ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/8240296

Normal-Mode Analysis Suggests Protein Flexibility Modulation throughout RNA Polymerase's Functional Cycle †

ARTICLE *in* BIOCHEMISTRY · NOVEMBER 2004

Impact Factor: 3.02 · DOI: 10.1021/bi049738+ · Source: PubMed

CITATIONS
49

READS
76

3 AUTHORS, INCLUDING:

Qiang Cui
University of Wisconsin–Madison
200 PUBLICATIONS **13,527** CITATIONS

SEE PROFILE

Available from: Qiang Cui
Retrieved on: 02 February 2016

Normal-Mode Analysis Suggests Protein Flexibility Modulation throughout RNA Polymerase's Functional Cycle†

Adam Van Wynsberghe,*,‡ Guohui Li,§ and Qiang Cui*,‡,§

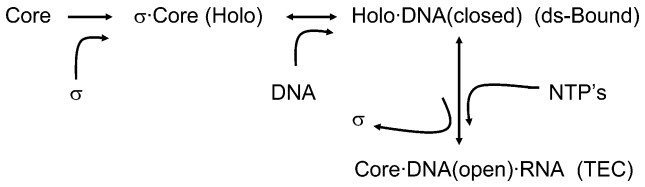
Graduate Program in Biophysics and Department of Chemistry and Theoretical Chemistry Institute, University of Wisconsin, 1101 University Avenue, Madison, Wisconsin 53706

Received February 4, 2004; Revised Manuscript Received July 30, 2004

ABSTRACT: To explore the domain-scale flexibility of bacterial RNA polymerase (RNAP) throughout its functional cycle, block normal-mode analyses (BNM) were performed on several important functional states, including the holoenzyme, the core complex, a model of RNAP bound to primarily duplex DNA, and a model of the ternary elongation complex. The calculations utilized a molecular mechanics (MM) force field with physical interactions; this is made possible by the use of BNM and the implementation of a sparse-matrix diagonalization routine. The use of homology models necessitated the MM force field rather than the simpler elastic network model (ENM). From the MM/BNM, we have systematically and semiquantitatively calculated the atomic fluctuations in the four functional states without bias due to crystal packing or other artifactual forces. We have observed that both α subunits and the ω subunit are rigid, in line with their roles as structural motifs that are not mechanistically involved in RNAP's functional cycle. It has been observed that the β subunit has two highly mobile domains; these are commonly known as the β1 and β2 domains. Our calculations suggest that the flexibility of these domains is modulated throughout the functional cycle and that they move entirely independently of each other unless DNA is bound. From an energetic perspective, we have shown the β2 domain can flex into and out of the cleft, forming interactions with DNA in the TEC as has been previously proposed. Our calculations also confirm that the β' subunit's likely flexibility into and out of the DNA binding cleft is energetically allowed. These two observations validate that both of the RNAP crab claw's pincers are mobile, as both β and β' have substantial flexibility.

From a biophysical perspective, the RNAP[1] system is a molecular machine. The intrigue of this machine lies in the number and variety of tasks that it must accomplish. Although the functional cycle (see Scheme 1) is often presented in the context of three simple steps, initiation, elongation, and termination, this hides the very intricate mechanical transitions that make up each of these stages. All RNAPs must be able to bind nonspecifically to dsDNA and slide along it in a one-dimensional search (*1*), bind specifically to double-stranded promoter DNA, melt the promoter, bind rNTPs complementary to the single-stranded DNA template, catalyze phosphodiester bond formation, elongate the nascent RNA chain processively, and then terminate when the RNA strand has been completed (*2*). In addition, it must recognize regulatory signals in the form of transcription factors, pausing sites in the DNA, and termination factors.

Scheme 1



The prokaryotic core RNA polymerase enzyme is a heteropentamer consisting of a β subunit, a β' subunit, and an ω subunit along with two α subunits, αI and αII (Figure 1). Only after the core (Figure 1a) associates with one of the σ transcription factors can it bind and melt dsDNA (*3*). With the σ subunit bound, the protein assembly is known as the holoenzyme (Figure 1b). Atomic-resolution crystal structures of this important functional state were determined in 2002 (*4*, *5*). Although their C-terminal domains (αCTDs) also interact with elements of the upstream DNA (*6*), the primary function of the two α subunits is to form a structural base on which the more functionally relevant β and β' subunits assemble. The ω subunit is thought to assist with this folding process (*7*). β and β' form a 70 Å deep and 25 Å wide channel with the active site at the base. DNA is

† This work was supported by Q.C.'s startup funds from the College of Letters and Science and the Department of Chemistry. A.V.W. was supported by a University of Wisconsin Graduate School Prize Fellowship from the Wisconsin Alumni Research Foundation and an NSF predoctoral fellowship.

‡ Graduate Program in Biophysics.

§ Department of Chemistry and Theoretical Chemistry Institute.

[1] Abbreviations: RNAP, RNA polymerase; BNM, block normal-mode analyses; MM, molecular mechanics; ENM, elastic network model; αCTD, α-carboxy-terminal domain; Taq, *Thermus aquaticus*; NMA, normal-mode analysis; rmsf, root-mean-square fluctuation; grms, gradient root-mean-square; C_α, α carbon; rmsd, root-mean-square deviation; TEC, ternary elongation complex; vdW, van der Waals; yRNAP II, yeast RNA polymerase II; SDSL, site-directed spin labeling; EPR, electron paramagnetic resonance; ds-bound, double-stranded DNA bound; FRET, fluorescence resonance energy transfer.

10.1021/bi049738+ CCC: $27.50 © xxxx American Chemical Society
Published on Web 00/00/0000 PAGE EST: 13.4

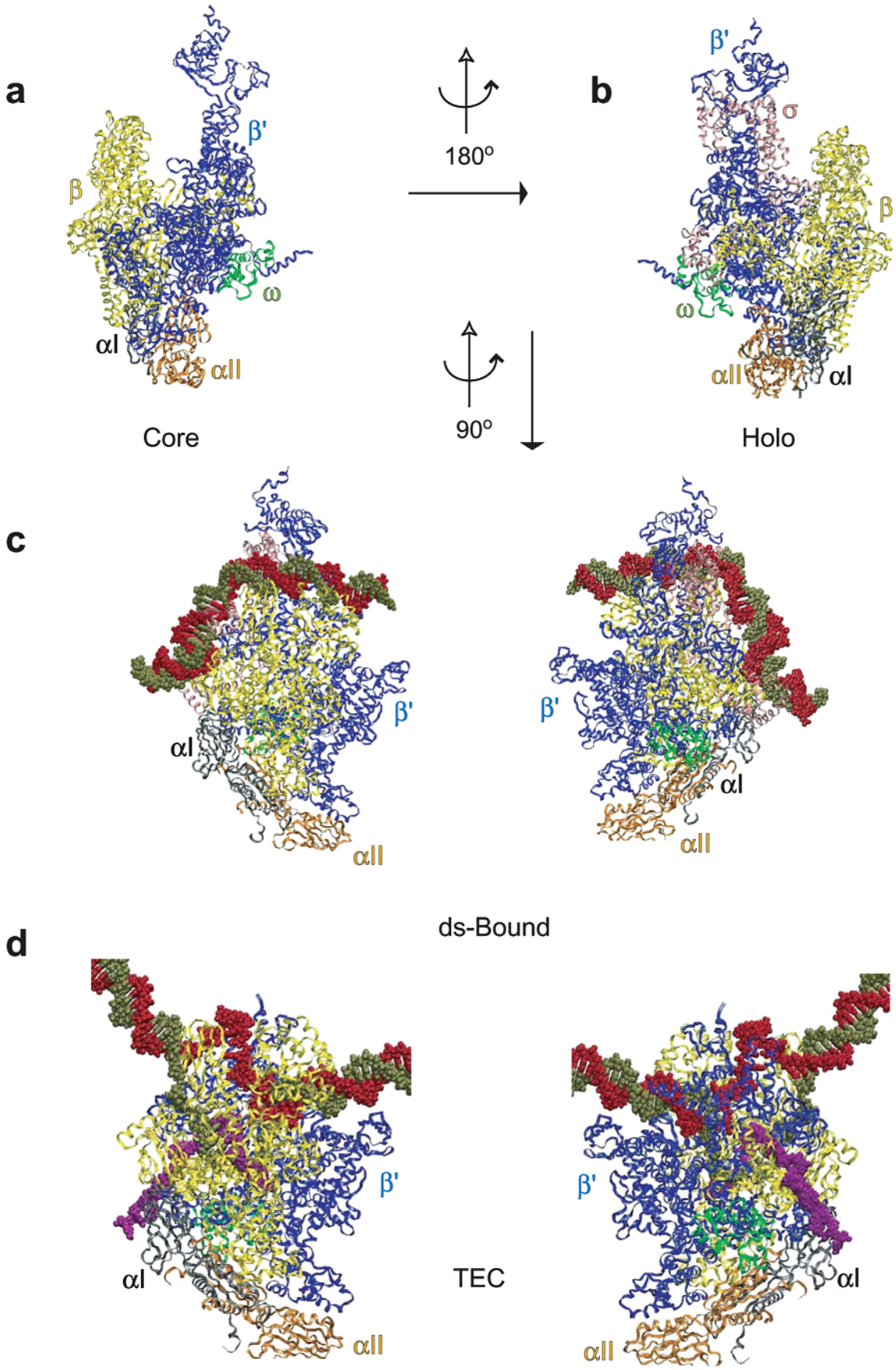


FIGURE 1: Ribbon representations of the four models of the functional states studied in this work. Individual subunits are colored differently: black for αI, orange for αII, yellow for β, blue for β', green for ω, and pink for σ. (a) The core model. (b) The holo model. In both panels c (the ds-bound model) and d (the TEC model), the nucleic acid atoms are shown as vdW spheres. This figure was created with VMD (*52*).

bound in this large main channel. The active site holds two catalytically essential Mg^{2+} ions, and the chemical mechanism of phosphodiester bond formation is thought to follow the two-metal model as proposed by Steitz (*8*, *9*). Along with the aforementioned interactions of the α subunits, two regions of the σ subunit help to identify promoter sequences. σ region 4.2 is known to recognize a consensus sequence located 35 bp upstream of the transcription start site (−35 region) (*10*), while σ region 2.4 has been implicated in binding to a sequence 10 bp upstream of the start site (−10 region or "Pribnow Box") (*11*).

The overall size of the bacterial RNAP is approximately 150 Å × 115 Å × 110 Å and roughly resembles a crab claw with the β and β' subunits each making up one of the "pincers" or "jaws" (*12*) (See Figure 1). From crystallographic structures (*12*−*14*), it has been observed that the jaws can adopt distinct conformations that open or close the main DNA binding channel. An open state and a closed state have been observed for both the eukaryotic (*13*) and prokaryotic (*12*, *14*) enzymes. In the yeast RNAP, two different crystal forms of the core enzyme exhibited a structural displacement of a region named the "clamp" (*13*). The relative position of this region, composed mostly of subunit Rpb1 [analogous to the β' subunit (*15*)], was also different in the free and the DNA- and RNA-bound elongation complex (*16*). In the unbound complexes, the clamp exhibited a hingelike bending motion with a maximal displacement of 14 Å, while in the elongation complex, the clamp's motion included a 30° rotation with a maximum displacement from the core structure of 30 Å. This large deviation results in the clamp swinging over the channel and trapping the DNA (*16*). Comparison of the crystal structure of the *Thermus aquaticus* (Taq) RNAP (*12*) with an electron diffraction map of the *Escherichia coli* RNAP reveals a hinging motion of the β' jaw that resulted in a 25 Å displacement of its tip (*14*). This motion in the bacterial system also resulted in the closing of the DNA binding channel by the jaws (β' and β for bacteria; Rpb1 and Rpb2 for yeast).

The experimental techniques described above as well as others have led to qualitative models describing how portions of the bacterial RNAP may move. However, none of these experiments have been a systematic search of the conformational freedom of RNAP, and the most telling observations have relied on potentially fortuitous forces from crystal packing interactions. Crystal structures can only provide a static look at individual snapshots and do not discern between structures stabilized internally or from the crystalline environment. In this work, we use simple but atomistic and energetically motivated computational techniques to explore the range of motion for all of the subunits of the bacterial RNAP in several important functional states. Our approach removes any artifacts from crystalline environments and provides a systematic and semiquantitative description of the conformational flexibility of the studied system.

To address the proposed questions, we have employed normal-mode analysis (NMA). In this technique, one approximates the potential energy of the system harmonically around a single local minimum. With this approximation, the equations of motion for the entire system may be solved analytically (*17*−*20*). The large-scale soft motions of large portions of the protein moving relative to each can be successfully captured with NMA. In the particular case of the bacterial RNAP (greater than 36 000 atoms with the σ subunit and DNA bound), its large size necessitates further simplifications.

Previous computational work on the conformational mobility of RNAP used a simplified potential energy function, the elastic network model (ENM) (*21*), to study the possibility of jaw (β and β' subunit) movement into and out of the DNA binding cleft (*22*). The ENM employs a simple phenomenological potential using two arbitrary parameters. It treats a protein system as a C_α trace and connects any two C_α atoms separated by less than a given cutoff radius (r_{cut}) with a Hookean spring with a force constant (k). After generation of this potential, the regular procedure for normal-mode analysis is carried out (*21*). The consequence of using such a potential is that ENM has no predictive power for the magnitude of the atomic fluctuations and must rely on comparison to crystal structure temperature factors to properly scale its results (*23*). This is obviously disadvantageous when dealing with homology models because they have no associated temperature factors. Using ENM, however, Delarue and Sanejouand showed that the bacterial RNAP, as well as many other DNA-dependent polymerases, could indeed flex into the DNA binding cleft. In their work, Delarue and Sanejouand were limited to only the holoenzyme and utilized a crystal structure (*5*) that was later shown to have been partially incorrect (*4*).

Our work extends and complements these previous studies while adding slight improvements. We have extended the scope of the normal-mode analyses of RNAP by studying several important functional states, and our improvements, although slight, come from using the more accurate holoenzyme crystal structure (*4*) and from using a molecular mechanics potential (MM) instead of the simpler ENM (see Experimental Procedures). While Delarue and Sanejouand mainly focused on the transition of the jaws between their distal and proximal conformations, this work focuses on the differences in flexibility between individual functional states.

As mentioned above, RNAP has a complex functional cycle as illustrated in Scheme 1. This is only a simple schematic, and only the four functional states that were studied are shown. Although there are only atomically detailed crystal structures for the core (*12*) and holo (*4*, *5*) forms, the other two may be reached through appropriate homology modeling (see Experimental Procedures for details). Through the refinement of crystal structure coordinates, one obtains temperature factors, also known as Debye−Waller or B factors (*24*), which are directly related to the flexibility of the protein. Although other factors may be calculated, protein flexibility is the main value calculated from NMA. Flexibilities computed from NMA generally match up well with those obtained through experiment (*25*). However, flexibilities calculated with NMA do not suffer from errors in the fitting of coordinates or artifactual bias from crystal packing interactions like those from crystallographic refinements. Fluctuations from NMA are also anisotropic, while only very high-resolution and well-refined crystal structures can have this property. More importantly with respect to this work, no experimental flexibilities can be determined from homology models, and thus, the flexibilities of these models can only be attained through computational methods.

EXPERIMENTAL PROCEDURES

Model Preparation. The holoenzyme model coordinates came directly from those contained within rnp_mol1_melt_dna.pdb as supplied to the authors by D. Vassylyev. This set of coordinates is a step in the refinement of deposited coordinates of PDB entry 1IW7. It should be noted that starting from these coordinates or one of the two sets in 1IW7 has no significant effect on the results of this analysis. BNM is a coarse-grained technique that is insensitive to small changes in atomic positions. The differences between the two sets of 1IW7 structures are small; the rmsd of the C_α atoms for any individual subunit is less than 1.0 Å, with the more rigid subunits (αI, αII, and ω; see Results and Discussion) having smaller rmsds than the more flexible subunits (β, β', and σ; see Results and Discussion). Disordered residues include both α C-terminal domains (residues 230−315), residues 1, 252−363, and 1506−1524 of the β' subunit, residues 1 and 97−99 of the ω subunit, and residues 1−73, 380−383, and 421−423 of the σ subunit. The core enzyme model was constructed from the holoenzyme model by simply deleting the σ subunit coordinates. In the double-stranded DNA-bound (ds-bound) model, a dsDNA segment was modeled into the active site DNA binding channel. It includes a large bend of the DNA to match available biochemical data that show the interactions between region σ 4.2 and the −35 sequence as well as region σ 2.4 and the −10 promoter DNA sequence (*26*). Three nucleotides are unpaired due to this distortion. This model is intended to represent an RNAP bound to a double-stranded DNA helix at the tip of the jaws. Since some bases are melted, it is inappropriate to refer to this as the canonical "closed" complex (*27*), but rather, it is a complex reached immediately before the template strand drops into the active site as described by Bushnell *et al*. (*28*) This model is included mainly to test how RNAP responds to a double helix bound at the tip of the jaws. These three models (core, holo, and ds-bound) are entirely the work of D. Vassylyev and co-workers (*4*). The ternary elongation complex was modeled from the deposited *Thermus thermophilus* RNAP coordinates by R. Landick in analogy to the yeast RNAPII TEC (yRNAPII TEC). Both the entirety of σ and some residues of the β' subunit (163−251 and 364−447) were deleted from the deposited coordinates. In addition, the clamp domain was moved (β residues 1058−1119 and β' residues 2−619 and 1436−1505) as defined by Murakami *et al*. (*5*). The flap domain (β residues 705−828) was adjusted to match the yeast RNAPII TEC (*16*) based on residues 734−762 with minor adjustments to leave the terminal residues of this region (705 and 828) in their original crystal structure positions. The flap tip (β residues 757−790) was replaced with the flap tip from the core *T. aquaticus* structure (*12*). Last, the F helix was adjusted to the yRNAPII TEC conformation to avoid any clash with the DNA template strand.

Minimization. All models were locally minimized using the adapted basis Newton−Raphson method both to remove local steric clashes and to achieve the local minimum as required by normal-mode analysis. Minimizations were accomplished using the computational package CHARMM (*29*) utilizing the CHARMM19 extended atom force field (*30*) supplemented to contain nucleic acid parameters from the CHARMM22 force field (*31*) as well as modified to use the EEF1 solvation model (*32*). All minimizations were carried out in cycles with a gradually decreasing positional harmonic constraint. This helped reduce very unfavorable steric contacts without perturbing the structure significantly. The final minimization was run with no positional constraint and terminated when the rms energy gradient reached a value of either 0.001 kcal mol^{-1} $Å^{-1}$ (ds-bound model) or 0.01 kcal mol^{-1} $Å^{-1}$.

Elastic Network Model Calculations. The anisotropic elastic network model (*33*) was implemented in the CHARMM framework and was used in all ENM calculations. ENM has been shown, in this work and elsewhere (*21*), with the proper scaling to reproduce the results of normal-mode analyses using MM. It also needs no minimization of the original crystal structure coordinates and is computationally inexpensive (∼4 h for 1000 modes). Several ENM calculations of the holoenzyme were conducted; each was started from a structure minimized to 0.01 kcal mol^{-1} $Å^{-1}$ grms using CHARMM to compare with MM/BNM results. For all calculations, the force constant equaled 0.95 kcal mol^{-1} $Å^{-2}$. The cutoff radius was varied for three different calculations; values of 10, 16, and 25 Å were used. For tests of convergence, different numbers of modes were used in the calculations. These tests used the 16 Å cutoff. After eigenvectors and eigenvalues were calculated, and rmsfs were calculated for all C_α atoms. These raw rmsfs were scaled to provide useful data; the scaling factor for each set was equal to the total sum of the MM/BNM fluctuations divided by the sum of the set's fluctuations.

Block Normal-Mode Calculations. In this work, we employ the block normal-mode (BNM) approach originally suggested by Tama *et al*. (*34*) and improved by Li and Cui. (*35*) Normal-mode calculations were performed with the same force field as the minimizations, again within the CHARMM suite. Each protein residue, nucleotide, and bound metal ion was considered a block. The core, holo, and ds-bound complexes contained three metal ions, two structural Zn^{2+} ions, and one active site Mg^{2+}, while the TEC contained only the two Zn^{2+} ions. The block normal-mode procedure was performed as discussed elsewhere (*35*). Briefly, the Hessian, the second-derivative matrix of the potential, was constructed using eq 1:

$$F_{ij} = \frac{\partial^2 V}{\partial q_i \, \partial q_j} \tag{1}$$

where F is the Hessian, V is the potential calculated from the force field, and the q_i values are mass-weighted displacement coordinates from the local energy minimum. This Hessian was projected onto the smaller space of rotational and translational degrees of freedom of every block in the direct manner described by Li and Cui (*35*). This "blocked" Hessian was then diagonalized using a sparse-matrix diagonalization routine implemented by Li and Cui (unpublished), and the resultant eigenvectors were transformed back from the blocked space to the all-atom space. This approximation of conventional NMA is modest if only the low-frequency modes are used (*34*, *35*). These eigenvectors were then used in the analysis of the fluctuations. All calculations were performed on desktop machines running the Linux operating system with 1 GB of RAM (only 300 MB was allocated to

BNM calculations) and 1.2 GHz AMD Athlon processors. The total time of the calculations (including minimization and normal-mode analysis) was 20−40 h of CPU time depending on the system.

Rmsf Calculations. The flexibility of the complex was described in terms of the rmsf of atoms. An analytical expression for the positional fluctuations of each atom of the system can be easily derived as well as easily calculated once the eigenvectors and eigenvalues of the Hessian are known. For a particular atomic displacement x_i, where i indexes over all degrees of freedom of the system, the mean-square fluctuation from the local energy minimum is given by eq 2.

$$\langle x_i^2 \rangle = \frac{k_B T}{m_i} \sum_j^{\text{modes}} \frac{U_{ij}^2}{\omega_j^2} \quad (2)$$

where m_i is the mass of the atom corresponding to the ith degree of freedom, k_B is Boltzmann's constant, T is the temperature, and U_{ij} is the ith component in the eigenvector with frequency ω_j. The sum only goes over relevant modes; modes corresponding to translation and rotation of the entire system are excluded as well as modes that have zero frequency. These zero-frequency modes come from the blocking scheme as any residues that only contain a single atom (in this case only metal ions are represented this way) have no rotational degrees of freedom.

RESULTS AND DISCUSSION

Comparison of Elastic Network Model and Block Normal-Mode Analyses

In this section, we first illustrate the differences of ENM and MM potentials in the context of normal-mode analysis as well as validate our use of BNM even with its computationally more demanding algorithm by describing several normal-mode analyses on the bacterial RNAP holoenzyme using both the ENM and MM/BNM. These calculations allowed us to investigate the influence of differing degrees of minimization (MM/BNM) and the effects of choosing different cutoff distances for virtual bonds (ENM). The comparison of ENM and MM/BNM is based mainly on the changes that occur in the root-mean-square fluctuations (rmsfs) upon changes in simulation parameters. Since the average displacement of each atom in this harmonic system is zero, the most straightforward way to discuss the inherent flexibility of a macromolecule in the normal-mode framework is through each atom's rmsf. Within the normal-mode framework, the rmsfs are analytically defined as a function of the eigenvectors and eigenvalues (see Experimental Procedures) and can easily be calculated for every atom in the system. These values give a rough description of the flexibility of the protein in the conformation that is being studied.

As stated above, BNM requires an initial minimization step before the normal-mode analysis is carried out. The effect of this minimization step has been investigated by minimizing to both a rms energy gradient (grms) of 0.01 and 0.001 kcal mol^{-1} Å^{-1} using the adapted basis Newton−Raphson method (*29*). These minimizations took 5.3 and 11.7 h on a single Linux machine (see Experimental Procedures), respectively, and resulted in structures that had 1.34 and 1.99 Å rmsds (C_α only), respectively, from the original crystal structure. Block normal-mode analyses were then performed on both of these minimized structures, and rmsfs were determined. The difference in the rmsf using these two sets of coordinates has a rms value of 0.261 Å, and the individual differences are shown in Figure 2a. This figure shows that the deviation is dominated by a few residues (∼230 and ∼1800−1850) whose deviations are very large. If 1% of the most deviated data is ignored, the rmsd drops to only 0.102 Å. The effect of minimization is fairly constant across the residue range. As shown in Figure 2a, the baseline deviation is a small positive number that results from a relaxation to more optimal (for the force field used) atomic positions with the longer minimization. The sharp peaks at residues 231 and 1827 occur at the N-terminus of the αII subunit and the C-terminus of the first fragment of β'. As discussed below, the fluctuations calculated near chain breaks and termini can give unphysical results, and as a result, these data are ignored. On the basis of the comparison in Figure 2a, we conclude that the minimization needed to perform BNM is predictable and rational and does not result in a large change in the calculated flexibility. For further verification, ENM calculations at both the crystal structure and minimized structure were also compared. No significant differences were observed (data not shown).

The dependence of ENM-calculated flexibilities on the radial cutoff distance was also investigated. Three ENM calculations were carried out with different cutoffs: 10, 16, and 25 Å. In Figure 2b, the difference [r_{cut} (10 Å) − r_{cut} (16 Å)] in the calculated fluctuations as a function of the residue index is shown. These data are similar to the difference data between 10 and 25 Å as well as between 16 and 25 Å cutoffs. Comparable to the rmsd between the aforementioned minimizations, the rms difference between these sets of data is 0.277 Å. Another similarity is that removal of 1% of the most deviated data reduces this rmsd to 0.106 Å. There is, however, a striking difference between panels a and b of Figure 2. The minimization differences for BNM in Figure 2a are for the most part skewed to positive numbers (the average deviation equals 0.090 Å) because the long minimization relaxes the structure and decreases short-range repulsions while also creating a more compact structure. This is consistent for most residues, and there are very few large changes in flexibility. For the cutoff ENM difference data, however, there is no systematic skewing of the data, and the average deviation is nearly zero (0.001 Å). A large cutoff merely introduces more connected springs as more C_α atoms are included in the interaction sphere. Larger cutoffs thus result in damped fluctuations in the raw data. However, to achieve reasonable fluctuations that match experimental data, the raw rmsf data are scaled. In this work, we have scaled the rmsfs from ENM so that the sum of the fluctuations matches the sum of the BNM fluctuations. This sum is dominated by core residues of the RNAP, those that give rise to the "baseline" fluctuation. A residue whose raw calculated fluctuation does not change proportionally to these baseline residues gives rise to a large difference in the *scaled* fluctuations, which are shown in Figure 2b. The residues that have this property are those that are on the protein surface and are far from densely packed regions. Figure 3 illustrates some of these residues. Increasing the size of the

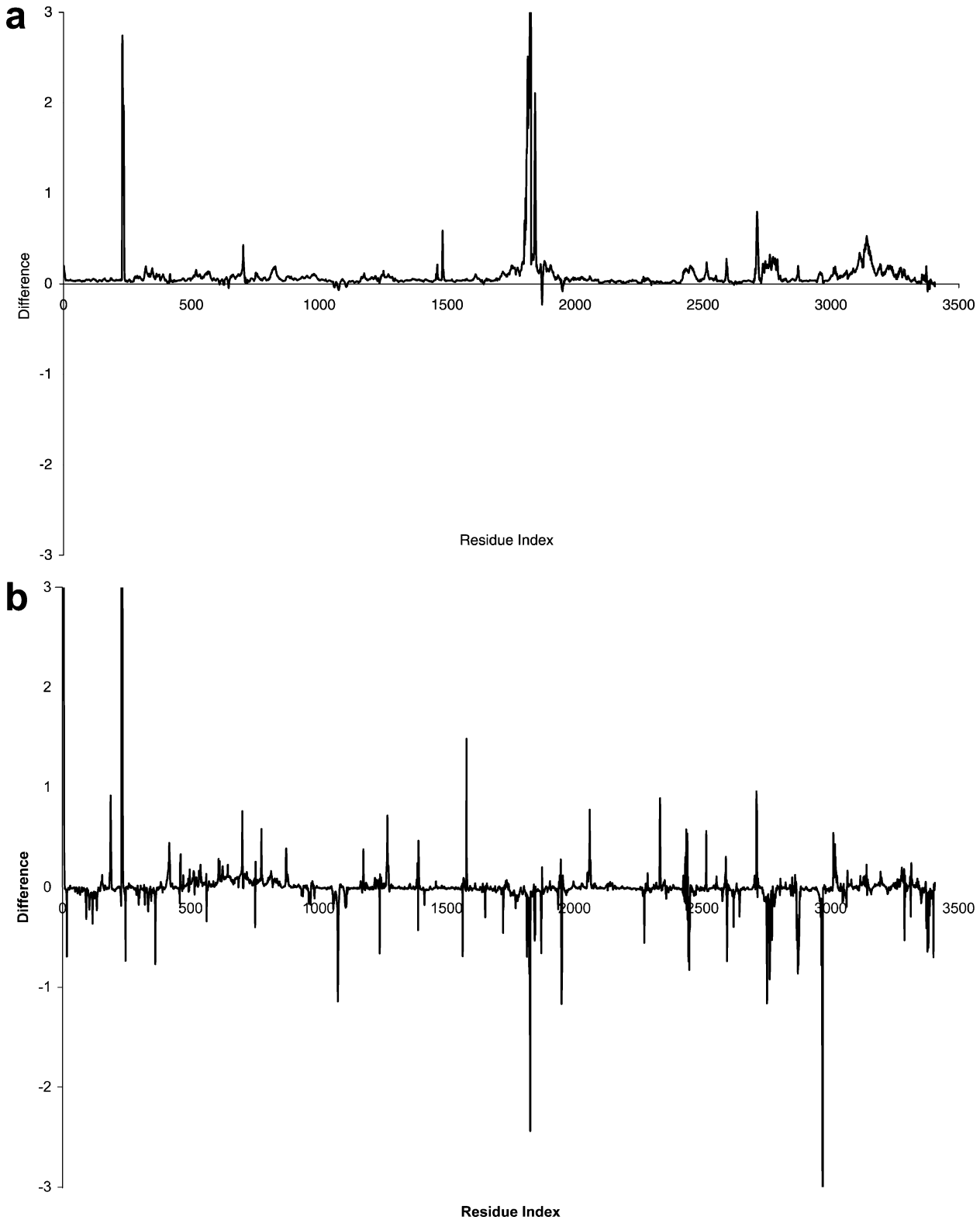


FIGURE 2: Differences in root-mean-square fluctuations (rmsfs) calculated from (a) MM/BNM analyses starting from different minimized structures and (b) ENM calculations with different cutoff radii. The rmsfs for the longer minimization are subtracted from the shorter minimization rmsfs, and the rmsfs with an r_{cut} of 16 Å are subtracted from those with an r_{cut} of 10 Å. The residue index spans all subunits in the system: αI residues 1−229 map to 1−229, αII residues 1−229 to 230−458, β residues 1−1119 to 459−1577, β' residues 2−251 to 1578−1827, β' residues 364−1505 to 1828−2969, ω residues 2−96 to 2970−3064, σ residues 74−379 to 3065−3370, and σ residues 384−420 to 3371−3407.

interaction sphere for these residues does not have the same effect as increasing the size of the interaction sphere for core residues. Core residues are surrounded by other residues on all sides and acquire large numbers of interactions with an increase in the cutoff radius. Surface residues, especially those on protrusions, increase their level of interaction very slowly with an increasing cutoff radius. Neither the magnitude nor the sign of the deviation is easily predicted before the calculation has been performed. This results in the very spiky trace of Figure 2b. These unpredictable changes in the

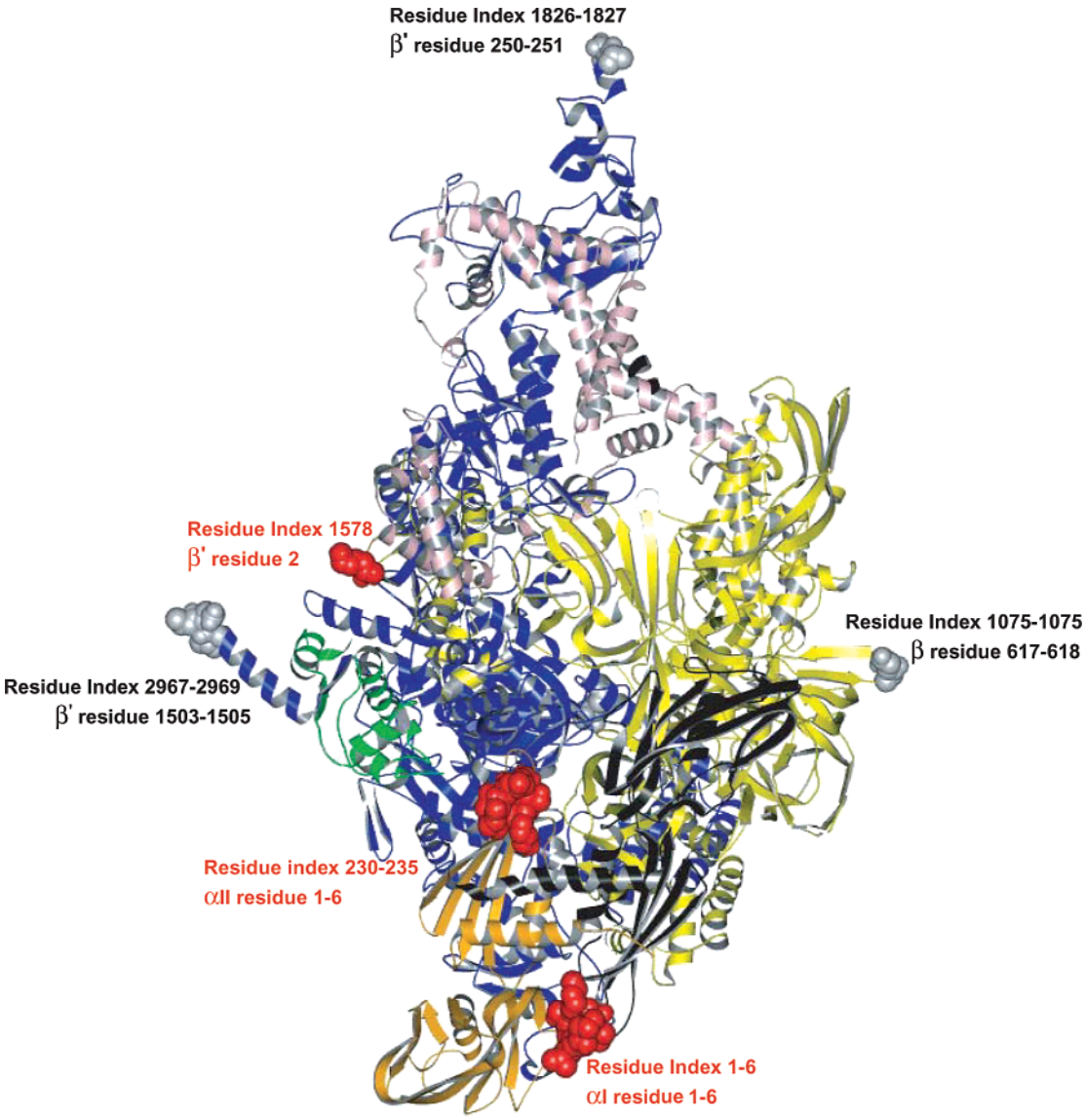


FIGURE 3: RNAP holoenzyme and the residues that correspond to several of the highest peaks in Figure 3b as van der Waal spheres. Red indicates a positive deviation, while gray indicates a negative deviation. This figure was created using MOLSCRIPT (*53*).

fluctuations of the surface residues with changes in the arbitrary parameter of cutoff distance is troublesome and, when compared with the relatively benign perturbation of minimization, is one of the reasons why the MM/BNM analysis method is preferred over ENM in the current study.

These comparisons strengthen our assertion that the MM/BNM approach is less arbitrary and more realistic than ENM even though the overall rmsfs obtained with both methods are very similar. For the current purpose of comparing the fluctuations of RNAP in different functional states, the improved realism as well as the nonarbitrary method that MM/BNM provides is advantageous. This is especially true for this problem as homology models of the RNAP are used, and they of course have no associated temperature factors to which ENM raw data may be scaled.

Comparison of Fluctuations from RNAP Functional States

Although the rmsf by itself gives no directional information, one may visualize the soft motions that give rise to the spatial fluctuations by examining the eigenvectors. An individual eigenvector defines the direction of motion of each atom in the system, and along with its corresponding eigenvalue determines the approximate thermal amplitude of the motion. We have calculated the rmsfs for all atoms (only the α carbon fluctuations are shown here) and examined the character of the eigenvectors for four important functional states in transcription: the core enzyme, the holoenzyme, a model with double-stranded DNA bound (ds-bound), and a model of the ternary elongation complex (TEC). In the following, we will discuss the flexibility of each subunit individually with respect to each of the different models.

α *Subunits.* As mentioned above, the α subunits are expected to have very little mobility. Figure 4 shows the rmsf data for both the αI and αII subunits. These data are typical of fluctuation data in that they have well-defined peaks superimposed on a variable baseline in the flexibility of the protein. The peaks are most often representative of unstructured regions connecting secondary structure; the strong hydrogen bonding in both α helices and *β* sheets tends to stabilize positional fluctuations. In this case, the majority of the peaks, especially in the αII subunit, correspond to

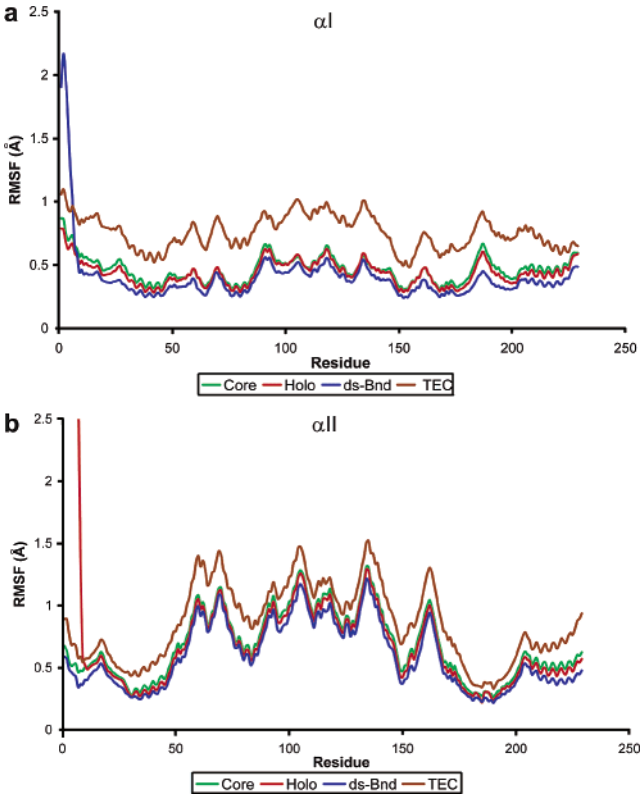


FIGURE 4: Root-mean-square fluctuation (rmsf) from MM/BNM for both the (a) αI and (b) αII subunits. Only the α carbon fluctuations are shown.

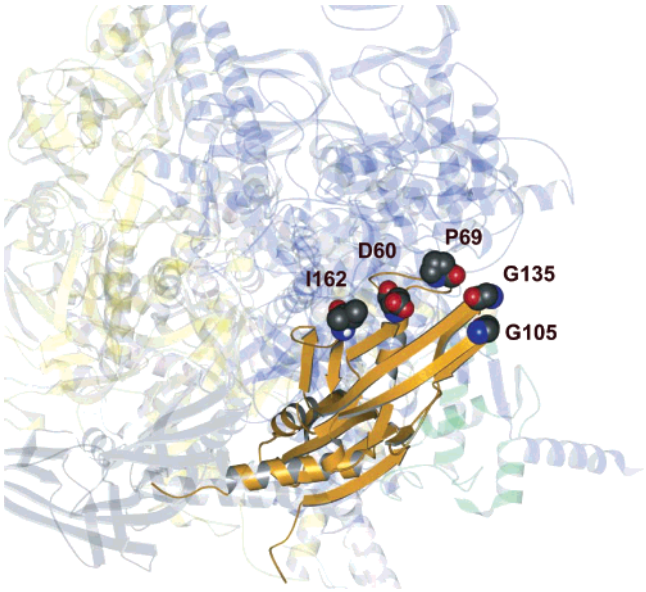


FIGURE 5: View of the αII subunit and its peaks in fluctuation. The residues shown as vdW spheres correspond to the peaks in Figure 4b. This figure was created using MOLSCRIPT (*53*).

turns in β sheets that point away from the protein into the solvent. Figure 5 shows the residues corresponding to the peaks in Figure 4b in van der Waals (vdW) radii. The variable baseline comes from different degrees of tertiary folding contacts and whether the atoms are buried in the core of the protein or near the surface. One notes that for the ds-bound and holo complexes in αII and for the ds-bound complex in αI the N-terminus shows a much larger flexibility than the rest of the subunit. This illustrates a weakness of normal-mode analysis. At termini (or chain breaks as will be shown below), there are fewer of the many packing interactions that stabilize protein structure. Consequently, the force constants that hold the protein together are very small at these points, and unphysical degrees of flexibility can be calculated. For termini (and chain breaks) that are relatively far from the core of the protein, any quantitative interpretation of the data must be ignored on physical grounds. These large calculated fluctuations, however, qualitatively suggest that these regions have high intrinsic flexibility.

Also noteworthy from these plots is the very small difference in flexibility among the four models. The TEC is the only model that could be considered different from the others. Although the ds-bound, holo, and core models differ in their DNA binding or σ binding states, they have the exact same starting coordinates for the α, β, β', and ω subunits. The TEC, on the other hand, comes from the deposited *T. thermophilus* RNAP coordinates but with several regions modified (see Experimental Procedures for details). These rather small changes are enough to change the local minimum reached in the minimization and to slightly perturb the amplitude of the fluctuations. The difference is small, however, and the semiquantitative trend of the α fluctuations in the other three models is reproduced in the TEC. For these types of calculations, it is the trend in the fluctuations that is important. The magnitudes are very sensitive to the calculated frequency (see eq 2), and the most meaningful interpretation of the data is obtained by examining the trends and relative flexibilities among different regions of the protein.

The small variation between the different functional states is consistent with the α subunits acting as the scaffold on which the active site formed by β and β' is docked. β and β', the subunits that make up the jaws, are what must actively change in the functional cycle; they bind the DNA and the σ subunit, and they are where catalysis takes place. With the exception of the αCTDs, which are not present in this analysis because they were not resolved in the crystal structure, the α subunits merely play a structural role in RNAP's catalytic cycle and are static; i.e., their mobility does not change between the states that were studied.

ω Subunit. The rmsfs for the ω subunit are shown in Figure 6a. When the flexibility is compared to those of the other subunits, the rmsf data show that ω is compact and does not have a wide range of motion. The small variability between the functional states also suggests that it does not play a mechanistic role in the transcriptional cycle. The low flexibility and the small variation between states are consistent with the suggestion that ω's main function is structural. It stays rigid, perhaps to simply lend stability to the overall structure. This finding helps support the conjecture that the stability of the ω subunit is helpful in assisting the correct folding of the jaws (*7*).

σ Subunit. The σ subunit is only present in two models studied in this work, the holo and ds-bound complexes. From the fluctuation data shown in Figure 6b, it is obvious that the trends in the flexibilities of σ do not change as the shapes of both the holo and ds-bound fluctuation curves are extremely similar. The baselines are substantially different, however. Because σ makes most of its contacts to the core protein on the β' subunit, the overall magnitude of its fluctuations (its baseline) is determined by the mobility of β'. β' is stabilized by contacts with the DNA in the ds-bound

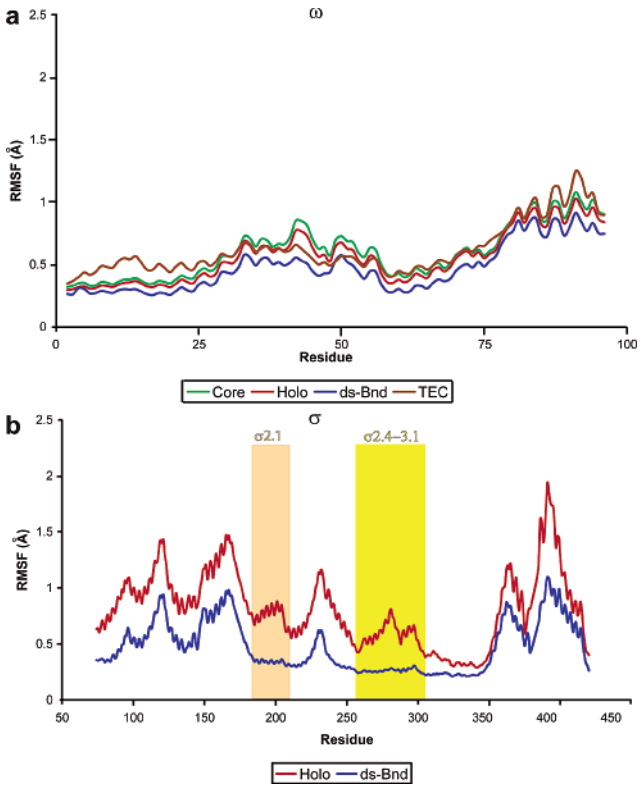


FIGURE 6: MM/BNM rmsfs for subunits ω (a) and σ (b). Region 2.1 of the σ subunit is highlighted in tan, and regions 2.4−3.1 are highlighted in yellow. The ω subunit's flexibility does not change over the states studied and suggests that its role is structural. The σ subunit has a change in its baseline between the holo and ds-bound states because the β' subunit is less mobile in the ds-bound state.

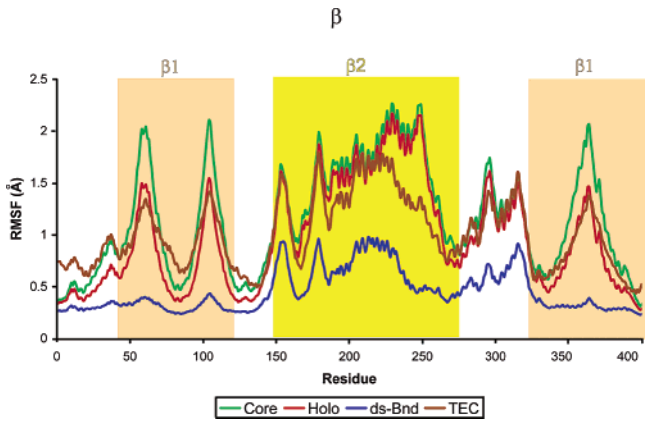


FIGURE 7: MM/BNM rmsfs for the first 400 residues of the β subunit. The residues in the β1 domain are highlighted in tan, while the β2 domain is highlighted in yellow. The β1 domain exhibits a cycle of flexibility within the studied functional states. The β2 domain is unaffected by the addition of the σ subunit (holo state) but is positionally stabilized in the ds-bound complex and in the TEC complex.

complex (see below) and has subsequently lower mobility. This in turn decreases the measured rmsf for every residue in the σ subunit.

There are only two features in the rmsf curves that differ between the holo and ds-bound states. There is a small quenching of the flexibility in the region of σ that interacts with the coiled coil region of β' (residues 540−585). These residues (σ residues 180−210, highlighted in tan in Figure 6b) are in region 2.1 of σ and show flexibility above the baseline in the holo state, but are consistent with the baseline in the ds-bound state. This can be rationalized by noting that the C-terminal portion of these residues makes direct contact with the bound DNA in the ds-bound state. In addition, residues 250 to approximately 300 (σ regions 2.4−3.1 highlighted in yellow in Figure 6b) have baseline flexibility in the ds-bound state but exhibit broad peaks of enhanced flexibility in the holo form. Region 2.4 has been implicated in recognizing the −10 promoter element, while the remainder of these residues span a helix that reaches across the DNA binding channel from the clamp to the β1 domain of the β subunit (also known as the "protrusion"). All of these residues make direct contact with the bound DNA, and thus, their number of atomic fluctuations is decreased.

β Subunit. The β subunit is one of the two subunits that make up the DNA binding channel and forms the active site. Interestingly, we see a cycle of flexibility in both the β1 (the protrusion) and β2 (the "lobe") domains of this subunit. The flexibilities of these domains are plotted in Figure 7 and mapped onto the four structures in Figure 8. The three peaks in the highlighted tan regions labeled β1 in Figure 7 correspond to two β sheets and an α helix that fold together to form the upstream portion of the jaw on the β side of the DNA binding channel. In the core complex (in the absence of σ or DNA), the fluctuation of β1 is larger than in the other models. The σ subunit is present in the holo complex (it is semitransparent in Figure 8b,c); it stabilizes the position of the β1 subunit through a hairpin turn (σ residues 289−294) that spans the binding cleft at its base. The fluctuations in the next functional state in the sequence of transcription, the ds-bound state (Figure 8c), are effectively negated. Most of the protein has very little fluctuation in any state; note the green coloring for the interior residues (all semitransparent) in all panels of Figure 8. When the DNA is bound in the ds-bound state, the β and β' jaws show the same flexibility as the interior, less than 0.5 Å. This suggests that the interactions between the β and β' jaws and the double-helical DNA bound in the cleft are favorable and that the DNA in this state holds the two jaws almost entirely rigid. However, as the RNAP moves from the ds-bound state to the TEC, the flexibility returns. In fact, the calculated magnitude of the fluctuation in the TEC is approximately equal to that of the holo state. This is intriguing as the interactions that stabilize the holo and TEC states relative to the core state are different. As mentioned above, β1 of the holo complex is stabilized by the σ subunit, but in this model of the TEC, it is missing. β1 of the TEC is stabilized by the nontemplate strand of the now melted DNA, shown partially in vdW radii in Figure 8d. These data taken as a whole suggest that the RNAP has significant conformational freedom in the β1 domain before it binds DNA, loses that flexibility upon initial binding, but then regains it again upon DNA melting and σ release.

Another interesting region of the β subunit is the β2 domain. This region (residues 150−330) is primarily made up of α helices that pack together to form the downstream side of the β jaw. There are several interesting observations to be made from the fluctuation data shown in Figures 7 and 8. First, the fluctuation for the core and holo states is the same throughout this entire domain. Unlike the β1 domain, the σ subunit makes no direct contact with this

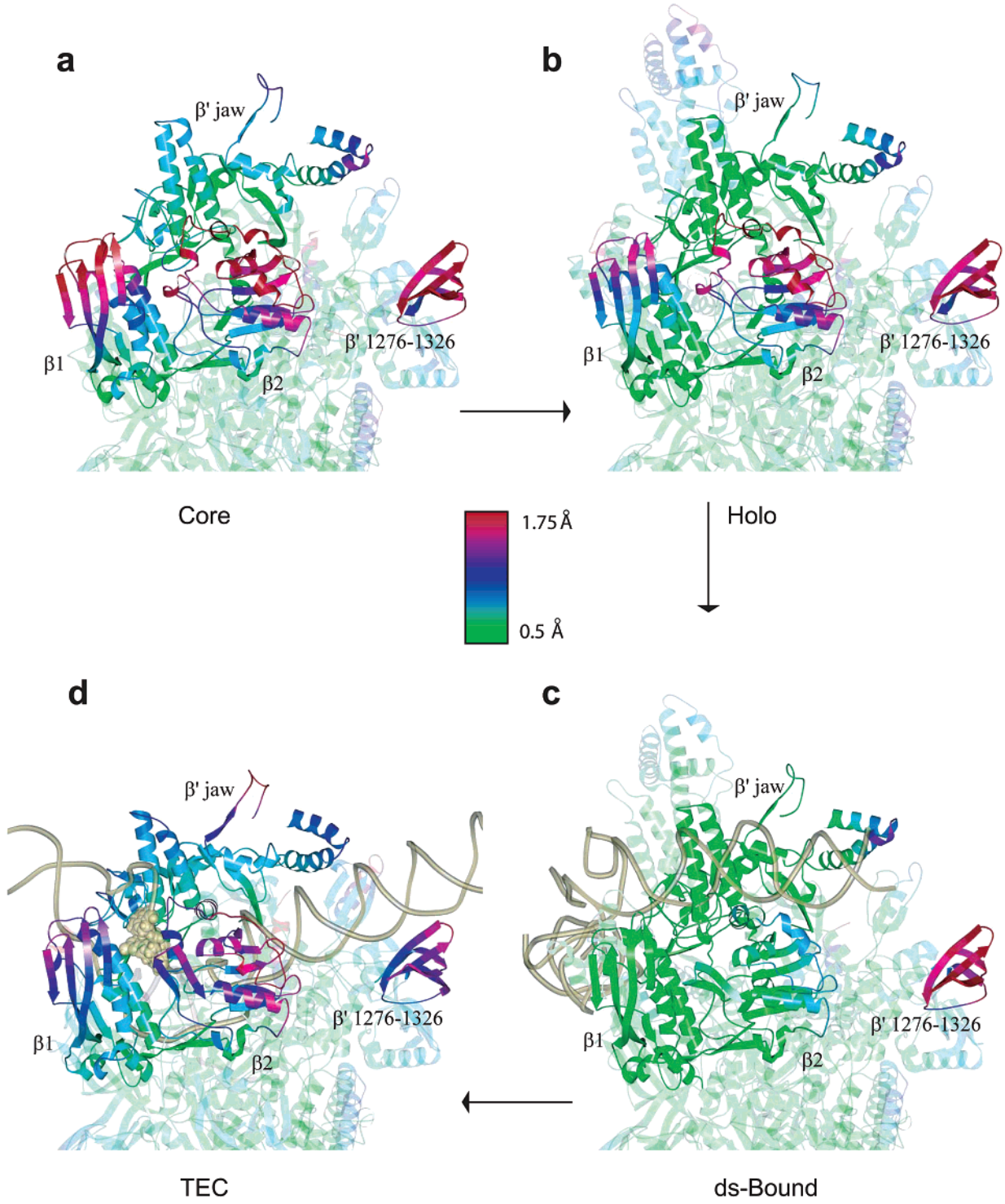


FIGURE 8: This plot maps the MM/BNM rmsf onto the structure of the RNAP in the (a) core, (b) holo, (c) ds-bound, and (d) TEC states. The structures are oriented so that the DNA binding cleft is running from left (upstream) to right (downstream). The $\beta 1$ and $\beta 2$ domains are closest to the reader, and the β' jaw is on the other side of the DNA binding cleft, furthest from the reader. The highlighted regions are discussed in the text, and their rmsf data are shown in other forms in Figure 7 ($\beta 1$ and $\beta 2$), Figure 9 (β' jaw), and Figure 10 (β' residues 1276−1326). This figure was created using MOLSCRIPT (*53*).

domain in the holo state; hence, the fluctuations are the same as in the core state. This also points out that the motions of the $\beta 1$ and $\beta 2$ domains are independent. This is an interesting observation because it has been previously proposed that these two domains bind DNA in a "concerted and interdependent manner" (*36*, *37*). In the functional states studied here, there seems to be no direct interaction between the two domains; i.e., stabilizing one of the domains does not affect the other. Any concerted motion of these two domains would have to be facilitated by either a conformational change of the domains not captured in the states studied here or additional protein factors that could bind to both domains. It is important to stress that the structures studied here are limited to those residues that could be resolved in the crystal structures that were used. Most pertinent to the immediate discussion is missing σ region 1.1, as well as *T. thermo-*

philus's lack of β dispensable regions 1 and 2; refs 36 and 37 describe the *E. coli* polymerase.

As in the β1 domain, as the RNAP binds a double-stranded piece of DNA in the ds-bound state, the flexibility of β2 is dramatically reduced. Again, favorable contacts between the DNA and the protein atoms hold the protein tightly to the bound DNA and the steric block of the DNA restricts the motion of the protein into the channel. Although the fluctuation certainly has been reduced, the magnitude of the reduction is less and thus the resultant flexibility is greater in β2 than in the β1 domain as can be seen by the aquamarine coloring on the extreme downstream portion of β2. Last, an examination of the TEC model also exhibits interesting behavior. Although the N-terminal (more upstream) regions of β2 show the same fluctuation as the core and holo complexes, the C-terminal (more downstream) portion of this domain has less flexibility than core and holo but more than the ds-bound complex. This can easily be understood by examining the structure of the TEC. The nontemplate strand effectively wedges itself between β1 and β2. The C-terminal (upstream) residues of β2 make direct contact with this single strand and are positionally stabilized, while the N-terminal (downstream) residues continue to have the flexibility present in the core and holo complexes; however, now the motion of the residues is best described as a hinging motion into and out of the cleft with the hinge being where β2 contacts the nontemplate strand. The motion observed is very similar to that proposed by Darst and co-workers in their model of the TEC (see Figure 3D of ref *38*).

β' Subunit. The β' subunit is the largest of all of the subunits that make up the RNAP. It forms one of RNAP's jaws, sitting directly across the DNA binding channel from the β subunit. Its overall topology is different from that of β; instead of the bilobal structure of β1 and β2, it forms a single pillar that rises from the α subunit base. In fact, the pillar is large enough so that its most extreme regions were unresolved in the crystal structure. One hundred twelve residues are missing (252−363) in the core, holo, and ds-bound complexes and 285 in the TEC model (163−447). This chain break complicates the normal-mode analysis. Much like the chain termini in the α subunits, because chain breaks are usually removed from a closely packed unit of the protein, they are not positionally stabilized. This results in unphysical fluctuation data that must be discarded. The result for these regions is at most a very qualitative statement that these regions are flexible. For the β' subunit, we have been forced to discard the data collected for residues 150−470. Most of these residues are nonconserved (163−452) (D. Vassylyev, personal communication). This arbitrary deletion of data was based on the first fluctuation peak from the chain break that corresponded to a fluctuation magnitude similar to that of the rest of the subunit. It is not expected that these missing residues will affect the results for the rest of the protein. BNM calculations were carried out with positional restraints applied to those residues of β' nearest the chain break, and no significant changes in the results were observed (data not shown).

The overall motion of the β' subunit shows it has a tendency to pivot into and away from the DNA binding channel. Figure 9 shows the extremes of the fourth meaningful normal mode ($f = 0.50$ cm^{-1}) of the core state. This mode defines an opening and closing of the RNAP jaws. Figure 9 shows that both the β and β' jaws have an open state (blue) and a closed state (red). It has been suggested previously through the work of Darst *et al.* (*14*) that the β' subunit indeed moves in this manner. Crystal structure comparisons coupled with a search for rigid body motions and prediction of molecular hinges allowed them to suggest that the β' subunit had large conformational mobility and could adopt an open and a closed state. In this work, we are able to confirm this prediction by showing that the motion they describe is indeed *energetically* feasible; i.e., the motion they captured can be realized without the crystalline environment. Their observations are of course limited by the differences in the structures that were obtained. In this work, however, we are able to predict the flexibility of the *entire* RNAP and do so from an energetic perspective without the confounding forces of crystal packing. We reproduce the motion suggested by Darst *et al.*, but we are also able to suggest that the β jaw has a wide range of conformational flexibility, comparable to if not greater than that of β'. This conclusion is supported by the magnitudes of the displacements in Figure 9 and can be seen even more clearly by comparing the rmsfs mapped onto the structures in Figure 8.

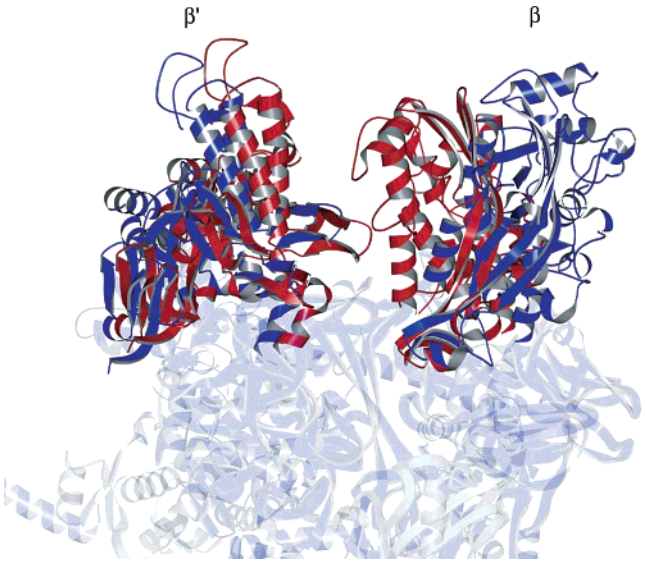


FIGURE 9: Extremes of meaningful mode number 4 ($f = 0.50$ cm^{-1}). Only the two jaws of the closed state (red) are shown. Both of the jaws have flexibility into and out of the DNA binding cleft. The magnitude of the motion of β' is smaller than that of β, but is still significant. This figure suggests that the RNAP is much like a crab claw with two flexible pincers. This figure was created using MOLSCRIPT (*53*).

Another interesting region to study is that of the trigger loop (residues 1238−1250), also known as the G loop. It has been suggested that this loop in the active site of the protein could play a part in the translocation of the RNAP during transcription (*4*). Figure 10 shows the fluctuations of the G loop (highlighted in tan) as well as the fluctuations of the residues that form the domain that packs up against the G loop (β' residues 1270−1330 highlighted in yellow). The data illustrate that upon DNA binding, the flexibility of the G loop that was present in the core and holo complexes is damped. Both the ds-bound and the TEC models show very little flexibility in this loop region. The ds-bound and TEC states are different in their fluctuation in the regions that pack behind the G loop, however. Because the DNA in the ds-bound complex has not yet been brought down into the

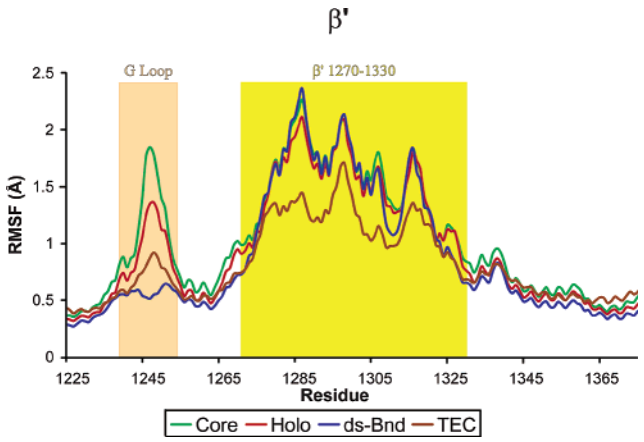


FIGURE 10: Fluctuations of a segment of the β' subunit. The G loop is highlighted in tan and shows a decrease in flexibility when DNA is bound. The yellow highlighted region, residues 1270−1330, packs behind the G loop, and the residues have large positional fluctuations. These fluctuations are damped somewhat in the TEC.

active site, its downstream double-helical portion does not make contact with residues 1270−1330. The TEC, on the other hand, has its template strand deep within the active site cleft so that RNA synthesis can take place. Its downstream DNA exits the active site channel from a deeper trajectory and thus makes stabilizing contacts with this portion of the β' subunit. This is best illustrated in Figure 8. Comparing the panels, one can see that the only state that does not have red coloring in this region is the TEC.

Conclusions

To our knowledge, the analysis performed here is the most detailed computational work performed on bacterial RNA polymerase to date. This work represents a first step in understanding the conformational changes in the functional cycle of transcription. With the block normal-mode analysis, we have been able to provide a systematic and semiquantitative picture of the structural flexibility of the RNAP system in several important functional states. Comparing the MM/BNM algorithm with the ENM has shown that MM/BNM has small but important advantages in investigations such as these. In particular, because it uses physically motivated interactions, the flexibilities of homology models may be assessed. It also provides a less arbitrary method of calculating protein flexibilities, a very useful result for questions that push the capabilities of normal-mode analysis. We note, however, that because ENM is more computationally efficient and gives reasonable values when used appropriately, it remains a useful method for quickly exploring protein structural flexibilities. For studies similar to those carried out by Delarue and Sanejouand (*22*), where a large number of structures were studied, ENM may provide the best combination of efficiency and reliability.

This highly approximate computational work is significant because it addresses the question of RNAP flexibility at each residue in four different functional states. This breadth of information is experimentally inaccessible for practical reasons. One experimental technique that in principle could be used to arrive at the same flexibility information provided here is site-directed spin labeling (SDSL) coupled with electron paramagnetic resonance (EPR). The spectra of nitroxide spin-labels are affected by both the polarity index of their immediate environment and their mobility (*39*). Several studies have utilized this technique to measure changes in mobility as substrate binds or as the protein of interest goes through its functional cycle (*39*−*43*). One could also, in principle, design single-molecule fluorescence resonance energy transfer (FRET) (*44*, *45*) experiments to monitor distance changes and thus flexibility. The latter approach is complicated by the need for appropriately sized distance changes, but the main practical difficulty in both of these experimental approaches is that it is not simple to introduce site-specific labels, especially without disrupting wild-type function (*46*−*48*). Creating the labeled proteins to use in such experiments in a systematic manner similar to the computational method presented here is of course impossibly work intensive. With the results presented in this work, however, several FRET or SDSL-EPR experiments could be rationally designed to confirm our predictions of the more interesting regions. An experimental measure of the β subunit's flexibility is perhaps the most intriguing; labels placed on the β1 and β2 domains could directly test our predictions for these regions. These experiments are imaginable in principle but in practice are extremely difficult. This underscores the value of the computational treatment in this paper: through appropriate use of theoretical techniques, we can begin to access information that is both interesting and difficult to experimentally determine.

We have studied the RNAP transcription cycle with MM/BNM, and each subunit has been examined. Several conclusions involving the rigidity or flexibility of those components have been drawn. First, the α subunits, specifically their N-terminal domains, have been shown to provide a rigid structural base on which the β and β' subunits can assemble to achieve catalysis. Similarly, the ω subunit has been shown to stay rigidly fixed in all of the studied functional states. This suggests that ω plays no mechanical role in the transcriptional cycle. Its hypothesized role, that of a folding chaperone, seems to be consistent with its rigid structure. We have also been able to show that the flexibilities of both the β and β' subunits are modulated throughout the functional cycle. In particular, we have shown that the β1 and β2 domains of the β subunit move independently and that their conformational mobility varies greatly depending on the presence or absence of DNA and the σ subunit. We have confirmed predictions of the flexibility of the β2 domain and β' subunit from an all-atom energetic perspective. In contrast to previous discussions dealing with the allowed conformational changes of RNAP, we have shown that the β subunit is perhaps the most mobile unit of the entire protein. This work presents a picture of the bacterial RNAP as a crab claw in which both pincers are flexible; i.e., RNAP is more like a pair of pliers than a crescent wrench. This has implications on our basic understanding of transcription initiation. It has been proposed by Astuerias and Craighead that double-stranded DNA never reaches the bottom of the active site cleft, i.e., that the dsDNA is initially held at the tip of the jaws and that the template strand melts away from the nontemplate strand and toward the active site (*49*). This was proposed on the basis of a crystal structure of an initiation-competent RNAPII complex with the clamp seemingly locked into the closed conformation (*50*, *51*). The crystal structures of bacterial initiation competent holo-

enzymes in their closed conformations were also used to bolster this hypothesis (*4*, *5*), as Asturias and Craighead argue that in these polymerases, dsDNA is wider than the observed cleft. Our results seem to disagree with their reasoning, but not necessarily with their conclusions. We show that in the holoenzyme, bacterial RNAP has a mobile clamp (β') which may open and allow dsDNA to fit inside the cleft near the active site. Although it is difficult to precisely determine the magnitude of the motion based on normal-mode analysis alone, our calculations also argue that the β subunit (the "other" jaw) is quite flexible and that even if the clamp domain is fixed, dsDNA may achieve access to the active site through widening of the cleft by β movement.

Overall, the major contribution of this work lies in the total description of these four functional states of RNAP. We hope that the pictures and trajectories created here will allow researchers both to have a better intuitive understanding of the range of RNAP motion as a molecular machine and to be able to design and understand experiments that probe the basic mechanism of transcription.

ACKNOWLEDGMENT

We thank Dr. Dmitry Vassylyev and Prof. Robert Landick for providing the homology models used in the work. In addition, Prof. Robert Landick, Dr. Ruth Saecker, and Rachel Mooney were particularly helpful in discussing the RNAP system. Last, A.V.W. thanks Mark Formaneck for his help with various computer issues.

REFERENCES

1. Guthold, M., Zhu, X., Rivetti, C., Yang, G., Thomson, N. H., Kasas, S., Hansma, H. G., Smith, B., Hansma, P. K., and Bustamante, C. (1999) Direct observation of one-dimensional diffusion and transcription by *Escherichia coli* RNA polymerase, *Biophys. J. 77*, 2284−2294.
2. Mooney, R. A., and Landick, R. (1999) RNA polymerase unveiled, *Cell 98*, 687−690.
3. Burgess, R. R., Travers, A. A., Dunn, J. J., and Bautz, E. K. (1969) Factor stimulating transcription by RNA polymerase, *Nature 221*, 43−46.
4. Vassylyev, D. G., Sekine, S., Laptenko, O., Lee, J., Vassylyeva, M. N., Borukhov, S., and Yokoyama, S. (2002) Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 Å resolution, *Nature 417*, 712−719.
5. Murakami, K. S., Masuda, S., and Darst, S. A. (2002) Structural basis of transcription initiation: RNA polymerase holoenzyme at 4 Å resolution, *Science 296*, 1280−1284.
6. Blatter, E. E., Ross, W., Tang, H., Gourse, R. L., and Ebright, R. H. (1994) Domain Organization of Rna-Polymerase α-Subunit: C-Terminal-85 Amino-Acids Constitute a Domain Capable of Dimerization and DNA Binding, *Cell 78*, 889−896.
7. Minakhin, L., Bhagat, S., Brunning, A., Campbell, E. A., Darst, S. A., Ebright, R. H., and Severinov, K. (2001) Bacterial RNA polymerase subunit ω and eukaryotic RNA polymerase subunit RPB6 ave sequence, structural, and functional homologs and promote RNA polymerase assembly, *Proc. Natl. Acad. Sci. U.S.A. 98*, 892−897.
8. Sosunov, V., Sosunova, E., Mustaev, A., Bass, I., Nikiforov, V., and Goldfarb, A. (2003) Unified two-metal mechanism of RNA synthesis and degradation by RNA polymerase, *EMBO J. 22*, 2234−2244.
9. Steitz, T. A. (1998) Structural biology: A mechanism for all polymerases, *Nature 391*, 231−232.
10. Gribskov, M., and Burgess, R. R. (1986) σ-Factors from *Escherichia coli*, *Bacillus subtilis*, Phage Sp01, and Phage T4 Are Homologous Proteins, *Nucleic Acids Res. 14*, 6745−6763.
11. Helmann, J. D., and Chamberlin, M. J. (1988) Structure and Function of Bacterial σ Factors, *Annu. Rev. Biochem. 57*, 839−872.
12. Zhang, G. Y., Campbell, E. A., Minakhin, L., Richter, C., Severinov, K., and Darst, S. A. (1999) Crystal structure of *Thermus aquaticus* core RNA polymerase at 3.3 Å resolution, *Cell 98*, 811−824.
13. Cramer, P., Bushnell, D. A., and Kornberg, R. D. (2001) Structural basis of transcription: RNA polymerase II at 2.8 Å resolution, *Science 292*, 1863−1876.
14. Darst, S. A., Opalka, N., Chacon, P., Polyakov, A., Richter, C., Zhang, G. Y., and Wriggers, W. (2002) Conformational flexibility of bacterial RNA polymerase, *Proc. Natl. Acad. Sci. U.S.A. 99*, 4296−4301.
15. Allison, L. A., Moyle, M., Shales, M., and Ingles, C. J. (1985) Extensive Homology among the Largest Subunits of Eukaryotic and Prokaryotic Rna-Polymerases, *Cell 42*, 599−610.
16. Gnatt, A. L., Cramer, P., Fu, J. H., Bushnell, D. A., and Kornberg, R. D. (2001) Structural basis of transcription: An RNA polymerase II elongation complex at 3.3 Å resolution, *Science 292*, 1876−1882.
17. Brooks, B., and Karplus, M. (1983) Harmonic Dynamics of Proteins: Normal-Modes and Fluctuations in Bovine Pancreatic Trypsin-Inhibitor, *Proc. Natl. Acad. Sci. U.S.A. 80*, 6571−6575.
18. Go, N., Noguti, T., and Nishikawa, T. (1983) Dynamics of a Small Globular Protein in Terms of Low-Frequency Vibrational-Modes, *Proc. Natl. Acad. Sci. U.S.A. 80*, 3696−3700.
19. Levitt, M., Sander, C., and Stern, P. S. (1983) The Normal-Modes of a Protein: Native Bovine Pancreatic Trypsin-Inhibitor, *Int. J. Quantum Chem.*, 181−199.
20. Hayward, S. (2001) Normal Mode Analysis of Biological Molecules, in *Computational Biochemistry and Biophysics* (Becker, O. M., MacKerell, A. D., Jr., Roux, B., and Watanabe, M., Eds.) pp 153−167, Marcel-Dekker, New York.
21. Tirion, M. M. (1996) Large amplitude elastic motions in proteins from a single-parameter, atomic analysis, *Phys. Rev. Lett. 77*, 1905−1908.
22. 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.
23. Valadie, H., Lacapcre, J. J., Sanejouand, Y. H., and Etchebest, C. (2003) Dynamical properties of the MscL of *Escherichia coli*: A normal mode analysis, *J. Mol. Biol. 332*, 657−674.
24. Blundell, T. L., and Johnson, L. N. (1976) *Protein Crystallography*, Academic Press, New York.
25. van Vlijmen, H. W. T., and Karplus, M. (1999) Analysis of calculated normal modes of a set of native and partially unfolded proteins, *J. Phys. Chem. B 103*, 3009−3021.
26. Dombroski, A. J., Walter, W. A., Record, M. T., Siegele, D. A., and Gross, C. A. (1992) Polypeptides Containing Highly Conserved Regions of Transcription Initiation-Factor Sigma-70 Exhibit Specificity of Binding to Promoter DNA, *Cell 70*, 501−512.
27. Chamberlin, M. J. (1974) The selectivity of transcription, *Annu. Rev. Biochem. 43*, 721−775.
28. Bushnell, D. A., Westover, K. D., Davis, R. E., and Kornberg, R. D. (2004) Structural basis of transcription: an RNA polymerase II-TFIIB cocrystal at 4.5 Å, *Science 303*, 983−988.
29. Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., and Karplus, M. (1983) CHARMM: a Program for Macromolecular Energy, Minimization, and Dynamics Calculations, *J. Comput. Chem. 4*, 187−217.
30. Neria, E., Fischer, S., and Karplus, M. (1996) Simulation of activation free energies in molecular systems, *J. Chem. Phys. 105*, 1902−1921.
31. Mackerell, A. D., Wiorkiewiczkuczera, J., and Karplus, M. (1995) An All-Atom Empirical Energy Function for the Simulation of Nucleic Acids, *J. Am. Chem. Soc. 117*, 11946−11975.
32. Lazaridis, T., and Karplus, M. (1999) Effective energy function for proteins in solution, *Proteins: Struct., Funct., Genet. 35*, 133−152.
33. Doruker, P., Jernigan, R. L., and Bahar, I. (2002) Dynamics of large proteins through hierarchical levels of coarse-grained structures, *J. Comput. Chem. 23*, 119−127.
34. Tama, F., Gadea, F. X., Marques, O., and Sanejouand, Y. H. (2000) Building-block approach for determining low-frequency normal modes of macromolecules, *Proteins: Struct., Funct., Genet. 41*, 1−7.
35. Li, G. H., and Cui, Q. (2002) A coarse-grained normal mode approach for macromolecules: An efficient implementation and application to Ca^{2+}-ATPase, *Biophys. J. 83*, 2457−2474.

36. Wigneshweraraj, S. R., Nechaev, S., Severinov, K., and Buck, M. (2002) β subunit residues 186−433 and 436−445 are commonly used by Eσ (54) and Eσ (70) RNA polymerase for open promoter complex formation, *J. Mol. Biol. 319*, 1067−1083.
37. Nechaev, S., Chlenov, M., and Severinov, K. (2000) Dissection of two hallmarks of the open promoter complex by mutation in an RNA polymerase core subunits, *J. Biol. Chem. 275*, 25516−25522.
38. Korzheva, N., Mustaev, A., Kozlov, M., Malhotra, A., Nikiforov, V., Goldfarb, A., and Darst, S. A. (2000) A structural model of transcription elongation, *Science 289*, 619−625.
39. Hubbell, W. L., Cafiso, D. S., and Altenbach, C. (2000) Identifying conformational changes with site-directed spin labeling, *Nat. Struct. Biol. 7*, 735−739.
40. Steinhoff, H. J., Mollaaghababa, R., Altenbach, C., Hideg, K., Krebs, M., Khorana, H. G., and Hubbell, W. L. (1994) Time-Resolved Detection of Structural-Changes During the Photocycle of Spin-Labeled Bacteriorhodopsin, *Science 266*, 105−107.
41. Sugata, K., Nakamura, M., Ueki, S., Fajer, P. G., and Arata, T. (2004) ESR reveals the mobility of the neck linker in dimeric kinesin, *Biochem. Biophys. Res. Commun. 314*, 447−451.
42. McHaourab, H. S., Lietzow, M. A., Hideg, K., and Hubbell, W. L. (1996) Motion of spin-labeled side chains in T4 lysozyme. Correlation with protein structure and dynamics, *Biochemistry 35*, 7692−7704.
43. Cornish, V. W., Benson, D. R., Altenbach, C. A., Hideg, K., Hubbell, W. L., and Schultz, P. G. (1994) Site-Specific Incorporation of Biophysical Probes into Proteins, *Proc. Natl. Acad. Sci. U.S.A. 91*, 2910−2914.
44. Deniz, A. A., Laurence, T. A., Dahan, M., Chemla, D. S., Schultz, P. G., and Weiss, S. (2001) Ratiometric single-molecule studies of freely diffusing biomolecules, *Annu. Rev. Phys. Chem. 52*, 233−253.
45. Stryer, L., and Haugland, R. P. (1967) Energy transfer: a spectroscopic ruler, *Proc. Natl. Acad. Sci. U.S.A. 58*, 719−726.
46. Severinov, K., Mustaev, A., Severinova, E., Bass, I., Kashlev, M., Landick, R., Nikiforov, V., Goldfarb, A., and Darst, S. A. (1995) Assembly of Functional *Escherichia coli* Rna-Polymerase Containing β-Subunit Fragments, *Proc. Natl. Acad. Sci. U.S.A. 92*, 4591−4595.
47. Mekler, V., Kortkhonjia, E., Mukhopadhyay, J., Knight, J., Revyakin, A., Kapanidis, A. N., Niu, W., Ebright, Y. W., Levy, R., and Ebright, R. H. (2002) Structural organization of bacterial RNA polymerase holoenzyme and the RNA polymerase-promoter open complex, *Cell 108*, 599−614.
48. Bergendahl, V., Heyduk, T., and Burgess, R. R. (2003) Luminescence resonance energy transfer-based high-throughput screening assay for inhibitors of essential protein−protein interactions in bacterial RNA polymerase, *Appl. Environ. Microbiol. 69*, 1492−1498.
49. Asturias, F. J., and Craighead, J. L. (2003) RNA polymerase II at initiation, *Proc. Natl. Acad. Sci. U.S.A. 100*, 6893−6895.
50. Armache, K. J., Kettenberger, H., and Cramer, P. (2003) Architecture of initiation-competent 12-subunit RNA polymerase II, *Proc. Natl. Acad. Sci. U.S.A. 100*, 6964−6968.
51. Bushnell, D. A., and Kornberg, R. D. (2003) Complete, 12-subunit RNA polymerase II at 4.1 Å resolution: Implications for the initiation of transcription, *Proc. Natl. Acad. Sci. U.S.A. 100*, 6969−6973.
52. Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD: Visual molecular dynamics, *J. Mol. Graphics 14*, 33.
53. Kraulis, P. J. (1991) Molscript: A Program to Produce Both Detailed and Schematic Plots of Protein Structures, *J. Appl. Crystallogr. 24*, 946−950.

BI049738+