, V567, Y571, and G572 are also conserved although they are not directly involved in the active site. Hydrophobic residues V554, L561, and Y571 are involved in intercluster interactions with clusters 1, 3, and 4.
3. Residues in cluster 3 (M635, A638, Y639, G640, S641, K642, and Q649) are on the fingers domain. It contains dNTP binding residues (M635 and Y639 of motif B), and helix Y, which is homologous to helix O in Taq pol I; therefore, it corresponds to cluster 3 of Taq pol. The conserved residues M635, A638, S641, and K642 are involved in intercluster interactions with clusters 1, 2, 4, and 6.
4. Cluster 4 (L680, W682, E683, S684, V685, S686, V687, T688, V690, A691, and N697) is on the fingers domain. It maps onto cluster 4 of Taq pol by structural homology alone (no conserved residues are identified). The conserved residues V685, T688, and N697 are involved in intercluster interactions with clusters 2 and 3.
5. Residues in cluster 5 (K741, P742, N762, and D770), that are on the fingers domain, map to cluster 5 of Taq pol by structural homology (G732 is identical with Taq pol's G725).
6. Residues (P780, N781, V783, and H784) in cluster 6 are on the palm/fingers interface, in proximity to the motif B-containing cluster 3. It maps onto cluster 6 of Taq pol by structural homology. All these residues except V783 are conserved.
7. Residues in cluster 7 (I810 and D812) are on the palm domain. It maps onto cluster 7 of Taq pol because of the conservation of active-site residues (H811 and D812 of T7 are identical with H784 and D785 of Taq pol).
8. Residues in cluster 8 (E830, T835, V841, P730, K793, and V796) are on the palm domain. It maps onto cluster 8 of Taq pol because M832 is identical with Taq pol's M807. All these residues are conserved.

The detailed analysis given above shows that residues in the eight clusters with high-δω (Figure 2A) are strongly conserved. Using structural homology, we found that they all have corresponding counterparts in Taq pol I. The high degree of conservation of some of the residues is for functional reasons, including interactions with DNA/dNTP. Just as in Taq pol I, we associate the highly conserved high-δω residues that do not interact with DNA/dNTP as being dynamically relevant in triggering the fingers-closing conformational change that ensures the formation of a tightly closed ternary complex of polymerase. We predict that mutations of these residues will compromise the efficiency of the function of the polymerase.

The critical role of residues in cluster 3, which serves as a hinge region to the fingers rotation motion, is evident from the plot of the displacement amplitude of mode 1. It corresponds to the deep valley between two sharp peaks corresponding to the tip regions of the fingers (Figure 2A, fourth panel, arrow) that undergoes the largest motion.

The combination of our NMA based method and the structural homology between T7 RNA pol and Taq pol allows us to link the eight clusters of interacting network of residues between the two polymerases. The network of residues in these two structures are highly conserved and are not only involved in the catalytic function of the polymerase but also in the dynamical function of the open/closed transition. Surprisingly, in both cases, a single normal mode is associated with the dynamical function of fingers closing.

We find only a weak correlation between S and δω in the rest of T7 RNA pol (see the region 1–400 in Figure 2A). The dynamics is not conserved between Taq pol and T7 RNA pol in this region, which is consistent with the observation that the structural homology between these two polymerases is limited only to the fingers/

palm domains. This suggests that the critical dynamic role of the fingers domain is not shared by the other domains.

Human/Rat DNA pol β

The network of high-δω residues in the hinge region is formed by six clusters (Figure 3B) where the spatial connectivity varies. The center of this network is formed by cluster 1, which is linked to four neighboring clusters (clusters 2, 3, 4, and 6). The residues in these clusters are directly or indirectly linked via the high-ranking links (Figure 3B, second panel). The interactions between residues from different clusters are likely to be important in the modulation of the open/closed conformational change. Now we discuss residues in those clusters:

1. Residues in cluster 1 (H135, F146, E147, and R152), cluster 2 (G179, S180, F181, G184, G189, D190, D192, and V193), and cluster 3 (V215, D226, T227, and L228) are in the palm domain. Several of these are conserved, but are not directly involved in dNTP binding (H135, F146, R152, T227, and L228).
2. Residues in cluster 4 (R254, R258, I260, K262, D263, Y265, G268, and F272) are at the palm/thumb interface. F272, which is located within dNTP binding pocket, is known to be critical for maintaining fidelity during the binding of dNTP (Li et al., 1999). Conserved residues Q264 and Y265, although not in contact with DNA or dNTP, are involved in the observed thumb movement. A recent mutational study (Opresko et al., 1998) showed that the mutant Y265C of pol β displays an increase in both base substitution and frame shift errors. This is in accord with the importance of Y265 in mediating a conformational change of fingers-closing transition to enhance the accuracy of DNA synthesis (Pelletier et al., 1996).
3. Residues in clusters 5 (Y296) and 6 (R333) are involved in the interaction with dNTP.

In summary, the six clusters of conserved residues seem to match the clusters of high-δω residues for mode 3. These clusters contain a number of residues whose strong conservation seems necessary for dynamical reasons. Similar to the above two pol I cases, for mode 3 of pol β, only weak correlation is found between S and δω in the 8-KD/fingers domains.

Compared with the pol I cases, where the motif B-containing cluster 3 serves as a critical hinge to the fingers movement, the hinge for the thumb movement of pol β seems to involve clusters 4, 5, and 6. This suggests important roles for both helix Y and a nearby β hairpin (residues 290–300) in mediating the open/closed transition. The dynamical roles of the other clusters may warrant further study. Such study will be useful in clarifying whether pol β employs induced fit mechanism similar to that observed in pol I family (Sweasy, 2003).

HIV-1 Reverse Transcriptase

The network of high-δω residues in the hinge region is formed by seven clusters (Figure 4B). The center of this network is formed by cluster 6, which is linked to four neighboring clusters (2, 3, 4, and 7). The residues in these clusters are directly or indirectly linked, by the high-ranking links, except for cluster 5 (Figure 4B, second panel). Residues from different clusters modulate the open/closed conformational change. Now we discuss residues in some of those clusters:

1. Residues in cluster 1 (F61 and K65) and cluster 2 (R72, L74, and V75) are in the fingers domain, forming the hinge of the projecting loop (58–77), which moves most in mode 4 (Figures 4A and 4B). It corresponds to cluster 3 of Taq pol and T7 RNA pol. Mutational study suggested that substitution of D76 (it is neighboring to V75 of cluster 2) by positively charged and nonpolar residues increases the replication accuracy significantly (Kim et al., 1998). These observations suggest that the above loop is likely to be the RT's analog of the O helix of Pol I in that both are important to the fingers closing transition. L74 is conserved but is not directly in contact with dNTP.
2. Residues in cluster 3 (Q85, V90, Q91, and H96), cluster 4 (A114, Y115, and F116), cluster 5 (D121, F124, and K126), cluster 6 (P150, Q151, P157, A158, and Q161), and cluster 7 (D185) are in the fingers/palm interface. Among these, cluster 6 contains several active site residues (Q151 and P157) and has the highest δω, so it corresponds to cluster 1 of Taq and T7 RNAP. Clusters 4 and 7 also contain some of the active site residues (A114, Y115, F116, and D185). Among these residues, F124, P150, and Q161 are conserved but not directly in contact with dNTP.

In a recent work of NMA on HIV-1 RT (Bahar et al., 1999), several regions severely constrained in the collective motion of RT were identified. They correspond to its hinge-bending centers. Two of these regions overlap with our clusters: region 87–110 overlaps with clusters 3 and 4, and region 177–192 overlaps with cluster 7. Both contain the active site of catalysis and the NNRTI binding site, which is consistent with their critical role in modulating the functionally relevant collective motion (Bahar et al., 1999). Our δω-based method enables us to identify additional clusters (clusters 1, 2, and 6) of hinge residues in the fingers domain.

Similar to Taq pol and T7 RNA pol, we find insignificant correlation between residue information and δω in the rest of RT (see the region 20 C-terminal in Figure 4A), which is in sharp contrast with the good matching discussed above. This further hints at the critical dynamic role unique to the fingers/palm domain which is not shared by the other domains.

Discussion

Relation to Other Studies

In several recent studies the powerful ENM have been extended in a number of novel ways (Chacon et al., 2003; Van Wynsberghe et al., 2004). As a complement to these studies, in this work we have introduced a method to probe long-range responses to a local perturbation within the ENM of polymerases. This tech-

nique allows us to identify a network of residues that are also strongly conserved for ensuring efficient open/closed transition. Although these new ideas have not been previously explored, our work is related to a few previous studies. Using the two lowest frequency modes of the Gaussian network model of HIV-1 RT, Bahar and coworkers (Bahar et al., 1999; Temiz and Bahar, 2002) computed the mean square fluctuations. They focused on the dynamics of DNA/inhibitor binding/releasing instead of the open/closed transition triggered by incoming nucleotide as studied here. Therefore, it is not feasible to directly compare our results for the fingers-closing mode 4 with the relevant modes identified in their study. However, both studies attempt to identify hinge region residues that are dynamically critical using different computational methods. It is interesting to note that the hinge residues (with high-$\delta\omega$ values) critical for the open/closed transition partly overlap with the hinge residues (with minimal mobility) identified in their work (Bahar et al., 1999).

Dynamically Relevant Residues Are Highly Conserved

Using the residue-dependent response to a local perturbation in the elastic network representation of polymerase structures we have identified a network of residues that modulate the functionally relevant open/closed transitions. Because the domain movements in large protein complexes occur essentially as rigid bodies, coarse-grained models that neglect the molecular details are remarkably successful (Bahar et al., 1999; Delarue and Sanejouand, 2002; Tama and Sanejouand, 2001; Zheng and Doniach, 2003). As argued earlier (Delarue and Sanejouand, 2002), the small amplitude motion of structural fluctuations in the open/closed transition permits a description of the low-frequency modes.

For the polymerases, there is a remarkable correlation between the distributions of conserved residues and elastic distortion of the relevant normal mode along the protein sequence. The clusters of high-$\delta\omega$ residues are mostly distributed around the flexible joints between relatively rigid structural elements, such as helices or larger compact subdomains, and especially at the interface between several subdomains. Thus, besides the highly conserved aspartate residues in the palm domain which bind to dNTP, and the two metal ions responsible for catalysis, we predict that there are strongly conserved residues that communicate over large spatial distances in the open/closed transition. These residues ensure that the domains movements are coordinated so that tight binding (required for fidelity) to incoming ligand can occur. Our results suggest that mutations of these residues can compromise the polymerase efficiency. Interestingly, there is no correlation between the coordination number of a residue and its extent of conservation. It appears that mechanically relevant residues are also encoded in the structure.

High-$\delta\omega$ Residues Are Responsible for Induced Fit

We find that the clusters of high-$\delta\omega$ residues match well with the clusters of conserved residues for a significant portion of the sequences. The clusters of high-$\delta\omega$ residues not only contain some catalytically important residues, but also other conserved residues not involved in substrate binding. The high-$\delta\omega$ residues form a network that is important in the modulation of the open/closed conformational change. Perturbations induced by ligand (dNTP) binding trigger deformations in ligand binding residues, which can then be transmitted via these links to the whole hinge region so that the functionally important motions (here the open/closed transition) can be initiated. The formation of a closed ternary complex for polymerization reactions involving fingers closing (induced fit mechanism) should be operative in all polymerases with high-fidelity replication. The accuracy of the induced fit conformational changes is ensured by conserving the high-$\delta\omega$ residues. The contacts between those residues and the substrates allow the relevant normal mode to be sensitive to their binding. This may ensure that only the correct pairing of dNTP bases can trigger the open/closed conformational change.

There are also other strongly conserved residues with low-$\delta\omega$ values. We speculate that these residues are conserved for reasons other than the specific dynamics associated with the open/closed transition. They may be needed for ligand binding, for example. Secondly there are non-conserved residues with high-$\delta\omega$. These are found in the hinge region of the thumb domain where there is less conservation pressure, perhaps, due to enhanced flexibility. Additional work is required to clarity their role in the polymerase functions.

Generality of the Proposed Method

In this work we have proposed a very general method to link mechanically important network of residues and sequence conservation. There are numerous other nanomachines (molecular motors, chaperones, etc.) that also undergo largely rigid body motions in response to ligand binding. The procedure we have developed, which combines multiple sequence alignment and the ENM of large structures, can be used to identify the network of residues that signal such domain movements. Identification of the mechanically "hot" residues allows design of experiments that can be used to assess the role of these key residues not directly involved in the active sites.

Experimental Procedures

Elastic Network Model

Given the C_α atomic coordinates for a protein's native structure, we build an elastic network model (Tirion, 1996) by using a harmonic potential with a single force constant to account for pairwise interactions between all C_α atoms that are within a cutoff distance (R_C = 10 Å). The energy in the elastic network representation is:

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

where d_{ij} is the distance between the dynamical coordinates of the C_α atoms i and j, and $d^0{}_{ij}$ is the distance between C_α atoms i and j, as given in the crystal structure. For the harmonic Hamiltonian in Equation 1, we perform the standard normal mode analysis (NMA). The eigenvectors of the lowest frequency normal modes are used to compute the overlaps with the conformational changes between two states with known structures (Zheng and Doniach, 2003).

Assessing the Effect of Point Mutation on Normal Modes

To probe the effect of a point mutation of a residue R (at position n) on a functionally relevant normal mode M, we calculate the response of the springs connected to R to a local perturbation. The harmonic springs connected to R are changed by a small amount δk. This local perturbation can be experimentally achieved by point mutations. The response to the perturbation in the frequency of mode M is calculated using $\delta\omega(M, n) = v_M^T \cdot \delta H \cdot v_M$, where v_M is the eigenvector of mode M and δH is the Hessian matrix of the following perturbation to the energy of the elastic network:

$$\delta E_R = \frac{1}{2} \sum_{d_{nj}^0 < R_c} \delta k (d_{nj} - d_{nj}^0)^2. \quad (2)$$

The response $\delta\omega(M, n)$ is numerically proportional to the contribution to mode M's elastic energy from the springs that are connected to residue R at position n. This procedure is reminiscent of the work by Brooks and Karplus (1985). The higher $\delta\omega(M, n)$ is the more sensitive is mode M to the point mutation of residue R. It is, therefore, natural to conjecture that residues with high $\delta\omega(M, n)$ are likely to be conserved evolutionarily if the mode M is associated with a biological function that requires coordination of the molecule's conformational changes.

To quantitatively test our hypothesis, we use the following procedure:

a. Perform multiple sequence alignment, with default setting of parameters, using Psi-Blast (Altschul et al., 1997). For the ensemble of aligned protein sequences, compute the information content using $S_n = \log 20 + \Sigma_{i=1\ldots20}\, p_i(n)\log(p_i(n))$, where $p_i(n)$ is probability of substitution by amino acid i at position n. A large information value means that a residue at position n is highly conserved.

b. For all positions n of the given protein sequence, compute the average $\langle S\rangle$ and the standard deviation σ_S over all N positions:

$$\langle S\rangle = \frac{\sum_{j=1,\ldots N} S_j}{N}, \sigma_s = \sqrt{\frac{\sum_{j=1,\ldots N} (S_j - \langle S\rangle)^2}{N}}. \quad (3)$$

c. Compute $\langle S\rangle_w$, which is a weighted average of information for all positions, using

$$\langle S\rangle = \frac{\sum_{j=1,\ldots N} S_j \cdot \delta\omega(M, j)}{\sum_{j=1,\ldots N} \delta\omega(M, j)}. \quad (4)$$

d. Calculate the weighted Z score:

$$Z_w = \frac{\langle S\rangle_w - \langle S\rangle}{\sigma_s}. \quad (5)$$

A positive value of Z_w implies a statistically positive correlation between S and $\delta\omega$. On the other hand, a negative or near-zero Z_w indicates a negative or zero correlation between them.

We identify functionally relevant modes using the crystal structures for two states (initial state and end state) of the protein. The modes that overlap significantly with the measured conformational changes between two known structures are relevant. We compute $\delta\omega(M, n)$ and corresponding Z_w only for the structure associated with the initial state. This is because the normal modes computed from this structure govern the transition. Empirically, it is known that NMA gets better overlaps when the transition is from open to closed conformations (Tama and Sanejouand, 2001).

To assess if there is a correlation between S and $\delta\omega$, we perform a coarse-grained average of S and $\delta\omega$. For each residue, we average the S and $\delta\omega$ over a set of residues including itself and all its spatial neighbors. This averaging allows to identify clusters of high-$\delta\omega$ residues that correspond to peaks in the averaged $\delta\omega$ as a function of sequence position (each peak has a center residue which has a local maximum of $\delta\omega$). Similarly, clusters of conserved residues that correspond to peaks in the averaged S can be determined. The regions where the averaged S and $\delta\omega$ match are assumed to be significant to the function associated with the given mode. Use of average value of S is justified because we want to compare with $\delta\omega$, which is computed using only the coarse-grained structural representation. The position-specific crosscorrelation coefficient for a segment of sequence centered at the given position is computed using

$$CC(i,W) = \frac{\sum_{j\in[i-W/2,i+W/2]} \frac{S_j \cdot \delta\omega(M,j)}{W} - \left\{\sum_{j\in[i-W/2,i+W/2]} \frac{S_j}{W}\right\} \cdot \left\{\sum_{j\in[i-W/2,i+W/2]} \frac{\delta\omega(M,j)}{W}\right\}}{\sqrt{\left\{\sum_{j\in[i-W/2,i+W/2]} \frac{S_j^2}{W} - \left\{\sum_{j\in[i-W/2,i+W/2]} \frac{S_j}{W}\right\}^2\right\} \times \left\{\sum_{j\in[i-W/2,i+W/2]} \frac{\delta\omega(M,j)^2}{W} - \left\{\sum_{j\in[i-W/2,i+W/2]} \frac{\delta\omega(M,j)}{W}\right\}^2\right\}}}, \quad (6)$$

with W = 50 as the window size. Significant value of $CC(i,W)$ (say CC > 0.5) serves as a guidance for detecting regions with positive correlation between $\delta\omega$ and S. Negative or near-zero value of CC indicates negative or weak correlation between them.

After identifying the regions with matching peaks (or significant CC) between $\delta\omega$ and S, we calculate the average $\delta\omega$ (denoted as $\langle\delta\omega\rangle$) for the identified regions. We select high-$\delta\omega$ residues as those with $\delta\omega \geq 2\langle\delta\omega\rangle$ (here the factor 2 is arbitrary). The selected high-$\delta\omega$ residues are then classified into clusters corresponding to the identified peaks in the plot of the averaged $\delta\omega$ as a function of sequence position. Each high-$\delta\omega$ residue is assigned to a cluster if it is a sequential neighbor of its center residue. In case where sequential neighbor criterion does not apply, we use spatial neighbor criterion.

We partition $\delta\omega_i$ for a given residue i further into elastic energy stored in each spring that links i to its neighbors j within R_c = 10 Å (denoted as $\delta\omega_{ij}$), which can be ranked and analyzed statistically for all pairs of residues within R_c = 10 Å (excluding j = i ± 1). We compute the average over all pairs $\langle\delta\omega_{ij}\rangle$, and retain only those pairs with $\delta\omega_{ij}$ significantly larger than the average (for example, $\delta\omega \geq 2\langle\delta\omega_{ij}\rangle$). These "high-ranking" pairwise links not only contribute to high-$\delta\omega$ values of hinge region residues, but also connect them together to form a network via dynamically important pairwise interactions. This network plays a critical role in transmitting local perturbations due to ligand binding to other regions so that the functionally important hinge motion can be initiated.

Applications

The above general strategy to identify the network of dynamically important residues for a specific mechanical function is applied to study the open/closed transition in a number of DNA/RNA polymerases. The specific structures used here are listed in Table 2. For each structure we test for correlation between S and $\delta\omega$ values to find regions that are linked dynamically in the open/closed transition.

Acknowledgments

We thank Ruxandra Dima for useful discussions. This work was supported by a grant to D.T. from the National Science Foundation through CHE-02-09340. It was also supported by the National Institutes of Health.

Received: October 25, 2004
Revised: December 20, 2004
Accepted: January 4, 2005
Published: April 11, 2005

References

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. *25*, 3389–3402.

Atilgan, A.R., Durell, S.R., Jernigan, R.L., Demirel, M.C., Keskin, O.,

and Bahar, I. (2001). Anisotropy of fluctuation dynamics of proteins with an elastic network model. Biophys. J. *80*, 505–515.

Bahar, I., Erman, B., Jernigan, R.L., Atilgan, A.R., and Covell, D.G. (1999). Collective motions in HIV-1 reverse transcriptase: examination of flexibility and enzyme function. J. Mol. Biol. *285*, 1023–1037.

Brooks, B., and Karplus, M. (1985). Normal modes for specific motions of macromolecules: application to the hinge-bending mode of lysozyme. Proc. Natl. Acad. Sci. USA *82*, 4995–4999.

Chacon, P., Tama, F., and Wriggers, W. (2003). Mega-Dalton biomolecular motion captured from electron microscopy reconstructions. J. Mol. Biol. *326*, 485–492.

Delarue, M., and Sanejouand, Y.H. (2002). Simplified normal mode analysis of conformational transitions in DNA-dependent polymerases: the elastic network model. J. Mol. Biol. *320*, 1011–1024.

Delarue, M., Poch, O., Tordo, N., Moras, D., and Argos, P. (1990). An attempt to unify the structure of polymerases. Protein Eng. *3*, 461–467.

Doublie, S., Sawaya, M.R., and Ellenberger, T. (1999). An open and closed case for all polymerases. Structure *7*, R31–R35.

Huang, H., Chopra, R., Verdine, G.L., and Harrison, S.C. (1998). Structure of a covalently trapped catalytic complex of HIV-1 reverse transcriptase: implications for drug resistance. Science *282*, 1669–1675.

Jeruzalmi, D., and Steitz, T.A. (1998). Structure of T7 RNA polymerase complexed to the transcriptional inhibitor T7 lysozyme. EMBO J. *17*, 4101–4113.

Kim, B., Hathaway, T.R., and Loeb, L.A. (1998). Fidelity of mutant HIV-1 reverse transcriptases: interaction with the single-stranded template influences the accuracy of DNA synthesis. Biochemistry *37*, 5831–5839.

Kundu, S., and Jernigan, R.L. (2004). Molecular mechanism of domain swapping in proteins: an analysis of slower motions. Biophys. J. *86*, 3846–3854.

Li, Y., Korolev, S., and Waksman, G. (1998). Crystal structures of open and closed forms of binary and ternary complexes of the large fragment of Thermus aquaticus DNA polymerase I: structural basis for nucleotide incorporation. EMBO J. *17*, 7514–7525.

Li, S.X., Vaccaro, J.A., and Sweasy, J.B. (1999). Involvement of phenylalanine 272 of DNA polymerase beta in discriminating between correct and incorrect deoxynucleoside triphosphates. Biochemistry *38*, 4800–4808.

Ollis, D.L., Brick, P., Hamlin, R., Xuong, N.G., and Steitz, T.A. (1985). Structure of large fragment of Escherichia coli DNA polymerase I complexed with dTMP. Nature *313*, 762–766.

Opresko, P.L., Sweasy, J.B., and Eckert, K.A. (1998). The mutator form of polymerase beta with amino acid substitution at tyrosine 265 in the hinge region displays an increase in both base substitution and frame shift errors. Biochemistry *37*, 2111–2119.

Patel, P.H., Kawate, H., Adman, E., Ashbach, M., and Loeb, L.A. (2001). A single highly mutable catalytic site amino acid is critical for DNA polymerase fidelity. J. Biol. Chem. *276*, 5044–5051.

Pelletier, H., Sawaya, M.R., Kumar, A., Wilson, S.H., and Kraut, J. (1994). Structures of ternary complexes of rat DNA polymerase beta, a DNA template-primer, and ddCTP. Science *264*, 1891–1903.

Pelletier, H., Sawaya, M.R., Wolfle, W., Wilson, S.H., and Kraut, J. (1996). Crystal structures of human DNA polymerase β complexed with DNA: implications for catalytic mechanism, processivity, and fidelity. Biochemistry *35*, 12742–12761.

Sarafianos, S.G., Clark, A.D., Jr., Das, K., Tuske, S., Birktoft, J.J., Ilankumaran, I., Ramesha, A.R., Sayer, J.M., Jerina, D.M., Boyer, P.L., et al. (2002). Structures of HIV-1 reverse transcriptase with pre- and post-translocation Aztmp-terminated DNA. EMBO J. *21*, 6614–6624.

Steitz, T.A. (1999). DNA polymerases: structural diversity and common mechanisms. J. Biol. Chem. *274*, 17395–17398.

Sweasy, J.B. (2003). Fidelity mechanisms of DNA polymerase beta. Prog. Nucleic Acid Res. Mol. Biol. *73*, 137–169.

Tama, F., and Sanejouand, Y.H. (2001). Conformational change of proteins arising from normal mode calculations. Protein Eng. *14*, 1–6.

Temiz, N.A., and Bahar, I. (2002). Inhibitor binding alters the directions of domain motions in HIV-1 reverse transcriptase. Proteins *49*, 61–70.

Tirion, M.M. (1996). Large amplitude elastic motions in proteins from a single-parameter, atomic analysis. Phys. Rev. Lett. *77*, 1905–1908.

Van Wynsberghe, A., Li, G., and Cui, Q. (2004). Normal-mode analysis suggests protein flexibility modulation throughout RNA polymerase's functional cycle. Biochemistry *43*, 13083–13096.

Zheng, W., and Doniach, S. (2003). A comparative study of motor-protein motions by using a simple elastic-network model. Proc. Natl. Acad. Sci. USA *100*, 13253–13258.