Available online at www.sciencedirect.com
SCIENCE DIRECT
Biochemical and Biophysical Research Communications 343 (2006) 90–98
BBRC
www.elsevier.com/locate/ybbrc

Structure and dynamics of RNA polymerase II elongation complex

Atsushi Suenaga [a,*], Noriaki Okimoto [a], Noriyuki Futatsugi [a], Yoshinori Hirano [b], Tetsu Narumi [a], Yousuke Ohno [a], Ryoko Yanai [a], Takatsugu Hirokawa [c], Toshikazu Ebisuzaki [b], Akihiko Konagaya [a], Makoto Taiji [a,*]

[a] *Bioinformatics Group, RIKEN Genomic Sciences Center, 61-1 Ono-cho, Tsurumi, Yokohama, Kanagawa 230-0046, Japan*
[b] *Computational Astrophysics Laboratory, RIKEN, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan*
[c] *Computational Biology Research Center (CBRC), National Institute of Advanced Industrial Science and Technology (AIST), 2-43-17F Aomi, Koutou-ku, Tokyo 135-0064, Japan*

Received 1 February 2006
Available online 2 March 2006

Abstract

RNA polymerase (Pol) II is a fundamental and important enzyme in the transcription process. However, two mysterious questions have remained unsolved: how an unwound bubble of DNA is established and maintained, and how the enzyme moves along the DNA. To answer these questions, we constructed a model structure of the Pol II elongation complex with the 50 base pairs of DNA-24 bases of RNA including the unwound bubble of DNA and performed a molecular dynamics simulation. We obtained a reliable model structure of the Pol II elongation complex in the pre-translocation state which has not yet been determined by the X-ray crystallographic study. The model structure revealed that multiple protein loops work concertedly to form and maintain the bubble structure. We also found that the conformational change of a loop in the Pol II, fork loop 1, couples with the unidirectional movement of the Pol II along the DNA.

Keywords: Molecular dynamics; Molecular modeling; RNA polymerase II; Transcription; Molecular motor

RNA polymerase (Pol) II catalyzes synthesis of messenger RNA in eukaryotes thereby playing important roles in the regulation of gene expression and cell growth. Within the Pol II–DNA transcribing complex, double-stranded DNA (dsDNA) is separated from unwound template and non-template strands, and the Pol II synthesizes messenger RNA based on the code on the template strand by moving along the DNA. Thus, the Pol II is an interesting enzyme not only as a molecular machine for transcription but also as a molecular motor. On the structural bases, several X-ray crystallographic studies have revealed the protein–DNA, protein–RNA, and DNA–RNA interactions of the Pol [1–13]. On the bacterial Pol, the model structures of the Pol–DNA/RNA complex have been built [14,15], and they provide insight into the protein–DNA interactions that is critical for understanding the mode of the actions. However, the complete information about the interactions between the eukaryotic Pol II and DNA/RNA, especially interactions between the Pol II and unwound non-template DNA strand/upstream rewound DNA, has remained unsolved.

According to current research on the Pol II, there are two mysterious questions: how an unwound bubble of DNA established and maintained at the active center and how the Pol II moved along the DNA. To solve these questions, we constructed a model structure of the Pol II complexed with the 50 base pairs of DNA-24 bases of RNA including the unwound bubble of DNA and investigated the complete structure of the Pol II elongation complex. As the template structure for our modeling, we used the (incomplete) X-ray structure of the Pol II elongation complex [3]. In the X-ray structure,

* Corresponding authors. Fax: +81 45 507 2524.
E-mail addresses: suenaga@gsc.riken.jp (A. Suenaga), taiji@gsc.riken.jp (M. Taiji).

0006-291X/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.bbrc.2006.02.124

coordinates of four important protein loops at the DNA and RNA/Pol II interface, so-called fork loop 1, fork loop 2, rudder, and lid, were absent, and furthermore the DNA–RNA hybrid was short (consisting of only a 13 base template strand and a 9 bases of RNA) [3]. The recent X-ray structures of the Pol II complexed with DNA–RNA hybrid (consisting of a 15 base template strand and a 10 base RNA [5] or a 19 base template strand, a 7 base non-template strand, and a 10 base RNA [6]) have revealed the structures of these loops, however, structures of the upstream rewound DNA and the unwound non-template DNA strand in the transcription bubble still have not been solved (Figs. 1 and 2A). The remarkable differences between the previous [3] and recent [5,6] X-ray structures were shown at the downstream end of the DNA–RNA hybrid, occupancy or vacancy of the nucleotide-addition site, respectively. The previous X-ray structure (or our model structure) indicates the state just after addition of nucleotide triphosphate and before translocation (''pre-translocation state'') [3], on the contrary, the recent X-ray structures indicate the ''post-translocation state'' [5,6]. Thus, to build a complete model structure of the Pol II elongation complex in the pre-transition state which is not determined by the X-ray crystallographic study is an attractive subject, therefore comparison of these two molecular states would be a good examination not only to unveil the structure of the transcription bubble and its interactions but also to investigate the mechanism of the dynamics from the pre- to the post-translocation states.

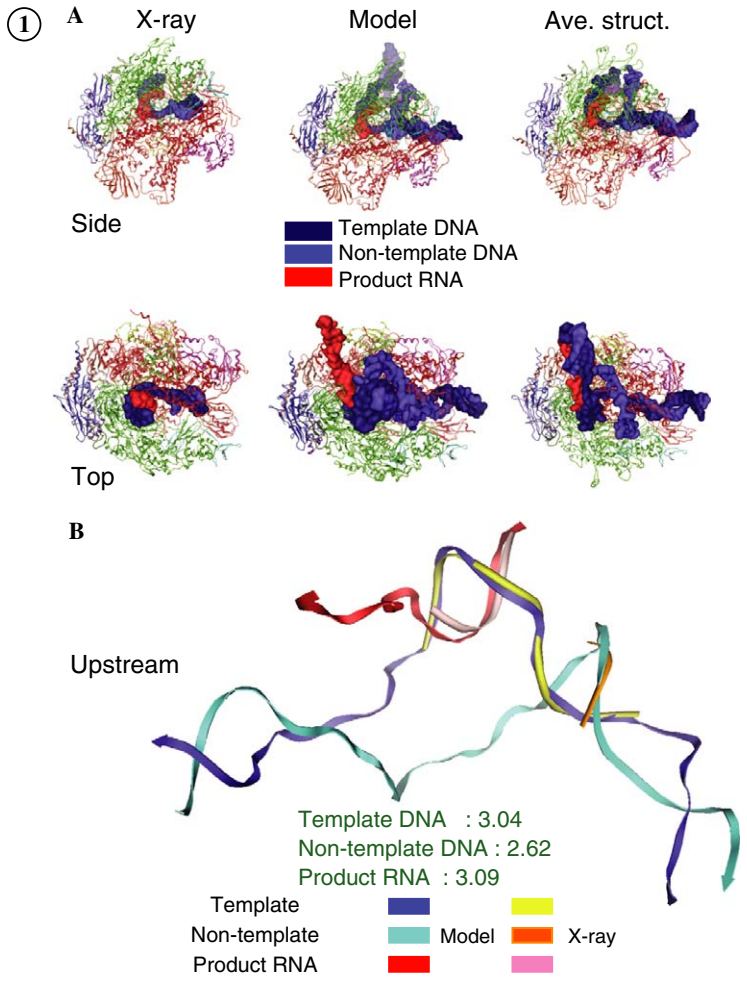
Here, we constructed the complete model structure of the Pol II elongation complex in the pre-translocation state. In the modeling structure, several protein loops and short nucleotide chains (<50 bases) which were absent in the previous X-ray structure [3] (template structure for our modeling) were modeled. Because these protein loops are short (consisting of about 10 amino acid residues) and comparison between previous [3] and recent [5,6] X-ray structures shows no structural differences of the Pol II, we thought that the relaxation of these modeled short loop structures may be the main point to obtain a correct model structure. Furthermore, since critical interactions between the Pol II and DNA (or RNA) proposed by Gnatt et al. [3] were carefully taken into account in our modeling of DNA and RNA, structural optimization of short DNA and RNA segments should be necessary. We performed a molecular dynamics (MD) simulation of the model structure of the Pol II elongation complex to relax these artificial effects in the model structure. In this paper, we present the complete Pol II elongation complex structure including upstream rewound DNA and unwound non-template DNA strand, and propose a novel mechanism of the movement of the Pol II along the DNA based on structural comparison of the pre- and post-translocation states of the Pol II elongation complex.

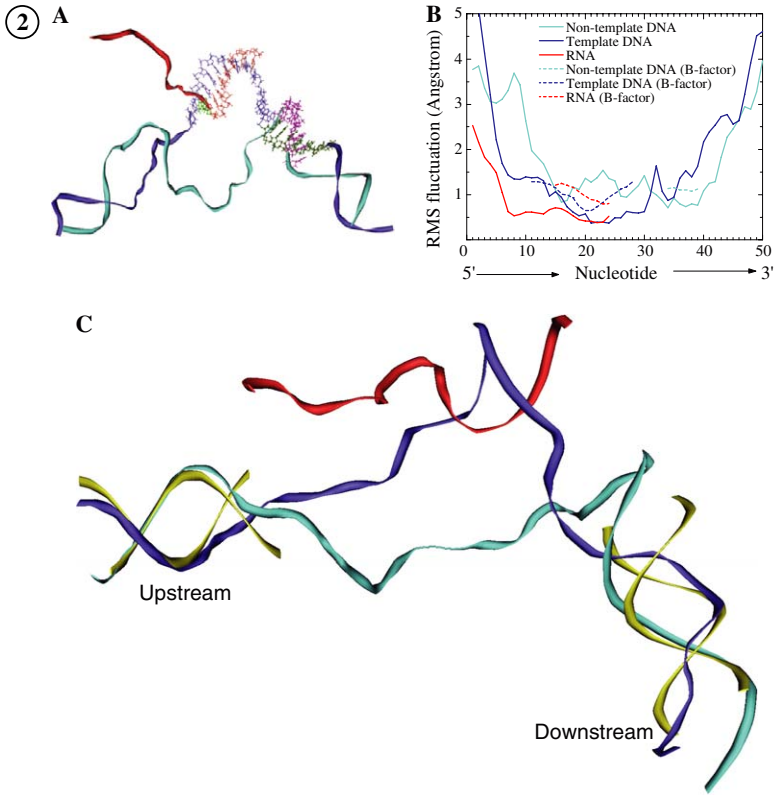
Materials and methods

The X-ray structures of nucleotide free Pol II [2] and transcribing complex [3] from yeast (Protein Data Bank entry 1l5O and 1l6H, respectively) were used as initial structures for this study. The coordinates of two small subunits, 4 and 7, were absent in the X-ray structures. However, we did not expect this absence to affect our results, as both subunits are dispensable for transcription [2,3,5]. In the X-ray structures, coordinates of several loops at Pol II/DNA and RNA interfaces (fork loop 1, fork loop 2, and rudder and lid) were also absent. Thus, we modeled these loops employing the segment match method [16] using the modeling software MOE (Chemical Computing Group Inc.). The 50 base pairs of DNA-24 bases of RNA including the transcription bubble was also modeled manually based on the proposal of Gnatt et al. [3] (Fig. 1). It is noted that, we assumed four points for modeling; (i) dsDNA is B-form, (ii) the dsDNA separates two single DNA strands at near the fork loop 2, (iii) these unwound single DNA strands rewind near the zipper, and (iv) the product RNA exits through the exit pore. The energies of the modeled nucleotide free Pol II and elongation complex structures were minimized by a 5000 step simulation in vacuum. Each energy-minimized structure was placed at a center of a sphere of TIP3P water molecules [17]. The sphere radii of 100 Å (for the nucleotide free Pol II) and 116 Å (for the Pol II elongation complex) were chosen to ensure that the atoms in the protein would be surrounded by at least five layers of water molecules. The total number of atoms included in the simulation was protein = 59,019 (for the nucleotide free Pol II) and 58,847 (for the Pol II elongation complex); $Zn^{2+} = 8$; $Mg^{2+} = 1$; DNA = 3165; RNA = 801; water molecules = 330,165 (for nucleotide free Pol II) and 454,713 (for Pol II elongation complex). During the simulation, water molecules were restrained by a soft half-harmonic potential (1.5 kcal mol^{-1} $Å^{-2}$) to avoid that they are free to boil off away from the protein.

Due to the large molecular size of the system of the Pol II system, it requires tremendous computational costs. Therefore, it has been difficult to perform the MD simulation even on a highly parallel supercomputer. Moreover, an accurate representation of the long-range interactions in an MD simulation is extremely important for representing the structure, dynamics, and energy of biomolecular systems properly, especially for highly charged molecules such as DNA and RNA [18,19]. The molecular dynamics machine (MDM) is a computer designed specifically for the calculation of large-scale, long-range interactions (van der Waals and Coulomb) with high speed and accuracy [20–26]. In this study, we used the MDM for the MD simulation of a large-scale biomolecule, Pol II. All MD simulations were carried out with the modified Amber 6.0 [27] for MDM, on a personal computer (Athlon 1.6 GHz) equipped with two MDGRAPE-2 boards, which are a main component of the MDM (32 chips, 512 Gflops) [20,21]. The CPU time usage was 13 s/step (1 ns/5 months) for the Pol II elongation complex. The parm96 force field [28] was adopted and the time step was set to 1.0 fs. 2,000,000 steps of the MD simulation (2.0 ns) was performed for each system (total of 4 ns). All non-bonded interactions, van der Waals and Coulomb forces and energies, were calculated directly and accurately using the MDM. The bond lengths involving hydrogen atoms were constrained to equilibrium lengths by the SHAKE method [29]. After gradually heating the system to 300 K over the first 50 ps (heating rate of 6 K/ps), the temperature was kept constant (NTV ensemble) by coupling to a temperature bath at 300 K with a coupling constant of 1.0 ps [30]. In the first 500-ps period of the MD simulation of the Pol II elongation complex, harmonic positional restraints were imposed on the Pol II atoms with harmonic force constant of 1.5 kcal mol^{-1} $Å^{-2}$ to allow the DNA, RNA, and solvent to equilibrate. After 500 ps, the harmonic force constant was gradually reduced to zero during the next 10-ps period. Another 1490 ps of MD simulation was performed for equilibration.

The conformations of the fork loop 1 in the average structure (pre-translocation state) and post-translocation state were used as an initial and a target in a targeted MD simulation [31] on Pol II elongation complex. Only five loops (fork loop 1, fork loop 2, lid, rudder, and zipper), four switches, a bridge helix, wall region, DNA, RNA, and a Mg^{2+} were


1
A
X-ray
Model
Ave. struct.
Side
Template DNA
Non-template DNA
Product RNA
Top
B
Upstream
Template DNA : 3.04
Non-template DNA : 2.62
Product RNA : 3.09
Template
Non-template
Product RNA
Model
X-ray


2
A
B
Non-template DNA
Template DNA
RNA
Non-template DNA (B-factor)
Template DNA (B-factor)
RNA (B-factor)
RMS fluctuation (Angstrom)
0
1
2
3
4
5
0
10
20
30
40
50
5'
Nucleotide
3'
C
Upstream
Downstream

considered in the simulation. The N- and C-termini of these loops and switches were capped by acetyl and *N*-methyl groups, respectively, for neutralization. The initial structure was surrounded by a 9 Å shell of TIP3P water molecules [17], and the energy of the solvated structure was minimized by a 1000 step simulation. During the simulation, the positions of these acetyl and *N*-methyl groups were fixed. The energy function for the targeted MD simulation was augmented by an artificial energy term that depended on the RMSD between the current conformation of the fork loop 1 and that in the target structure. The conformational transition was driven by applying RMSD restraints to main chain atoms in the fork loop 1, with a force constant of 0.5 kcal mol^{-1} $Å^{-2}$. The choice of force constant did not affect the results because the simulations, used the force constants of 0.25, 0.5, and 1.0 kcal mol^{-1} $Å^{-2}$, provide almost the same results.

Results and discussion

We used the trajectory of the last 1.0 ns simulation (from 1.0 to 2.0 ns) for data analysis and the structures sampled by the trajectory were averaged. Since the simulation time scale in this study is too short to observe large conformational change of the protein, we compared structural feature between the averaged (pre-translocation state) and recent X-ray (post-translocation state) [6] structures, and extracted dynamical properties of the averaged structure. Although the simulation was too short, so that we did not directly observe the large scale motion from the pre- to post-translocation states in this simulation, the structural comparison between pre- and post-transition states provided us the aspect of the dynamic property regarding to the Pol II movement. Our aim of this modeling and simulation is not only to build model structure but also to show effectiveness of the MD simulation for modeling of flexible loop structures. Our primary goals are also to provide insight into the protein–DNA (or RNA) interactions of the Pol II elongation complex and to guide further biological experiments.

From our simulation, we obtained a reliable model structure of the Pol II elongation complex in the pre-translocation state involving the transcription bubble which has not yet been determined by the X-ray crystallographic study (Fig. 1). During the simulation, the Pol II and the transcription bubble near the active center were stable. The transcription bubble was well maintained during the simulation, which indicates that the bubble structure is stabilized in the Pol II elongation complex interacting with the Pol II.

In the averaged structure, the Pol II was relatively stable, while the modeled DNA and RNA showed large backbone RMSD values from the initial model structure (Fig. 3 and Table 1). These results suggest the reason why upstream and downstream dsDNA and the unwound non-template strand were absent or largely disordered in the X-ray structures. Moreover, the terminal regions (upstream and downstream) of the DNA/RNA and the unwound non-template strand showed large fluctuations (Fig. 2B), indicating that equilibration of these terminal regions of the DNA and RNA was not sufficient during 2.0 ns period. The structural deviation of the Pol II (main-chain RMSD < 3.0 Å) originated from several loop regions which have been partially ordered or disordered in the X-ray structure [3] (Fig. 3B), which means that the core regions of the Pol II were very stable in the simulation. Furthermore, each subunit showed rigid body motion (Table 2).

Surprisingly, the downstream dsDNA in the averaged structure which is held by the Pol II clamp deviated from

Table 1
Main-chain RMSD values from initial model structure

	Pol II	DNA	RNA
RMSD[a]	2.413	14.252	6.211
SD[b]	0.053	0.331	0.159

[a] Averaged values from 1.0 to 2.0 ns simulation.
[b] Standard deviation.

Table 2
RMSD value of each subunit from initial model structure

Subunit	RMSD[a]
Rpb 1	2.167 ± 0.206
Rpb 2	2.771 ± 0.136
Rpb 3	1.361 ± 0.076
Rpb 5	1.208 ± 0.155
Rpb 6	1.234 ± 0.188
Rpb 8	2.162 ± 0.150
Rpb 9	1.547 ± 0.174
Rpb 10	1.554 ± 0.235
Rpb 11	1.236 ± 0.121
Rpb 12	1.587 ± 0.176

[a] Averaged values from 1.0 to 2.0 ns simulation. RMSD values of each subunit were calculated after all Cα atoms in each subunit were fitted. Rpb1, Rpb2, and Rpb8 showed relatively large values (see text in detail).

Fig. 1. Structures of the Pol II elongation complex. (A) X-ray (left) [6], initial model (center), and averaged (right) structures of the Pol II elongation complex. Backbone of the Pol II is drawn by ribbons. Each subunit is specified by colors. Template, non-template DNA strands, and RNA are represented by surface model and colored in dark-blue, blue, and red, respectively. (B) Comparison of the DNA and RNA structures between recent X-ray [6] and averaged structures. Backbone RMSD values between the X-ray and averaged structures are also represented.

Fig. 2. Structures of DNA and RNA. (A) The modeled DNA and RNA are represented by ribbons and the template structure for our model building [3] is drawn by sticks. Template, non-template DNA, and RNA strands are colored in blue, cyan, and red, respectively. The structures of the parts of DNA, and RNA which are modeled in this study but published in recent X-ray structure [6] are shown by ball-stick models (template strand, dark green; non-template strand, magenta; RNA, light green). (B) RMS fluctuations obtained from simulation as a function of nucleotide. The nucleotides are numbered from its 5′ to 3′ ends. Experimental RMS fluctuations are calculated from B-factor in the recent X-ray structural data [6] and drawn by dashed lines. (C) Upstream and downstream dsDNAs in the averaged structure, which are superposed on a canonical B-form dsDNA. The canonical B-form dsDNA is colored in yellow. Template, non-template, and RNA strands are colored in blue, cyan, and red, respectively.

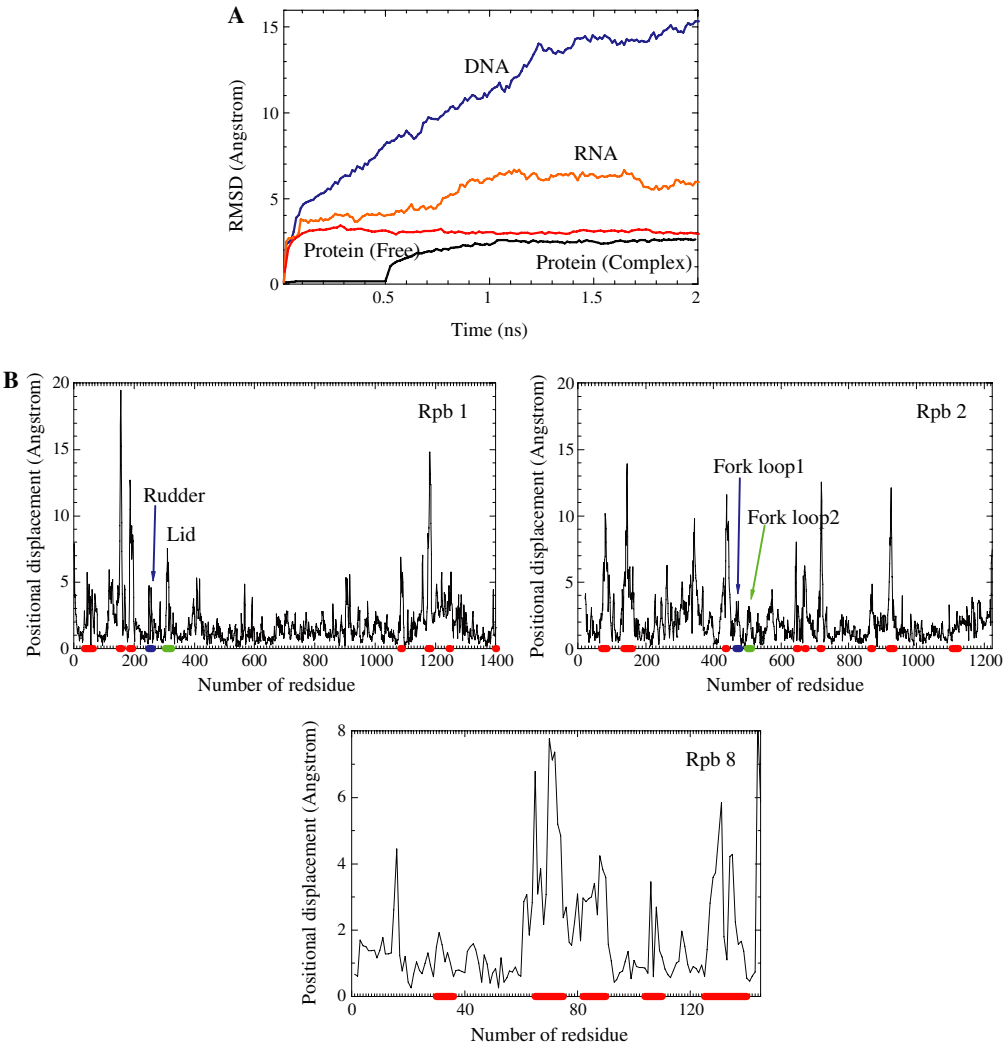


Fig. 3. Time series of main-chain RMSDs (A) and the positional displacement of each residue between initial model and averaged structures (B). Only the Cα atom of each residue was considered for calculating the displacement. Color bars on the *X*-axis indicate the partially ordered and disordered loops in the X-ray structure [3].

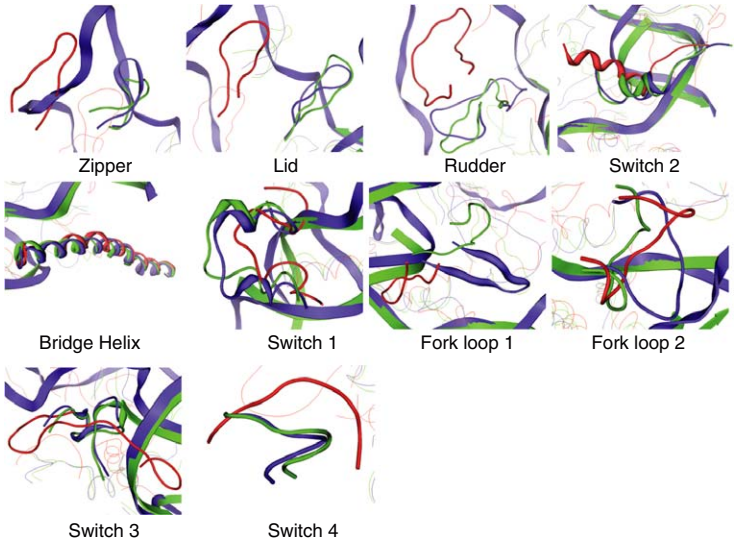


Fig. 4. Comparison of the important loop structures at the Pol II/DNA (RNA) interface between recent X-ray (green) [6] and averaged structures obtained from free Pol II (red) and Pol II elongation complex (blue) simulations.

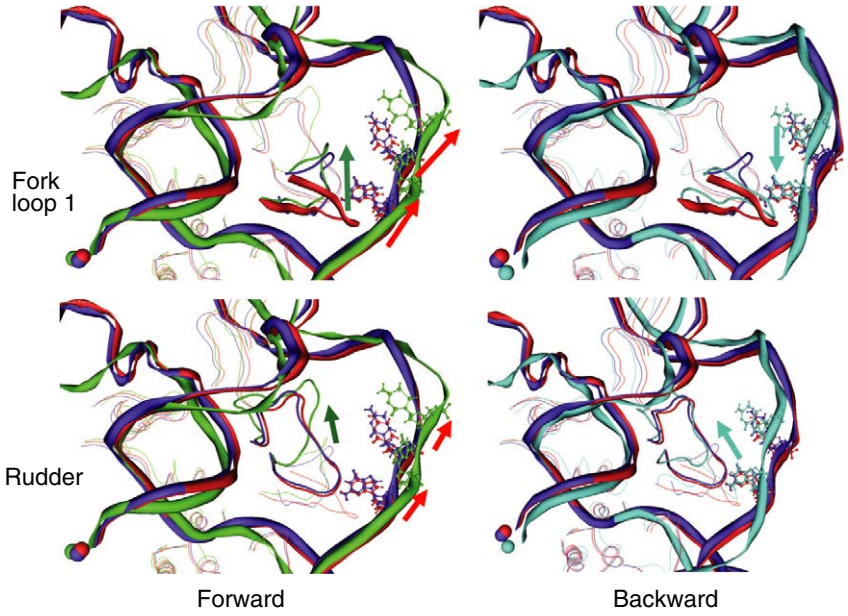


Fig. 5. Responses of the fork loop 1 and rudder in targeted MD trajectories that move from the pre- to the post-translocation states (forward) and from the post- to the pre-translocation states (backward). Averaged structure (pre-translocation state) obtained from simulation is colored in red and post-translocation state is colored in blue. Final structures obtained from forward and backward targeted MD simulations are colored in green and cyan, respectively. The fork loop 1, rudder, and DNA/RNA are highlighted by ribbons. Mg^{2+} is represented by a sphere. Seeral bases are drawn by sticks as markers.

the canonical B-form dsDNA (Fig. 2C), although the region was modeled as B-form dsDNA in the initial model structure. This indicates that the downstream dsDNA is distorted and destabilized in the Pol II before unwinding. Although the modeled DNA and RNA were considerably moved from the initial positions and largely fluctuated as described above, the fluctuations and structures of the parts of DNA and RNA (structures of the parts of DNA and RNA were modeled in this study but published recently by other research groups [6], Fig. 2A) were well reproduced compared to those of the recent X-ray structure [6] (RMSD $\approx$ 3.0 Å, Fig. 1B). These results indicate that the artificial structures of the modeled DNA and RNA were relaxed.

The transcription bubble is established and maintained by the multiple protein loops and an α-helix in the Pol II. These loops and the α-helix consist of fork loop 1, fork loop 2, rudder, lid, zipper, switch 1–4, and bridge helix. Making a comparison of the modeled loop structures (fork loop 1, fork loop 2, rudder, and lid) between recent X-ray and averaged structures, these loop structures in the averaged structure were well reproduced compared to those in the recent X-ray structure, except for fork loop 1 (Fig. 4), although any information of the recent X-ray structure was not used in our modeling procedure. Values of main-chain RMSD of these loop structures (except for the fork loop 1) between the recent X-ray and averaged structures were in the range of 1.3–2.9 Å. This result strongly suggests that our simulation was performed properly and several artifacts in the model structure were

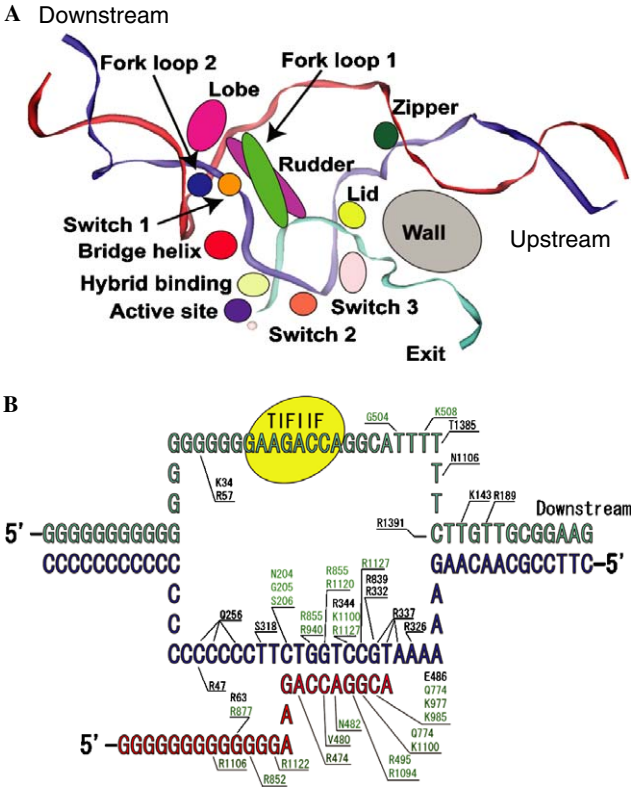


Fig. 6. Structure and formation of transcription bubble. (A) Schematic representation of the structure and formation of the transcription bubble. (B) Hydrogen bond network in the Pol II elongation complex. Template, non-template DNA, and RNA are colored in blue, cyan, and red, respectively. Residues are numbered according to [3]. Residues in subunit 2 (Rpb 2) are colored in green.

eliminated in our 2.0 ns-simulation. Interestingly, the structure of the fork loop 1 in the averaged structure was different from that in the recent X-ray structure (RMSD = 9.0 Å, Fig. 4). In our simulation, we used the X-ray structure of the pre-translocation state [3] as the template structure for modeling and the recent X-ray structure has been reported as in the post-translocation state [6]. Thus, that structural difference in the fork loop 1 may be caused by the movement (at least movement from pre- to post-translocation states) of the Pol II along the DNA, and the fork loop 1 should help the Pol II movement.

As a control simulation, we also performed an additional MD simulation of nucleotide free Pol II. In the averaged structure obtained from the nucleotide free Pol II simulation, all loops important for the formation and maintenance of the transcription bubble showed larger fluctuations than those of the Pol II elongation complex (values of RMS fluctuations of these loops in the nucleotide free Pol II were ranging from 0.6 to 2.3 Å; on the other hand, those obtained from the Pol II elongation complex simulation were ranging from 0.5 to 1.3 Å) and structures of these loops were different from those in the Pol II elongation complex (values of RMSD were ranging from 2.3 to 9.0 Å, Fig. 4). Especially, the structure of the fork loop 1 was remarkably different (Fig. 4). These results suggest that these loops that are flexible in the nucleotide free Pol II are stabilized by interacting with the DNA (or RNA) in the Pol II elongation complex and the conformation of the fork loop 1 dramatically changes via the DNA binding. Taken together, we speculate that the fork loop 1 plays a very important role in the binding to DNA and movement of the Pol II accompanied by its conformational change.

In order to examine how the fork loop 1 plays a part in the Pol II movement mechanism, we performed a targeted MD simulation [31] of the Pol II elongation complex in which an artificial force was applied to the fork loop 1 for its conformational change from that seen in the pre- to the post-translocation states (forward simulation) during the course of a short simulation (10 ps). The transition of the fork loop 1 conformation in the targeted MD simulation resulted in a movement of the non-template strand in the transcription bubble (Fig. 5). At the split and knot region of the unwound bubble of DNA, no sliding of the strand was observed. This result indicates that the conformational change of the fork loop 1 is closely connected with sliding of the non-template strand in the transcription bubble (between split and knot). In addition, a conformation of another loop, rudder, changed together with the conformational change of the fork loop 1. An experimental study reported that the rudder makes critical contributions to elongation complex stability through direct interactions with the nascent RNA [32], thus the conformational change of the rudder accompanied by the sliding of the DNA–RNA is consistent with the experimental facts. Interestingly, no sliding of the non-template strand in the transcription bubble was observed in the backward simulation (an artificial force was applied to the fork loop 1 for its conformational change from that seen in the post- to the pre-translocation states) (Fig. 5). These results indicate that the conformational change of the fork loop 1 resembles a "heart valve," which seems to act as a trigger of the unidirectional movement. To understand the mechanism of movement of the Pol II along the DNA, further simulations and its energetic analyses should be essential, and these simulations and analyses using our model structure obtained from this study being planned in next simulations.

Previous studies reported that the oscillation of the bridge helix (F-bridge or O-helix) from straight to bent form is very important for moving the Pol along DNA [3,10,12,13,33,34]. The template and non-template strands will slide completely by the conformational change of the fork loop 1, which acts in concert with the bridge helix oscillation. The structural change of the fork loop 1 coupled to the movement of the Pol II along the DNA, especially to the unidirectional movement, is a novel mechanism to explain the function of the Pol II as a molecular motor. The unique and complex mechanism for moving the Pol II along the DNA may be necessary for processive movement and fidelity of transcription process.

How an unwound bubble of DNA is formed and maintained at the active center has been a mysterious question in the transcription process. Our study based on model building and MD simulation of the Pol II elongation complex will answer the question. First of all, the formation of the transcription bubble is initiated by the interception of a stream of dsDNA. The bridge helix stands in the stream of the dsDNA after which the helix separates the template and non-template strands (Figs. 6A and 7A, left). The non-template strand is repelled by two negatively charged residues, E833 in the bridge helix and D503 in the fork loop 2 (Fig. 7B, left), and the template strand is pulled by four positively charged residues, R1386 in the switch 1, K330, K332, and R337 in the switch 2 (Fig. 7B, left).

Each separated template and non-template strand is guided in the correct direction by switch 1 and fork loop 2, respectively. These loops also prevent the re-association of these DNA strands. The non-template strand goes away from the active center along the lobe, on the contrary, the template strand goes to the active site (Fig. 6A). Next, at the unwound DNA strands/Pol II interfaces, the template strand is preserved and stabilized by switch 2 and 3 by forming 17 hydrogen bonds in the bubble structure (Figs. 6A and 7A, right). The product RNA forms a nine base pair DNA–RNA hybrid together with the template strand, and the DNA–RNA hybrid structure is maintained by the active site region, the hybrid binding region, fork loop 1 and rudder, consistent with previous studies [3,5,6,34,35] (Fig. 6A). The unwound non-template strand and the RNA are bridged and kept a distance between them by the fork loop 1 and rudder, which results in the bubble structure being maintained [32] (Fig. 6A).

The product RNA is separated from the template strand by a lid (Figs. 6A and 7A, bottom). The lid also prevents

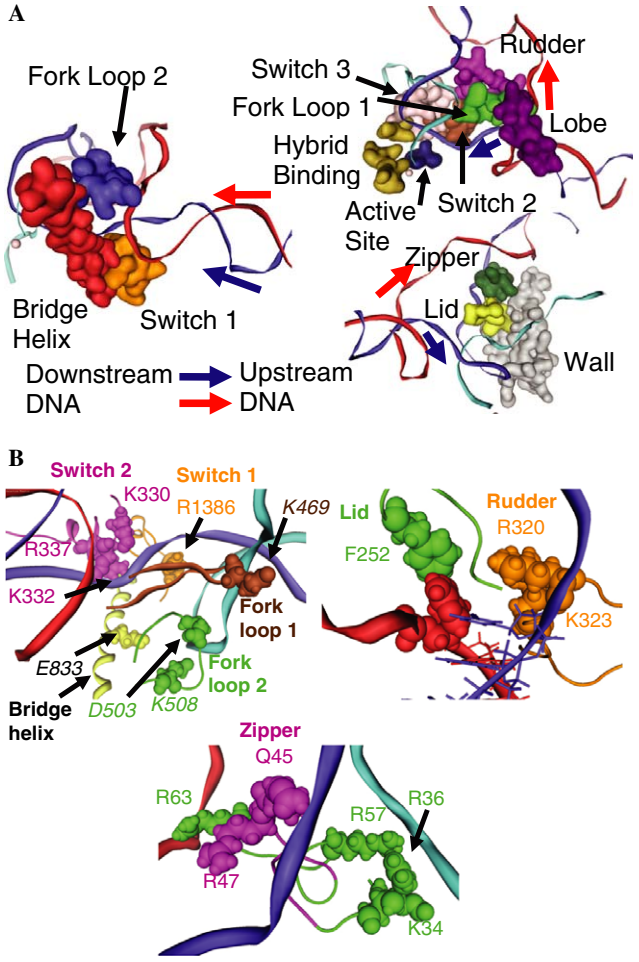


Fig. 7. Structure and formation of transcription bubble. (A) Separating the double-stranded DNA region (left), "bubble" region (top right), and rewound DNA region (bottom right) are shown separately. Several loops and DNA–RNA hybrid are represented by surface and ribbon models, respectively. (B) Interactions between the Pol II loops and DNA/RNA. The template, non-template DNA strands, and RNA (red) are colored in blue, cyan, and red, respectively. Residues are numbered according to Gnatt et al. [3] and residues in subunit 2 (Rpb 2) are shown in italics.

re-association of the RNA and the template strand. The phenylalanine 252 (F252) in the lid contacts with a base in the RNA at the position of the slit of the DNA–RNA hybrid by an edge-to-face aromatic–aromatic interaction. Thus, the F252 plays a role of "blocker" to prevent the formation of the base-paring between the template strand and the RNA at the slit point (Fig. 7B, right). The separated RNA goes to exit pore and exits along RNA exit groove 1 [3]. The wall bends the separated template strand, after which the template strand goes away from the active center (Fig. 6A).

Finally, the template strand meets the non-template strand and the zipper ties these strands (Fig. 6A). Four positively charged residues (R34, R36, R47, and R57) and a glutamine (Q45) lead each DNA strand to the appropriate position to rewind these strands (Fig. 7B, bottom). In the transcription bubble, 61 hydrogen bonds between Pol II and DNA/RNA are formed (Fig. 6B). In the hydrogen bond network, 35 hydrogen bonds are formed between Pol II and the template strand, 7 hydrogen bonds are formed between Pol II and the non-template strand, and 19 hydrogen bonds are formed between Pol II and RNA. Most of the hydrogen bonds are non-specific (residues in Pol II contact with phosphate of DNA) and it is consistent with the fact that the Pol II must move along the DNA. The fewer number of hydrogen bonds seen at the Pol II-non template strand interface compared to the Pol II-template strand interface may be supplemented by interacting with another protein; transcription factor IIF-non template strand interaction (Fig. 6B).

In summary, our study unveiled the complete Pol II elongation complex structure and its dynamic property, and the protein–DNA, protein–RNA, and DNA–RNA interactions in the transcription bubble at the atomic level. Multiple important loops in the Pol II work concertedly to form and maintain the transcription bubble, and it seems to be a sophisticated molecular machine. Additionally, the Pol is a kind of "molecular motor" too. From our results, we propose a novel mechanism of movement of the Pol II along DNA. On the novel mechanism, the conformational change of the fork loop 1 plays an important role not only to bind the Pol II with the DNA but also to move the Pol II along the DNA. Interestingly, our results show a possibility that the conformational change of the loop is one of the sources for the unidirectional movement of the Pol II. It will act concertedly with the oscillation of the bridge helix and then the Pol II will move completely along the DNA.

Acknowledgments

We sincerely thank Dr. T. Koishi, Dr. R. Susukita, Dr. K. Yasuoka, Dr. A. Kawai, and Mr. H. Furusawa for the use of the MDM. We also thank Dr. M. Hatakeyama and Dr. D. Nerukh for critical reading of the manuscript. We are grateful to Prof. R.D. Kornberg for his encouragement. This work is supported by the contracted research "Protein 3000 Project" by the Ministry of Education, Culture, Sports, Science and Technology of Japan, and Special Postdoctoral Research Program from RIKEN (to Y.H.).

References

[1] P. Cramer, D.A. Bushnell, J. Fu, A.L. Gnatt, B. Maier-Davis, N.E. Thompson, R.R. Burgess, A.M. Edwards, P.R. David, R.D. Kornberg, Architecture of RNA polymerase II and implications for the transcription mechanism, Science 288 (2000) 640–649.
[2] P. Cramer, D.A. Bushnell, R.D. Kornberg, Structural bases of transcription: RNA polymerase II at 2.8 Å resolution, Science 292 (2001) 1863–1876.
[3] A.L. Gnatt, P. Cramer, J. Fu, D.A. Bushnell, R.D. Kornberg, Structural bases of transcription: RNA polymerase II elongation complex at 3.3 Å resolution, Science 292 (2001) 1876–1882.
[4] D.A. Bushnell, K.D. Westover, R.E. Davis, R.D. Kornberg, Structural bases of transcription: an RNA polymerase II-TFIIB cocrystal at 4.5 Å, Science 303 (2004) 983–988.

[5] K.D. Westover, D.A. Bushnell, R.D. Kornberg, Structural bases of transcription: separation of RNA from DNA by RNA polymerase II, Science 303 (2004) 1014–1016.
[6] H. Kettenberger, K.-J. Armache, P. Cramer, Complete RNA polymerase II elongation complex structure and its interactions with NTP and TFIIS, Mol. Cell 22 (2004) 955–965.
[7] G. Zhang, E.A. Campbell, L. Minakhin, C. Richter, K. Severinov, S.A. Darst, Crystal structure of *Thermus aquaticus* core RNA polymerase at 3.3 Å resolution, Cell 98 (1999) 811–824.
[8] K.S. Murakami, S. Masuda, S.A. Darst, Structural bases of transcription initiation: RNA polymerase holoenzyme at 4 Å resolution, Science 296 (2002) 1280–1284.
[9] K.S. Murakami, S. Masuda, E.A. Campbell, O. Muzzin, S.A. Darst, Structural bases of transcription initiation: RNA polymerase holoenzyme–DNA complex, Science 296 (2002) 1285–1290.
[10] D.G. Vassylyev, S. Sekine, O. Laptenko, J. Lee, M.N. Vassylyeva, S. Borukhov, S. Yokoyama, Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 Å resolution, Nature 417 (2002) 712–719.
[11] I. Artsimovitch, V. Patlan, J. Sekine, M.N. Vassylyeva, T. Hosaka, K. Ochi, S. Yokoyama, D.G. Vassylyev, Structural basis of transcription regulation by alarmone ppGpp, Cell 117 (2004) 299–310.
[12] D. Temiakov, V. Patlan, M. Anikin, W.T. McAllister, S. Yokoyama, D.G. Vassylyev, Structural basis for substrate selection by T7 RNA polymerase, Cell 116 (2004) 381–391.
[13] W.Y. Yin, T.A. Steitz, The structural mechanism of transcription and helicase activity in T7 RNA polymerase, Cell 116 (2004) 393–404.
[14] N. Naryshkin, A. Revyakin, Y. Kim, V. Mekler, R.H. Ebright, Structural organization of the RNA polymerase–promoter open complex, Cell 101 (2000) 601–611.
[15] N. Korzheva, A. Mustaev, M. Kozlov, A. Malhotra, V. Nikiforov, A. Goldfarb, S.A. Darst, A structural model of transcription elongation, Science 289 (2000) 619–625.
[16] M. Levitt, Accurate modeling of protein conformation by automatic segment matching, J. Mol. Biol. 226 (1992) 507–533.
[17] W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, M.L. Klein, Comparison of simple potential functions for simulating liquid water, J. Chem. Phys. 79 (1983) 926–935.
[18] M. Saito, Molecular dynamics simulations of proteins in water without the truncation of long-range Coulomb interactions, Mol. Simul. 8 (1995) 321–333.
[19] D.M. York, W. Yang, H. Lee, T.A. Darden, L.G. Pedersen, Toward the accurate modeling of DNA—the importance of long-range electrostatics, J. Am. Chem. Soc. 117 (1995) 5001–5002.
[20] T. Narumi, R. Susukita, T. Ebisuzaki, G. McNiven, B. Elmergreen, Molecular dynamics machine: special-purpose computer for molecular dynamics simulations, Mol. Simul. 21 (1995) 401–415.
[21] R. Susukita, T. Ebisuzaki, B.G. Elmegreen, H. Furusawa, K. Kato, A. Kawai, Y. Kobayashi, T. Koishi, G.D. McNiven, T. Narumi, K. Yasuoka, Hardware accelerator for molecular dynamics: MDGRAPE-2, Comput. Phys. Comm. 155 (2003) 115–131.
[22] A. Suenaga, M. Hatakeyama, M. Ichikawa, X. Yu, N. Futatsugi, T.N. arumi, K. Fukui, T. Terada, M. Taiji, M. Shirouzu, S. Yokoyama, A. Konagaya, Molecular dynamics, free energy and SPR analyses of the interactions between the SH2 domain of Grb2 and ErbB phosphotyrosyl peptides, Biochemistry 42 (2003) 5195–5200.
[23] A. Suenaga, A.B. Kiyatkin, M. Hatakeyama, N. Futatsugi, N. Okimoto, Y. Hirano, T. Narumi, A. Kawai, R. Susukita, T. Koishi, H. Furusawa, K. Yasuoka, N. Takada, Y. Ohno, M. Taiji, T. Ebisuzaki, J.B. Hoek, A. Konagaya, B.N. Kholodenko, Tyr317 phosphorylation increases Shc structural rigidity and reduces coupling of domain motions remote from the phosphorylation site as revealed by molecular dynamics simulations, J. Biol. Chem. 279 (2004) 4657–4662.
[24] A. Suenaga, N. Takada, M. Hatakeyama, M. Ichikawa, X. Yu, K. Tomii, N. Okimoto, N. Futatsugi, T. Narumi, M. Shirouzu, S. Yokoyama, A. Konagaya, M. Taiji, Novel mechanism of interaction of p85 subunit of PI3K and ErbB3 receptor-derived phosphotyrosyl peptides, J. Biol. Chem. 280 (2005) 1321–1326.
[25] N. Okimoto, T. Nakamura, A. Suenaga, N. Futatsugi, Y. Hirano, I. Yamaguchi, T. Ebisuzaki, Cooperative motions of protein and hydration water molecules: molecular dynamics study of Scytalone dehydratase, J. Am. Chem. Soc. 126 (2004) 13132–13139.
[26] T. Koishi, S. Yoo, K. Yasuoka, X.C. Zeng, T. Narumi, R. Susukita, A. Kawai, H. Furusawa, A. Suenaga, N. Okimoto, N. Futatsugi, T. Ebisuzaki, Nanoscale hydrophobic interaction and nanobubble nucleation, Phys. Rev. Lett. 93 (2004) 185701.
[27] D.A. Case, D.A. Pearlman, J.W. Caldwell III, T.E. Cheatham, W.S. Ross, C. Simmerling, T. Darden, K.M. Merz, R.V. Stanton, A.L. Cheng, J.J. Vincent, M. Crowley, V. Tsui, R.J. Radmer, Y. Duan, J. Pitera, G.L. Seibel, U.C. Singh, P.K. Weiner, P.A. Kollman, Amber 6.0, University of California, San Francisco, 1999.
[28] P.A. Kollman, R. Dixon, W. Cornell, T. Fox, C. Chipot, A. Pohorille, The development/application of a 'minimalist' organic/biochemical molecular mechanic force field using a combination of ab initio calculations and experimental data, in: A. Wilkinson, P. Weiner, W.F. van Gunsteren (Eds.), Computer Simulations of Biological Systems, vol. 3, Elsevier, 1997, pp. 83–96.
[29] J.-P. Ryckeart, G. Ciccotti, H.J.C. Berendsen, Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of *n*-alkanes, J. Comput. Chem. 23 (1997) 327–341.
[30] H.J.C. Berendsen, J.M.P. Postma, W.F. van Gunsteren, A. DiNola, J.R. Haak, Molecular dynamics with coupling to an external bath, J. Comput. Phys. 81 (1984) 3684–3690.
[31] C. Guilbert, D. Perahia, L.A. Mouawad, A method to explore transition paths in macromolecules—applications to hemoglobin and phosphoglycerate kinase, Comput. Phys. Comm. 91 (1995) 263–273.
[32] K. Kuznedelov, N. Korzheva, A. Mustaev, K. Severinov, Structure-based analysis of RNA polymerase function: the largest subunit's rudder contributes critically to elongation complex stability and is not involved in the maintenance of RNA–DNA hybrid length, EMBO J. 21 (2002) 1369–1378.
[33] G. Bar-Nahum, V. Epshtein, A.E. Ruckenstein, R. Rafikov, A. Mustaev, E. Nudler, A ratchet mechanism of transcription elongation and its control, Cell 120 (2005) 183–193.
[34] V. Epshtein, A. Mustaev, V. Markovtsov, O. Bereshchenko, V. Nikiforov, A. Goldfarb, Swing-gate model of nucleotide entry into the RNA polymerase active center, Mol. Cell 10 (2002) 623–643.
[35] E. Nudler, A. Mustaev, E. Lukhtanov, A. Goldfarb, The DNA–RNA hybrid maintains the register of transcription by preventing backtracking of RNA polymerase, Cell 89 (1997) 33–41.