



Conformational selection and induced fit for RNA polymerase and RNA/DNA hybrid backtracked recognition

Jian Wu^{1†}, Wei Ye^{1†}, Jingxu Yang¹ and Hai-Feng Chen^{1,2*}

¹ State Key Laboratory of Microbial Metabolism, Department of Bioinformatics and Biostatistics, College of Life Sciences and Biotechnology, Shanghai Jiaotong University, Shanghai, China, ² Independent Institute, Shanghai Center for Bioinformation Technology, Shanghai, China

OPEN ACCESS

Edited by:

Wei Yang,
Florida State University, USA

Reviewed by:

Guohui Li,
Chinese Academy of Sciences, China
Ray Luo,
University of California, Irvine, USA

*Correspondence:

Hai-Feng Chen
haifengchen@sjtu.edu.cn

[†]These authors have contributed
equally to this work.

Specialty section:

This article was submitted to
Molecular Recognition,
a section of the journal
Frontiers in Molecular Biosciences

Received: 02 September 2015

Accepted: 11 October 2015

Published: 05 November 2015

Citation:

Wu J, Ye W, Yang J and Chen H-F
(2015) Conformational selection and
induced fit for RNA polymerase and
RNA/DNA hybrid backtracked
recognition. *Front. Mol. Biosci.* 2:61.
doi: 10.3389/fmolb.2015.00061

RNA polymerase catalyzes transcription with a high fidelity. If DNA/RNA mismatch or DNA damage occurs downstream, a backtracked RNA polymerase can proofread this situation. However, the backtracked mechanism is still poorly understood. Here we have performed multiple explicit-solvent molecular dynamics (MD) simulations on bound and apo DNA/RNA hybrid to study backtracked recognition. MD simulations at room temperature suggest that specific electrostatic interactions play key roles in the backtracked recognition between the polymerase and DNA/RNA hybrid. Kinetics analysis at high temperature shows that bound and apo DNA/RNA hybrid unfold via a two-state process. Both kinetics and free energy landscape analyses indicate that bound DNA/RNA hybrid folds in the order of DNA/RNA contracting, the tertiary folding and polymerase binding. The predicted Φ -values suggest that C7, G9, dC12, dC15, and dT16 are key bases for the backtracked recognition of DNA/RNA hybrid. The average RMSD values between the bound structures and the corresponding apo ones and Kolmogorov-Smirnov (KS) *P*-test analyses indicate that the recognition between DNA/RNA hybrid and polymerase might follow an induced fit mechanism for DNA/RNA hybrid and conformation selection for polymerase. Furthermore, this method could be used to relative studies of specific recognition between nucleic acid and protein.

Keywords: backtracked RNA polymerase, DNA/RNA hybrid, induced fit, conformational selection, *P*-test

INTRODUCTION

RNA polymerase efficiently catalyzes the template-directed transcription with a low error rate (Shaevitz et al., 2003). This high fidelity may be the result of a “proofreading” mechanism. Backtracked transcribing complex plays a key role in the proofreading mechanism of transcription (Jeon and Agarwal, 1996; Thomas et al., 1998; Shaevitz et al., 2003; Kostrewa et al., 2009; Wang et al., 2009; Martinez-Rucobo et al., 2011). RNA polymerase does not move forward monotonously. It oscillates between three stable states to move forward or backward at every step in transcription (Thomas et al., 1998; Toulmé et al., 1999; Darzacq et al., 2007; Wang et al., 2009). If the forward movement is blocked, for example, by DNA/RNA mismatch or DNA damage, a backtracked state of the transcribing complex will form. Then the transcription blocking may be recovered by transcription factor SII (TFIIS) in eukaryotes and GreA (and/or GreB) in prokaryotes (Wind and Reines, 2000; Fish and Kane, 2002; Shilatfard et al., 2003; Sims et al., 2004).

Besides naturally occurrence, ultraviolet light, toxic chemical agents, and ionizing radiation could cause DNA damage with a relatively higher possibility (Jackson and Bartek, 2009). DNA damage would lead to many kinds of cancers (Jackson and Bartek, 2009). In addition, DNA damage and its repair are important hypotheses of aging mechanism (Hoeijmakers, 2009). There are several ways of DNA damage repair, including nucleotides excision induced by backtracked RNA polymerase II (Hanawalt, 2002; Foustieri and Mullenders, 2008; Wang et al., 2009), cell cycling arresting or apoptosis induced by p53 (Wang, 1998; Cortez et al., 2001; Kruse and Gu, 2009) or Chk2 (Lim et al., 2000; Li and Zou, 2005). Therefore, it is critical to reveal the specific recognition between the backtracked RNA polymerase II and damaged DNA/RNA hybrid.

In the backtracked state, RNA polymerase retreats on the template, extruding the 3'-end of the RNA (Wang et al., 2009). In eukaryotes, TFIIS binds to backtracked RNA polymerase II and stimulates a ribonuclease activity which cuts off the 3'-end of nascent RNA. As a result, RNA polymerase II elongation complex can repeatedly shorten and extend the nascent RNA, which is referred as cleavage and resynthesis (Reines et al., 1993). Therefore, RNA polymerase timely recovers the blocked transcription.

The X-ray structure of backtracked RNA polymerase II and DNA-RNA hybrid complex was released in 2009 (pdb code: 3GTQ) (Wang et al., 2009). The complex in this study consists of one chain of damaged DNA, one chain of 11 nt RNA, and one bridge helix from Thr809 to Glu846. The bridge helix is bound to the 3'-end of the RNA. In order to compare the difference between DNA damage and DNA mismatch, the complex between RNA polymerase II and mismatched DNA-RNA hybrid (pdb code: 3GTK) was also investigated. The complex includes 13 nt of the RNA with two mismatches at its 3'-end (C-dC and G-dG), 13 nt of the DNA, and the bridge helix (shown in **Table 1** and **Figure 1**).

X-ray experiments suggest that the uridine 5'-monophosphate base of RNA was paired with its complement in the DNA template but tilted 15°–20° out of the plane. The backbone of the backtracked RNA between the last match base and the mismatched base was bent over 120° out of the path of

the hybrid helix (Wang et al., 2009). Therefore, mismatched RNA/DNA undergoes significant conformational adjustment through specific recognition. However, the recognition between RNA/DNA hybrid and polymerase is poorly understood. Molecular dynamics (MD) simulation may be a powerful tool for analyzing the structural and dynamic features of biomacromolecules. Several simulations were performed on RNA polymerase II. Suenaga et al. (2006) did whole molecule MD simulation on normal RNA polymerase II and DNA/RNA hybrid to discover the elongation mechanism in 2006. The transcription bubble is formed and maintained by multiple protein loops. MD simulation on backtracked RNA polymerase II and mismatched DNA/RNA hybrid was mentioned in work of Wang et al. (2009). The two nanoseconds' simulation reveals that the mismatched bases are highly mobile and unstable. Here we focus on the folding kinetics of DNA/RNA hybrid upon polymerase-binding. Our research intends to reveal a series of interesting questions. (1) How does the conformation of the DNA/RNA hybrid change upon the binding of RNA polymerase II? (2) Which mechanism does the folding of DNA/RNA and polymerase obey? (3) How does the enzyme recognize the DNA damage site of the DNA/RNA hybrid? To answer these questions, we utilize atomic molecular dynamics simulation in explicit solvent to analyze the coupling mechanism between DNA/RNA folding and polymerase binding.

However, all atomic MD simulations are currently restricted to the timescale of less than 1 μ s, which is much shorter than the folding half-time of most biomacromolecules at room temperature (at least 1 ms) (Baker, 1998; Fersht and Daggett, 2002). Fortunately, the rate of unfolding can accelerate by approximately six orders of magnitude at high temperature (usually 498 K) (Shea and Brooks, 2001), so most biomacromolecules unfold in the nanosecond's time scale at this temperature (Fersht and Daggett, 2002). Furthermore, experiment confirms the transition state for folding and unfolding is expected to be the same from the principle of microscopic reversibility (Fersht and Daggett, 2002). Based on these previous works, unfolding simulations at high temperature have been used in the current study. In this paper, we carried out multiple molecular dynamics ran in explicit water (Interlandi et al., 2006) to study the folded states and the unfolding kinetics for both bound and apo-states to understand the specific recognition between DNA/RNA hybrid and RNA polymerase.

TABLE 1 | Information of DNA-RNA hybrid and polypeptide.

Model	DNA-RNA hybrid	Polypeptide
DNA/RNA hybrid with DNA damage	Apo RNA 5' 1----6----11 AUGGCGGGAG 3' DNA Bound DNA 3' CTACCTCTCCTC 5' --21--16--	/
Protein	/	809--814--819--824-- H ₂ N ⁺ -TPQEFFFHAMGGREGLIDT AVKTAETGYIQRRLVKALE-COO ⁻ -829--834--839--844
Bound DNA/RNA hybrid with mismatches	1----6----11- RNA 5' AUCGAGAGGAUGC 3' : DNA 3' TAGCTCTCCTAGC 5' 26-21-16-	

METHODS

Molecular Dynamics Simulations

The atom coordinates of DNA/RNA-polymerase complex were obtained from the Protein Data Bank, with the accessory code 3GTQ (Wang et al., 2009). All simulations and most analyses procedures were conducted using the AMBER11 software package (Case TAD TEC et al., 2010). Hydrogen atoms were added using the LEaP module of AMBER11. Counter-ions were used to maintain system neutrality. All systems were solvated in a truncated octahedron box of TIP3P waters with a buffer of 10 Å. Particle Mesh Ewald (PME) (Darden et al., 1993) was

employed to treat long-range electrostatic interactions with the default setting in AMBER11. The improved parm99SBildn force field (Lindorff-Larsen et al., 2010) was used for the intramolecular interactions. The SHAKE algorithm (Ryckaert et al., 1977) was used to constrain bonds involving hydrogen atoms. 1000-step steepest descent minimization was performed to relieve any structural clash in the solvated systems. This was followed by heating up and brief equilibration for 20 ps in the NVT ensemble at 298 K with PMEMD of AMBER11. Langevin dynamics with a time step of 2 fs were used in the heating and equilibration runs with a friction constant of 1 ps^{-1} .

Folded state can be used to compare conformational differences between different systems, to determine inter-/intra-chain interactions and to predict the recognition mechanism. A previous work suggests that a small number of trajectories of MD simulation (5–10) is sufficient to capture the average properties of the protein (Day and Daggett, 2005). Therefore, to study the folded state of each solvated system, six independent trajectories of 10.0 ns each in the NPT ensemble at 298 K were then simulated with PMEMD of AMBER8. Six

independent trajectories of 15.0 ns each were performed for the apo-DNA/RNA, apo-polymerase, and DNA/RNA-polymerase complex in the NVT ensemble at 498 K to investigate the unfolding kinetics of each solvated system. Trajectories of 450 ns total were collected for three systems (DNA/RNA-polymerase complex, apo-DNA/RNA, and apo-polymerase) at both 298 K and 498 K, taking about 41,000 CPU hours on the Xeon (3.0 GHz) cluster. Detailed simulation information was listed in **Table 2**.

Native contacts between DNA/RNA hybrid and polymerase were monitored to determine the unfolding kinetics. Kinetics analysis suggested that 10 ns at 498 K were needed to reach the equilibrium stage for the bound and apo states, so that the first 10 ns (a total of 60 ns) were used to study the unfolding kinetics.

Data Analysis

Tertiary contact assignment was handled with in-house software (Chen and Luo, 2007; Chen, 2008; Qin et al., 2012; Ye et al., 2012a,b). Residues and nucleotides are in hydrophobic contact when mass centers of their side chains are closer than 6.5 \AA

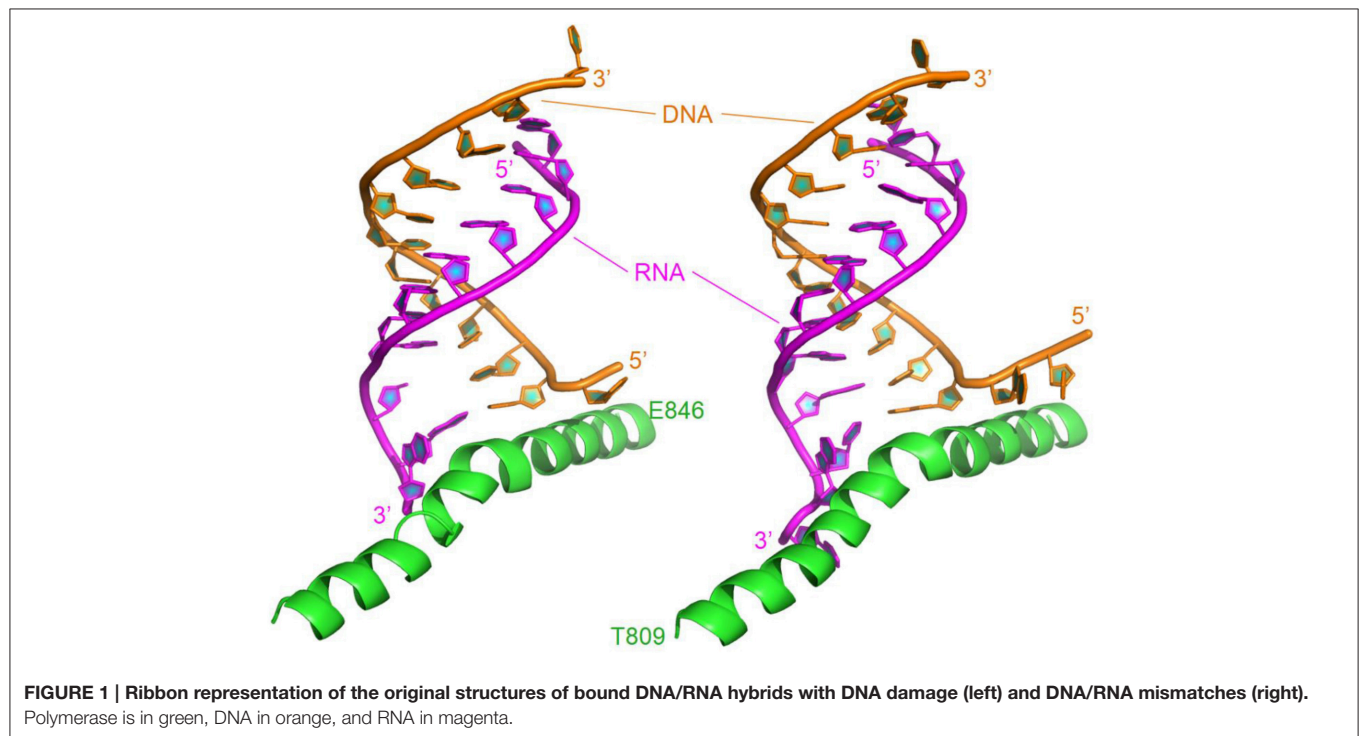


TABLE 2 | Simulation condition for five systems.

Model	Number traj	Simulation time (ns)		Water molecules number	Water box size (\AA^3)	Initial density (g/mL)	
		298 K	498 K				
DNA damage	Bound	6	10.0	15.0	12789	445921	0.903
	Apo	6	10.0	15.0	6572	237294	0.883
	Protein	6	10.0	15.0	8566	158607	0.894
Mismatch	Bound	6	10.0	–	11779	418192	0.895
	Apo	6	10.0	–	7933	283023	0.890

for the complex. Electrostatic (i.e., charge-charge) interactions are assigned when the distance between the mass center of positive charge residue and the DNA/RNA hybrid phosphate backbone is less than 11 Å. Hydrogen bond is defined that the distance between two polar heavy atoms either of which has a hydrogen atom are less than 3.5 Å. The non-adjacent nucleotides and residues are in contact (termed “native contact”) when their side chains were closer than 7.5 Å. All the 3D molecular representations were visualized and rendered with PyMOL 1.5.

The energy landscape was performed by calculating normalized probability from a histogram analysis (Pande and Rokhsar, 1999). Here we used RMSD and the radius of gyration (Rg) to map the energy landscape. Unfolding kinetics were reflected and fitted in 2 indices: (1) Qf, the fraction of intra-chain (tertiary) native contacts in unfolding state, comparing to folded state (unfolding kinetics); and (2) Qb, the fraction of inter-chain native contacts (unbinding kinetics). Binding and folding pathways could be restored as a reverse procedure of unfolding pathway. The free energy landscapes and unfolding kinetics was plotted and fitted in Origin 8.6.

Landscapes were also used to analyze the distance differences to detect the structural changes between bound and apo-DNA/RNA hybrid. Distance between every pair of C5' (DNA/RNA) or Cα (protein) was calculated in both bound and apo structure, respectively. Then the landscape of base-to-base (or residue-to-residue) distances in apo-structure minus the corresponding distances in bound structure was plotted. This indicates a structural contraction when the color in red is positive and an extension when the color in blue is negative. In order to exclude the effect of thermal fluctuations, we averaged all the distances in the whole 10 ns' simulation instead of a certain time point.

According to the definition of protein Φ-values which computed with a strategy similar to those used in other studies (Cafilisch and Karplus, 1994; Vendruscolo et al., 2001; Gsponer and Cafilisch, 2002), this method might be suitable for the Φ-values calculation for DNA/RNA hybrid. Therefore, the equation below was used to process the Φ-values of DNA/RNA hybrid.

$$\Phi_{\text{DNA/RNA}}^{\text{calc}} = \frac{N_i^{\text{TS}} - N_i^{\text{U}}}{N_i^{\text{F}} - N_i^{\text{U}}}$$

N_i^{TS} is the number of native contacts for base i at transition state, N_i^{F} and N_i^{U} is the number of native contacts for base i at folded and unfolded states, respectively.

Binding Mechanism Evaluation

There are two mainstream hypotheses to explain binding process: “Induced fit” model (Koshland, 1958; Boehr and Wright, 2008) and “Conformational selection” model (Ma et al., 1999, 2002; Kumar et al., 2000; Tsai et al., 2001; Boehr and Wright, 2008; Tsai and Nussinov, 2011). The former model emphasizes that the conformational change of one part in the complex is induced by the binding of the other part (binding first, Turjanski et al., 2008), while the latter points out that one part selects the most optimal ligand among several chaperones with different conformations (Lange et al., 2008). If the folding obeys a mechanism of induced

fit, a significant conformational change must occur upon the ligand binding, especially near the binding site. Otherwise, if the conformational change is not so significant, the binding tends to obey a conformational selection mechanism (Wlodarski and Zagrovic, 2009). Recent studies have demonstrated that the binding process may employ a mixed mechanism, selecting an optimal conformation globally and inducing the local structure of the ligand to be adjusted (Anthis et al., 2011; Bucher et al., 2011; Silva et al., 2011). Following the concepts above, induced fit could be calculated from the RMSD between a bound structure and its most “like” apo-structure (which has the lowest RMSD from the bound structure). In the induced-fit model, relative magnitude of conformational selection could be characterized from the average RMSD between this most similar apo-structure and the other apo-structures (Wlodarski and Zagrovic, 2009). In this study, binding site for the ligand was defined as the mass center of the receptor. RMSD for induced fit of 6 trajectories was given as a function of distance from the binding site. Because the RMSD values are not normally distributed, which is not suitable for a t -test, a standard two-sample Kolmogorov-Smirnov (KS) test was used to evaluate the statistical significance of RMSDs' variations. The equation below was used to quantitatively describe relative differences between conformational selection and induced fit.

$$\Delta = \frac{\sum_{X_i, f_i \in \text{CS}} X_i f_i}{N_{\text{CS}}} - \frac{\sum_{X_j, f_j \in \text{IF}} X_j f_j}{N_{\text{IF}}}$$

X is the RMSD value and f is the frequency for a particular RMSD value region in a distribution of conformational selection (CS) or induced fit (IF), and N_{CS} and N_{IF} are the numbers of data points for the distribution of CS or IF, respectively.

Transition State Identification

The transition state reflects a representative snapshot at the free energy maxima of the unfolding and folding pathways. The structures at the free energy maxima constitute the transition state ensemble (TSE). TSE structures can either fold or unfold, and the transition probability (P) will be 50%. In order to determine TSE, we have scanned MD snapshots for TSE structures in all six unfolding trajectories for each of bound DNA/RNA and apo-DNA/RNA, respectively (Pande and Rokhsar, 1999). The transition probability curves are further fitted by the Boltzmann Equation,

$$P = \frac{1}{1 + e^{\frac{t - \tau_{\text{TS}}}{\tau_{\text{trans}}}}}$$

P is the transition probability, t is simulation time, τ_{TS} is the fitted time when $P = 50\%$, and τ_{trans} sets the period when P is ranging from 0.4 to 0.6 (Pande and Rokhsar, 1999; Gsponer and Cafilisch, 2002; Chong et al., 2005). Average TS structure was selected from TSE near τ_{TS} of each trajectory.

RESULTS

To collect enough snapshots for statistically meaningful structural analysis, six trajectories of 10.0 ns were collected

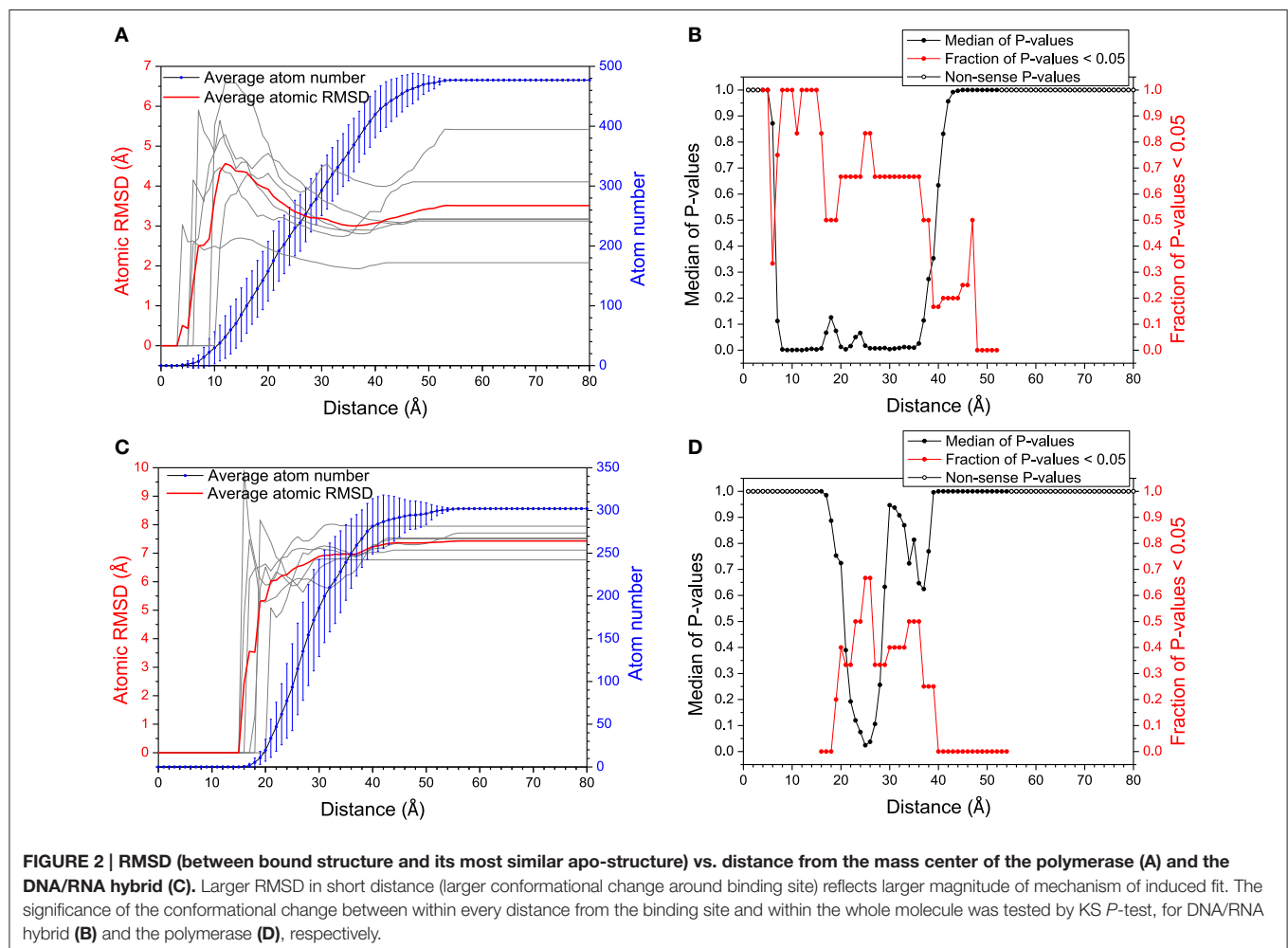
for apo- and polymerase-bound DNA/RNA hybrid with DNA damage. Additionally, bound DNA/RNA hybrid with mismatches was also simulated. Average $C\alpha$ RMSDs with respect to simulation time (shown in Figure S1) indicate that 10 ns' simulation time at 298 K was sufficient for the three systems to equilibrium in the fully solvated environment. Bound DNA/RNA hybrid shows a lower RMSD than the apo-structure, which reflects a stabilizing role of the binding between the two components.

The average structures of bound and apo DNA/RNA hybrid are shown in Figure S1. The 3'-end of the RNA was extruded by the backtracked protein and the 5'-end of the DNA was incurved. The protein lost some α -helices at the N-terminal and rotated about 30°. Similar changes can be found in the complex of RNA/DNA mismatch.

Induced Fit or Conformational Selection in Polymerase and DNA/RNA Recognition

The mechanism of specific binding is one of the most attractive issues in our study. Recently, residual dipolar coupling experiment suggests that the folding of ubiquitin complex obeys conformational selection rather than induced-fit mechanism

(Lange et al., 2008). Nevertheless, both mechanisms have been observed in the same system (James and Tawfik, 2003; Okazaki and Takada, 2008). To evaluate the mechanism of specific recognition between DNA/RNA hybrid and polymerase from MD simulation, **Figure 2** shows the average atom RMSD values between the bound structures and its most "like" apo structure vs. the distance from DNA/RNA or polymerase binding site. The RMSD gradually decreases for DNA/RNA hybrid, however, the RMSD of polymerase has the propensity of increases. These observations suggest DNA/RNA adjusts the conformation near the binding site. This conformational changes were then tested by two-sample Kolmogorov-Smirnov (KS) P -test (Włodarski and Zagrovic, 2009) to see the significance of these structural differences. The conformational difference is statistically significant up to ≈ 30 Å away from the polymerase binding site with the median P -values typically < 0.05 and the fraction of typical $P > 0.5$, which suggests that the local conformational changes are significant for DNA/RNA. Therefore, bound DNA/RNA hybrid might obey an induced fit mechanism. However, the conformational difference is not significant for the polymerase which can be indicated with the low fraction with $P < 0.05$ for the whole distance range. Different



to DNA/RNA hybrid, induced fit may not so significant on the polymerase.

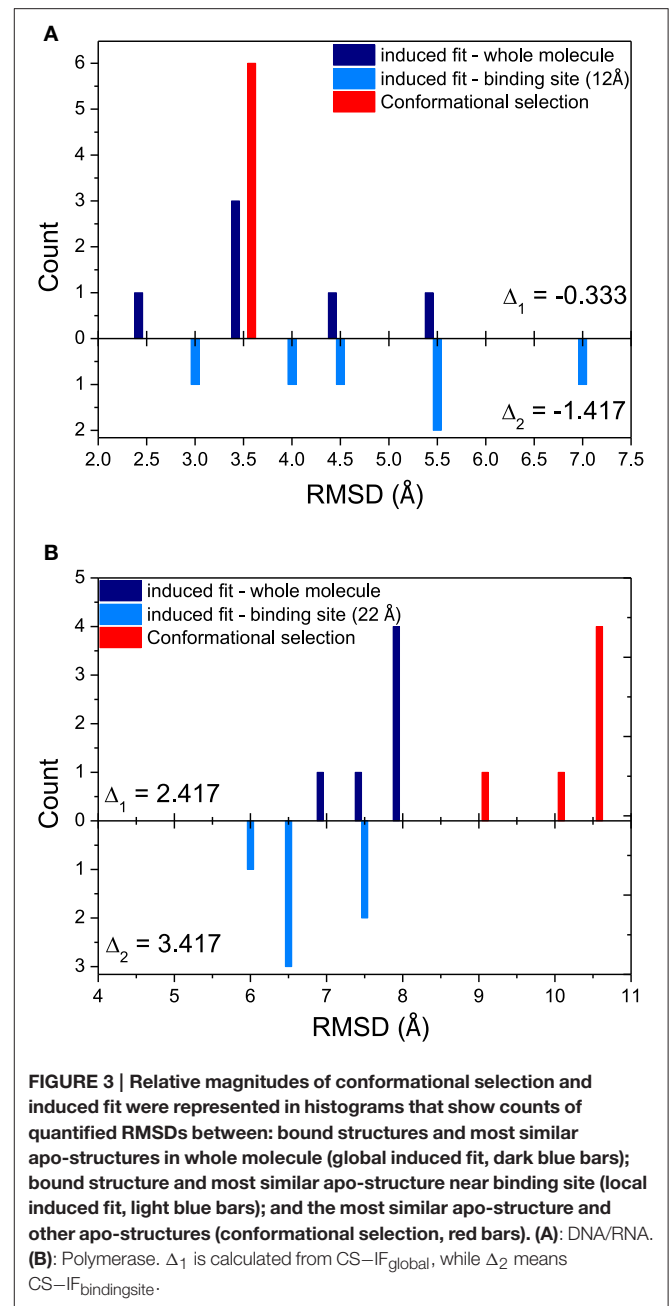
Since both mechanisms would exist in the binding process, relative magnitude for these mechanisms is another interesting issue in our study. The histograms comparing the magnitude of conformational selection with induced fit at 12 Å (just the binding site) and 40 Å (the whole molecule) for DNA/RNA hybrid and polymerase, and shown in **Figure 3**. Parameter Δ represents the difference between conformational selection and induced fit, Δ_1 for the whole molecule, Δ_2 for the binding site. For DNA/RNA hybrid, both Δ_1 and Δ_2 are negative. This indicates that the magnitude of induced fit is higher than that of conformational selection. That is, the folding of DNA/RNA hybrid might predominantly follow an induced fit mechanism upon the protein binding. Consistent as the observation above, protein might obey different mechanisms from the view of the positive value for both whole molecule and the binding site. Polymerase might follow a conformational selection mechanism when binding onto DNA/RNA hybrid.

Distance Analyses

Distance analyses could be used to further confirm the conformation adjustment in the binding process. C5' and C α fluctuations for bound and apo states are illustrated in **Figure 4**. The C α fluctuation of bound polymerase is significantly lower than that of apo-polymerase, especially in the region of Residue 829–846, which suggests that polymerase becomes more stable upon DNA/RNA binding. The conformations of 3'-end of RNA and 5'-end of DNA for bound DNA/RNA hybrid also have significant change upon the interaction of polymerase.

To investigate the relative movement in the binding process, distances from last 4 base pairs of bound DNA/RNA hybrid to the mass center of the protein are analyzed and shown in **Figure 5**. The last base pair (G11-dC12) has the propensity of constriction, however other three base pairs keep stable from each other. From the result of distance difference (**Figure 5B**), the relative location of protein between DNA and RNA has large fluctuation and keeps dynamics equilibrium. Protein keeps much closer to the G11 of RNA than to dG12 of DNA. This is consistent with the previous work (Gnatt et al., 2001).

The landscape of distance difference for base pairs between apo and bound DNA/RNA is shown in **Figure 6A**. Distance difference landscape can reflect the relative conformational change of the DNA/RNA backbone. The deep red areas indicate that the distance difference for bases dC12-U2, dC12-C3, and dC12-G4 is positive value, which suggests that these base pairs are contracted upon polymerase-binding. The blue areas represent that the distance difference is negative value. It suggests that dC23 of DNA becomes far away from G11 of RNA. The distance landscape for residue pairs between apo and bound polymerase is shown in **Figure 6B**. The result shows that the most regions of polymerase backbone did not undergo an obvious conformational change except the residue pairs of Glu831/Gln837, Glu831/Val842, Glu831/Leu845, and Phe814, Phe815, His816 with Arg840, Lys843. These regions might play a key role in the formation of DNA/RNA/polymerase complex.



Driving Forces for the Conformation Adjustment

The conformation adjustment in the binding process could be attributed to the electrostatic, hydrophobic and hydrogen-binding interactions between DNA/RNA and polymerase. The electrostatic interactions in the simulation of six trajectories are shown in **Figure 7**. There are 12 electrostatic interactions between positively charged amino acids and the phosphates of DNA/RNA with populations higher than 30%. The positive charged residues, such as Arg821, Arg839, Arg840, and Lys843, provide electrostatic interacts with the phosphates of the DNA/RNA terminal. More than 80% electrostatic interactions are located between polymerase and DNA. The hydrophobic

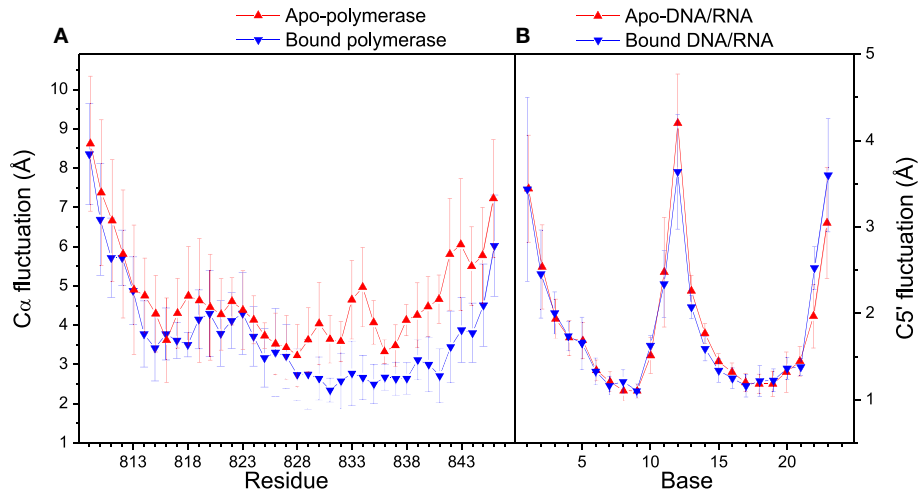


FIGURE 4 | C_{α} and $C5'$ fluctuations in folded state. (A) Comparison of C_{α} fluctuations between apo- (red) and bound (blue) polymerase; **(B)** Comparison of $C5'$ fluctuations between apo- (red) and bound (blue) polymerase. Standard deviations among six trajectories are displayed with error bars.

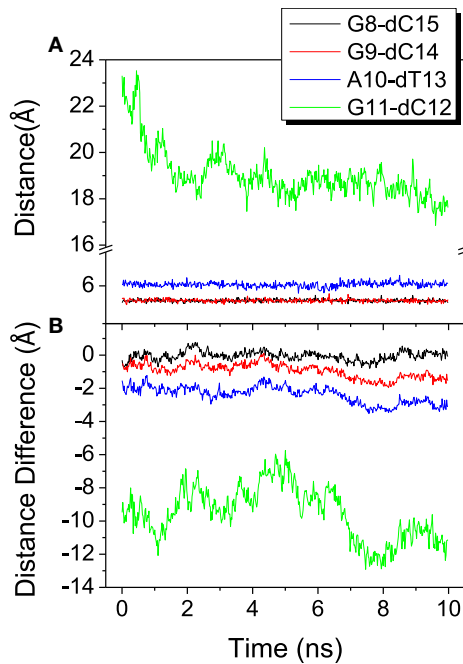


FIGURE 5 | Distance among the mass center of the polymerase and the last 4 base pairs of the DNA/RNA hybrid. (A) Distance between each base pair. **(B)** Distance difference for each base pair, calculated from distance of RNA to protein minus that of DNA to protein.

contacts in the simulation of six trajectories are also shown in **Figure 7**. Two stable hydrophobic interactions can be found: Ala832-dT13 and Ala828-G11, with populations higher than 30%. **Figure 7** also shows two hydrogen bonds with populations higher than 30%. These hydrogen bonds are between residues Arg839 and Watson-Crick edge of dC14. Comparing among these three kinds of interactions, electrostatic interactions is in predominance, which suggests that the electrostatic interactions

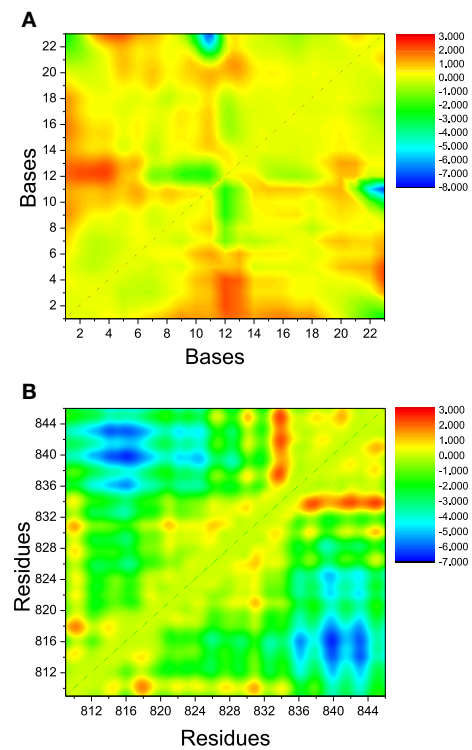
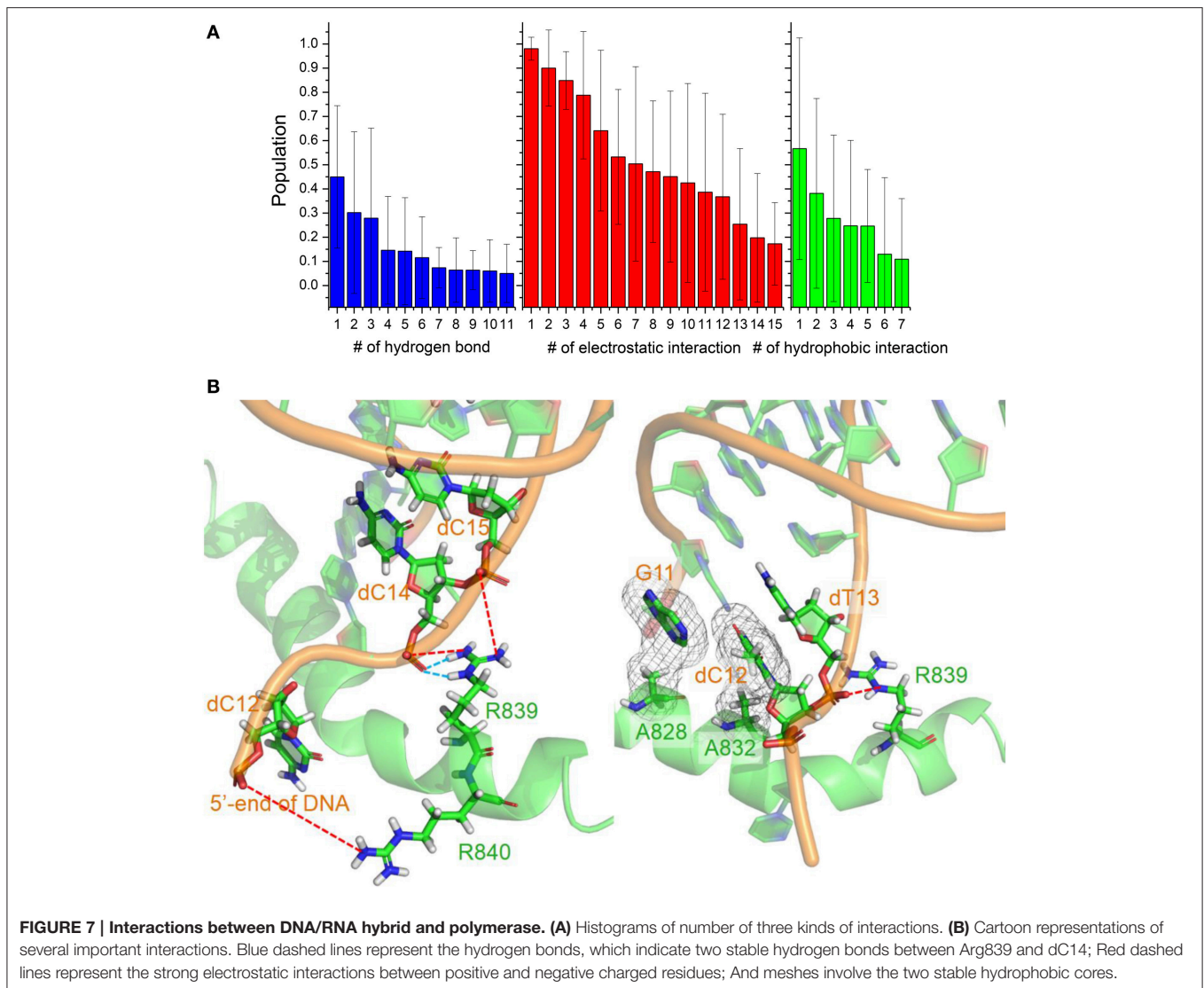


FIGURE 6 | The landscape of distance difference between bound and apo states. (A) DNA/RNA, **(B)** polymerase. The color in red represents positive distance difference. The color in blue represents negative distance difference.

play key roles in the backtracking and proofreading of polymerase.

Unfolding Kinetics

Unfolding state along high temperature simulations was performed to detect the binding and folding details of DNA/RNA



hybrid and the polymerase. The fraction of tertiary native contacts (Q_f) and binding native contacts (Q_b) is used to monitor unfolding and unbinding kinetics, respectively. The stable native contacts were calculated for the folded state, including inter-chain and intra-chain. Q_f is defined as the fraction of intrachain native contacts which are still stable at unfolding state; Similarly, Q_b is the fraction of interchain native contacts at 498 K. Time evolutions of Q_b and Q_f for bound DNA/RNA are shown in **Figure 8**, which suggests that the tertiary unfolding and unbinding kinetics can be represented well by single exponential functions, indicating first order kinetics in the NVT ensemble at high temperature. The fitted kinetics data are listed in **Table 3**. The kinetics analysis shows that the unbinding half-time is 1.21 ± 0.08 ns and the tertiary unfolding half-time is 1.71 ± 0.04 ns. This indicates that the unbinding is slightly faster than the tertiary unfolding for bound DNA/RNA. The time evolution of Q_f for apo-DNA/RNA is also shown in **Figure 8**. It is found that the tertiary unfolding of apo-DNA/RNA also obeys first

order kinetics, with a half-time of 1.85 ± 0.04 ns, which is a little slower than the bound DNA/RNA. Time evolution of the radius of gyration (R_g) for bound DNA/RNA is also presented in **Figure 8**. The extension of bound DNA/RNA obeys first order kinetics, and the half-time is 4.24 ± 0.40 ns. This indicates that the extension of bound DNA/RNA is slower than the tertiary unfolding and unbinding. There were 14 native contacts between the DNA and the protein at folded state. These native contacts sharply decreased during unfolding (5 in transition state and only 1 in unfolded state). There were 13 native contacts between the RNA and the enzyme at folded state, but only 2 in transition state. At unfolded state, RNA separated from the enzyme.

Unfolding Landscape

To further understand the interdependence between the DNA/RNA folding and polymerase binding, the unfolding landscape of bound DNA/RNA was analyzed with the variables

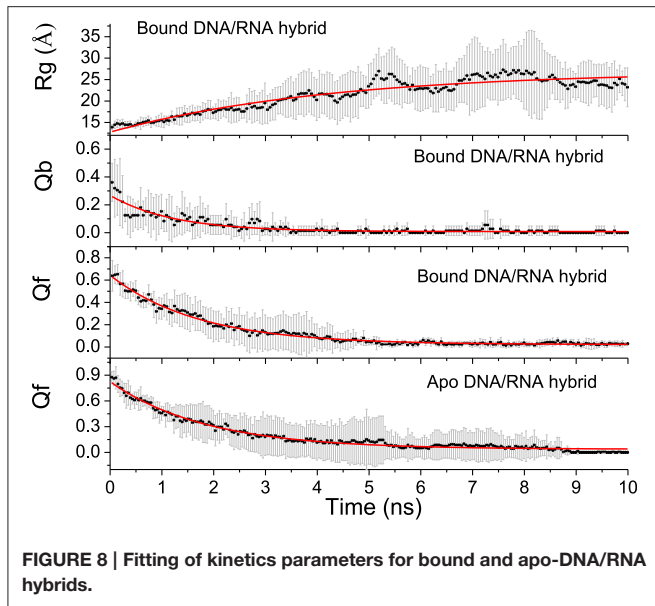


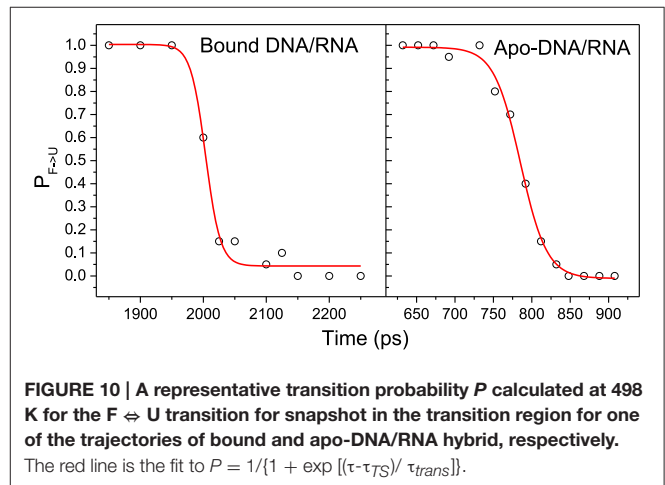
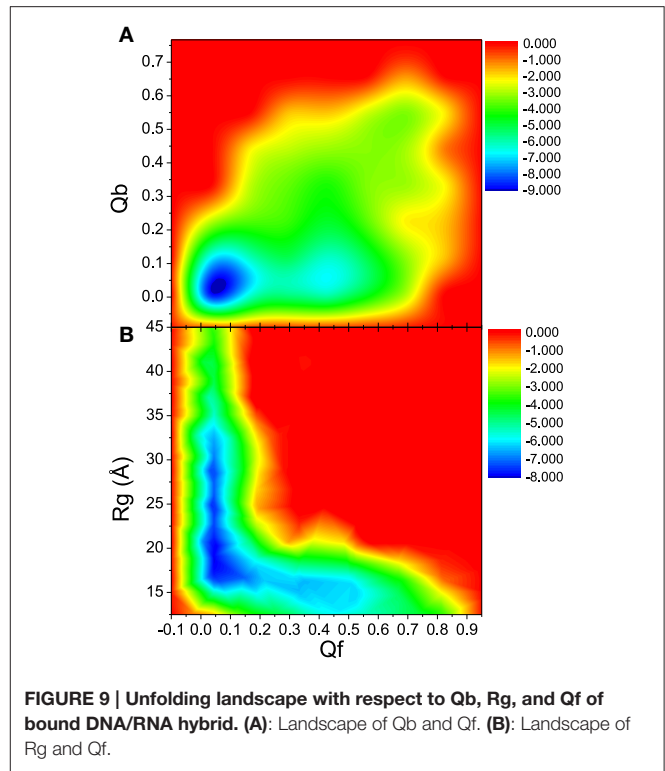
TABLE 3 | Unfolding kinetic constants for bound and apo DNA/RNA hybrid.

Structure	Property	τ (ns)	A	B	R^2
Bound DNA/ RNA hybrid	Rg	4.24 ± 0.40	-14.22 ± 0.42	26.96 ± 0.50	0.8852
	Qb	1.21 ± 0.08	0.26 ± 0.01	0.01 ± 0.00	0.8490
	Qf	1.71 ± 0.04	0.62 ± 0.01	0.02 ± 0.00	0.9835
Apo-DNA/ RNA hybrid	Qf	1.85 ± 0.04	0.79 ± 0.01	0.04 ± 0.00	0.9797

Qf and Qb, and shown in **Figure 9**. The unfolding landscape shows that the unbinding proceeds first while the tertiary contacts are held stable, and then followed by the tertiary unfolding. This is consistent with the unfolding kinetics analysis. The formation of the tertiary contacts might be important to the binding of polymerase. The coupling between the tertiary unfolding and the extension of DNA/RNA hybrid is also illustrated in **Figure 9**. It is found that Qf decreases first while the radius of gyration does not change, and then followed by the extension of DNA/RNA, which suggests that the extension of DNA/RNA is followed by the tertiary unfolding. This is also consistent with the unfolding kinetics analysis.

Transition State

Kinetics analysis shows that the unfolding of bound and apo-DNA/RNA obeys first order kinetics, which suggests that bound and apo-DNA/RNA unfold via a two state process. Therefore, there is a transition state ensemble (TSE) corresponding to the free energy maximum. TSE structures can either fold or unfold, and the transition probability (P) will be 50%. In order to determine TSE, we have scanned MD snapshots for TSE structures in all six unfolding trajectories for each of bound DNA/RNA and apo-DNA/RNA, respectively (Pande and Rokhsar, 1999). The transition probability curves are further fitted by the Boltzmann equation and shown in **Figure 10**. The



analysis yields 50 snapshots for the bound DNA/RNA and 61 snapshots for the apo-DNA/RNA TSE, respectively.

The average structures of TSE for bound and apo states are shown in Figure S2. There are 22 native contacts within bases for bound DNA/RNA TSE and 18 native contacts for apo-DNA/RNA TSE. Apparently, it can be concluded that the TSE of bound DNA/RNA is more native-like than that of apo-DNA/RNA. The stability of the complex is the foundation for backtracking.

Φ -Value Prediction

Φ -values have been widely used to determine key residues in the protein folding by theoretical and experimental investigations (Fersht, 2000; Fersht and Daggett, 2002; Fernández-Escamilla

et al., 2004; Fersht and Sato, 2004). In this study, all TSE structures were used to predict Φ -values of bases for bound and apo-DNA/RNA. The predicted Φ -values of bases for bound and apo-DNA/RNA are shown in **Figure 11**. Φ -Values of the bases at 3'-end of the RNA for bound DNA/RNA (C7, G9, dC12, dT13, dC15, and dT16), are larger than those for apo-DNA/RNA, which indicates these bases play a key role in the folding of bound DNA/RNA. This is consistent with the interactions analyses that these critical bases of G9, dC12, dT13, and dC15 form electrostatic interactions, native contacts and hydrogen bond network with polymerase. These predicted Φ -values could be further confirmed by experiments.

DISCUSSION

Three states of RNA polymerase are only different at the translocation of bridge-helix of the polymerase. Most interactions and specific recognitions are between the RNA's 3'-end of DNA/RNA hybrid and the bridge-helix of the polymerase. In molecular modeling and MD simulations, it is quite normal to simplify the system that is object of study to make it more computationally affordable. For the same reason, we selected the core piece of bridge helix, Thr809 to Glu846, which is the most important polypeptide in backtracked polymerase.

Comparison with Experiments

Many experiments released the structural basis of transcription for RNA polymerase II, template DNA, and nascent RNA (Cramer et al., 2000, 2001; Gnatt et al., 2001; Shaevitz et al., 2003; Hahn, 2004; Kettenberger et al., 2004; Wang et al., 2009). The main observation is focused on high electron density at the 3'-end of the RNA, indicating the extrusion and instability of the 3'-terminal of RNA. Gnatt et al. (2001) reported that the specific recognition for RNA polymerase II backtracking may prefer RNA to DNA. In our simulation, RNA polymerase approaches closer to the 3'-end of the RNA than to the 5'-end of the DNA, which is in agreement with the experimental observations.

The structural analyses have shown that Thr831 and Ala832 of polymerase are the critical residues in stabilizing the complex. Our room temperature simulations detected a stable hydrophobic interaction for Ala832/dT13 and two relatively

weak hydrophobic interactions (Ala832/dC12 and Ala832/A10) with a population around 25%. For Thr831, there are two stable native contacts (Thr831/A10 and Thr831/G11) with the population about 55%. This is consistent with the observation of X-ray experiment.

Comparison between DNA Damage and DNA/RNA Mismatch

Backtracked state could also be caused by the nascent mismatched RNA. In order to compare DNA damage and DNA/RNA mismatch, six trajectories of 10.0 ns each were also simulated for the DNA/RNA mismatch. In the distance analyses (Figure S3), the protein was found to retreat and approach two mismatched base pairs of RNA. Polymerase also moves closer to the 3'-end of RNA than the 5'-end of DNA. This is consistent with the distance analysis for DNA damage.

Interactions between the protein and the mismatched DNA/RNA hybrid were also analyzed and shown in Figure S4. There are 12 stable electrostatic, 4 stable hydrophobic, and 1 hydrogen bond interactions between DNA/RNA and polymerase, with populations higher than 30%. Similar to DNA damage, electrostatic interactions form the main interactions stabilizing the binding between the polymerase and mismatched DNA/RNA hybrid. Also, these non-specific interactions play a key role in recognition and backtracking of the enzyme. Average structure of mismatched DNA/RNA hybrid was also analyzed and shown in Figure S1. This figure suggests that the 5'-end of the DNA was incurved and the 3'-end of the RNA was extruded. This is also in agreement with the structural analysis.

Comparison with Other Simulations

The RMSD for A10, G12, and C13 of bound DNA/RNA mismatch was analyzed respectively, and shown in **Figure 12**. The C5' RMSD of A10 is the lowest with small fluctuation. For mismatched G12 and C13, the C5' RMSD is significant high with large fluctuation. This result is qualitatively consistent with the previous simulation (Wang et al., 2009). Furthermore, low stability of mismatched DNA/RNA hybrid leads to easier excision of the nascent and mismatched ribonucleotides for polymerase.

In our simulation, the RMSF of bound polymerase is significant smaller than that of apo polymerase. This is also

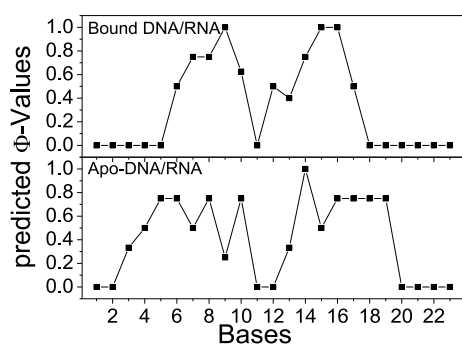


FIGURE 11 | Predicted Φ -values of bases for bound and apo-DNA/RNA hybrid, respectively.

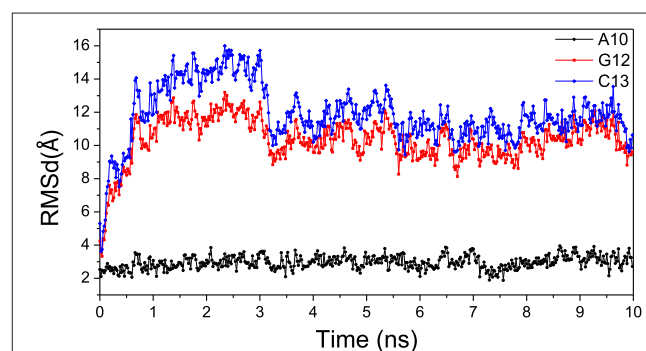
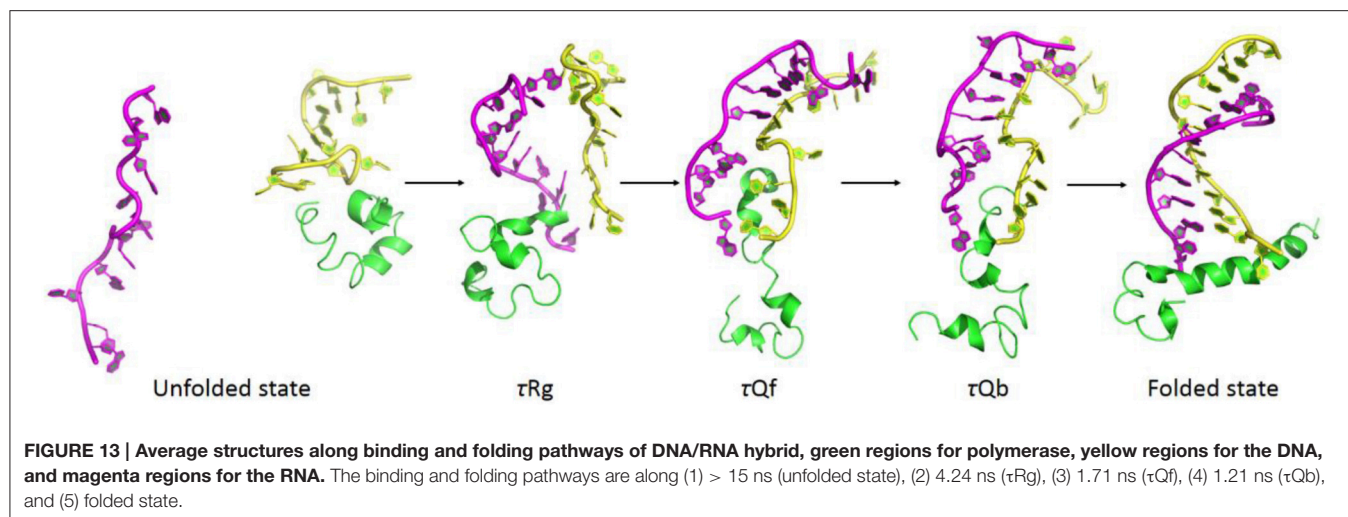


FIGURE 12 | RMSD for A10, G12, and C13 of the RNA in bound mismatch DNA/RNA.



consistent with the previous work that RMSFs of free polymerase are ranging from 0.6 to 2.3 Å and bound polymerase from 0.5 to 1.3 Å (Suenaga et al., 2006).

Further Discussions on Binding and Folding Pathways

By reversing the unfolding pathway, the binding and folding process could be constructed as (1) unfolded state, (2) tertiary contraction (τ_{Rg} , 4.24 ns), (3) tertiary folding (τ_{Qf} , 1.71 ns), (4) tertiary binding (τ_{Qb} , 1.21 ns), and (5) folded state. Average structures along binding and folding pathway are shown in **Figure 13**. In the unfolded state, RNA was far away from the complex of DNA and the polymerase. After its binding onto polymerase (τ_{Rg}), G9 and G11 kept tightly interacting with the polymerase. G9 was considered as a key base from the results of interaction analyses and Φ -value prediction because of the electrostatic interaction between it and the polymerase. This observation also indicate that the binding process is initiated by the non-specific electrostatic interaction to get the RNA and the protein approach to each other, then is stabilized with the specific interactions (hydrogen bonding and hydrophobic interactions). Upon binding, DNA/RNA hybrid undergoes a significant conformational adjustment with these interaction. Then 3'-end of the RNA was extruded by the polymerase, which is the key step in the nascent RNA's cleavage and excision.

CONCLUSIONS

Both 298 K and 498 K molecular dynamics simulations are performed for bound and apo DNA/RNA hybrid. At folded state, conformation of bound DNA/RNA hybrid shows significant adjustment and becomes more unstable upon the polymerase binding, especially the 3'-end of the RNA. Kinetics analysis of MD simulations at 498 K shows that bound

and apo DNA/RNA hybrid unfold via a two-state process. Bound DNA/RNA hybrid folds in the order of DNA/RNA contracting, the tertiary folding, and the protein binding. C7, G9, dC12, dC15, and dT16 are key bases to bound DNA/RNA hybrid folding. Inter-chain interaction analyses and binding pathway construction show that the recognition is in the order of non-specific closing, specific binding and conformation adjustment. dC14, G9, A10, and G11 are key bases to the protein recognition. The average RMSD values between the bound structures and the most "like" apo-structure and Kolmogorov-Smirnov (KS) *P*-test analyses indicate that the DNA/RNA hybrid might follow an induced fit mechanism and the polymerase with the conformational selection in their binding process. Furthermore, this method could be used to relative studies of specific recognition between nucleic acid and protein.

ACKNOWLEDGMENTS

This work was supported by Center for HPC at Shanghai Jiao Tong University, by grants from the Ministry of Science and Technology of China (2012CB721003), the National High-tech R&D Program of China (863 Program) (2014AA021502), the National Natural Science Foundation of China (J1210047 and 31271403), the Innovation Program of the Shanghai Education Committee (Grants No. 12ZZ023), Medical Engineering Cross Fund of Shanghai Jiao Tong University (YG2013MS68 and YG2014MS47).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fmolb.2015.00061>

REFERENCES

- Anthis, N. J., Doucleff, M., and Clore, G. M. (2011). Transient, sparsely populated compact states of apo and calcium-loaded calmodulin probed by paramagnetic relaxation enhancement: interplay of conformational selection and induced fit. *J. Am. Chem. Soc.* 133, 18966–18974. doi: 10.1021/ja2082813
- Baker, D. (1998). Metastable states and folding free energy barriers. *Nat. Struct. Biol.* 5, 1021–1024. doi: 10.1038/4130
- Boehr, D. D., and Wright, P. E. (2008). Biochemistry. How do proteins interact? *Science* 320, 1429–1430. doi: 10.1126/science.1158818
- Bucher, D., Grant, B. J., and McCammon, J. A. (2011). Induced fit or conformational selection? The role of the semi-closed state in the maltose binding protein. *Biochemistry* 50, 10530–10539. doi: 10.1021/bi201481a
- Cafilisch, A., and Karplus, M. (1994). Molecular-dynamics simulation of protein denaturation—solution of the hydrophobic cores and secondary structure of barnase. *Proc. Natl. Acad. Sci. U.S.A.* 91, 1746–1750. doi: 10.1073/pnas.91.5.1746
- Case TAD TEC, D. A. III., Simmerling, C. L., Wang, J., Duke, R. E., Luo, R., Walker, R. C., et al. (2010). *AMBER 11*. 11 Edn. San Francisco, CA: University of California.
- Chen, H. F. (2008). Mechanism of coupled folding and binding in the siRNA-PAZ complex. *J. Chem. Theory Comput.* 4, 1360–1368. doi: 10.1021/ct800030s
- Chen, H. F., and Luo, R. (2007). Binding induced folding in p53-MDM2 complex. *J. Am. Chem. Soc.* 129, 2930–2937. doi: 10.1021/ja0678774
- Chong, L. T., Snow, C. D., Rhee, Y. M., and Pande, V. S. (2005). Dimerization of the p53 oligomerization domain: identification of a folding nucleus by molecular dynamics simulations. *J. Mol. Biol.* 345, 869–878. doi: 10.1016/j.jmb.2004.10.083
- Cortez, D., Guntuku, S., Qin, J., and Elledge, S. J. (2001). ATR and ATRIP: partners in checkpoint signaling. *Science* 294, 1713–1716. doi: 10.1126/science.1065521
- Cramer, P., Bushnell, D. A., Fu, J., Gnat, A. L., Maier-Davis, B., Thompson, N. E., et al. (2000). Architecture of RNA polymerase II and implications for the transcription mechanism. *Science* 288, 640–649. doi: 10.1126/science.288.5466.640
- Cramer, P., Bushnell, D. A., and Kornberg, R. D. (2001). Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. *Science* 292, 1863–1876. doi: 10.1126/science.1059493
- Darden, T., York, D., and Pedersen, L. (1993). Particle mesh Ewald: an N(log N) method for Ewald sums in large systems. *J. Chem. Phys.* 98, 10089–10092. doi: 10.1063/1.464397
- Darzacq, X., Shav-Tal, Y., de Turris, V., Brody, Y., Shenoy, S. M., Phair, R. D., et al. (2007). *In vivo* dynamics of RNA polymerase II transcription. *Nat. Struct. Mol. Biol.* 14, 796–806. doi: 10.1038/nsmb1280
- Day, R., and Daggett, V. (2005). Ensemble versus single-molecule protein unfolding. *Proc. Natl. Acad. Sci. U.S.A.* 102, 13445–13450. doi: 10.1073/pnas.0501773102
- Fernandez-Escamilla, A. M., Cheung, M. S., Vega, M. C., Wilmanns, M., Onuchic, J. N., and Serrano, L. (2004). Solvation in protein folding analysis: combination of theoretical and experimental approaches. *Proc. Natl. Acad. Sci. U.S.A.* 101, 2834–2839. doi: 10.1073/pnas.0304180101
- Fersht, A. R. (2000). Transition-state structure as a unifying basis in protein-folding mechanisms: contact order, chain topology, stability, and the extended nucleus mechanism. *Proc. Natl. Acad. Sci. U.S.A.* 97, 1525–1529. doi: 10.1073/pnas.97.4.1525
- Fersht, A. R., and Daggett, V. (2002). Protein folding and unfolding at atomic resolution. *Cell* 108, 573–582. doi: 10.1016/S0092-8674(02)00620-7
- Fersht, A. R., and Sato, S. (2004). Phi-value analysis and the nature of protein-folding transition states. *Proc. Natl. Acad. Sci. U.S.A.* 101, 7976–7981. doi: 10.1073/pnas.0402684101
- Fish, R. N., and Kane, C. M. (2002). Promoting elongation with transcript cleavage stimulatory factors. *Biochim. Biophys. Acta* 1577, 287–307. doi: 10.1016/S0167-4781(02)00459-1
- Fousteri, M., and Mullenders, L. H. (2008). Transcription-coupled nucleotide excision repair in mammalian cells: molecular mechanisms and biological effects. *Cell Res.* 18, 73–84. doi: 10.1038/cr.2008.6
- Gnat, A. L., Cramer, P., Fu, J., 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. doi: 10.1126/science.1059495
- Gsponer, J., and Cafilisch, A. (2002). Molecular dynamics simulations of protein folding from the transition state. *Proc. Natl. Acad. Sci. U.S.A.* 99, 6719–6724. doi: 10.1073/pnas.092686399
- Hahn, S. (2004). Structure and mechanism of the RNA polymerase II transcription machinery. *Nat. Struct. Mol. Biol.* 11, 394–403. doi: 10.1038/nsmb763
- Hanawalt, P. C. (2002). Subpathways of nucleotide excision repair and their regulation. *Oncogene* 21, 8949–8956. doi: 10.1038/sj.onc.1206096
- Hoeljmackers, J. H. (2009). DNA damage, aging, and cancer. *N. Engl. J. Med.* 361, 1475–1485. doi: 10.1056/NEJMra0804615
- Interlandi, G., Settanni, G., and Cafilisch, A. (2006). Unfolding transition state and intermediates of the tumor suppressor p16INK4a investigated by molecular dynamics simulations. *Proteins* 64, 178–192. doi: 10.1002/prot.20953
- Jackson, S. P., and Bartek, J. (2009). The DNA-damage response in human biology and disease. *Nature* 461, 1071–1078. doi: 10.1038/nature08467
- James, L. C., and Tawfik, D. S. (2003). Conformational diversity and protein evolution—a 60-year-old hypothesis revisited. *Trends Biochem. Sci.* 28, 361–368. doi: 10.1016/S0968-0004(03)00135-X
- Jeon, C., and Agarwal, K. (1996). Fidelity of RNA polymerase II transcription controlled by elongation factor TFIIS. *Proc. Natl. Acad. Sci. U.S.A.* 93, 13677–13682. doi: 10.1073/pnas.93.24.13677
- Kettenberger, H., Armache, K. J., and Cramer, P. (2004). Complete RNA polymerase II elongation complex structure and its interactions with NTP and TFIIS. *Mol. Cell* 16, 955–965. doi: 10.1016/j.molcel.2004.11.040
- Koshland, D. E. (1958). Application of a theory of enzyme specificity to protein synthesis. *Proc. Natl. Acad. Sci. U.S.A.* 44, 98–104. doi: 10.1073/pnas.44.2.98
- Kostrewa, D., Zeller, M. E., Armache, K. J., Seizl, M., Leike, K., Thomm, M., et al. (2009). RNA polymerase II-TFIIB structure and mechanism of transcription initiation. *Nature* 462, 323–330. doi: 10.1038/nature08548
- Kruse, J. P., and Gu, W. (2009). Modes of p53 regulation. *Cell* 137, 609–622. doi: 10.1016/j.cell.2009.04.050
- Kumar, S., Ma, B., Tsai, C. J., Sinha, N., and Nussinov, R. (2000). Folding and binding cascades: dynamic landscapes and population shifts. *Protein Sci.* 9, 10–19. doi: 10.1110/ps.9.1.10
- Lange, O. F., Lakomek, N. A., Farès, C., Schröder, G. F., Walter, K. F., Becker, S., et al. (2008). Recognition dynamics up to microseconds revealed from an RDC-derived ubiquitin ensemble in solution. *Science* 320, 1471–1475. doi: 10.1126/science.1157092
- Li, L., and Zou, L. (2005). Sensing, signaling, and responding to DNA damage: organization of the checkpoint pathways in mammalian cells. *J. Cell. Biochem.* 94, 298–306. doi: 10.1002/jcb.20355
- Lim, D. S., Kim, S. T., Xu, B., Maser, R. S., Lin, J., Petrini, J. H., et al. (2000). ATM phosphorylates p95/nbs1 in an S-phase checkpoint pathway. *Nature* 404, 613–617. doi: 10.1038/35007091
- Lindorff-Larsen, K., Piana, S., Palmo, K., Maragakis, P., Klepeis, J. L., Dror, R. O., et al. (2010). Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins* 78, 1950–1958. doi: 10.1002/prot.22711
- Ma, B., Kumar, S., Tsai, C. J., and Nussinov, R. (1999). Folding funnels and binding mechanisms. *Protein Eng.* 12, 713–720. doi: 10.1093/protein/12.9.713
- Ma, B., Shatsky, M., Wolfson, H. J., and Nussinov, R. (2002). Multiple diverse ligands binding at a single protein site: a matter of pre-existing populations. *Protein Sci.* 11, 184–197. doi: 10.1110/ps.21302
- Martinez-Rucobo, F. W., Sainsbury, S., Cheung, A. C., and Cramer, P. (2011). Architecture of the RNA polymerase-Spt4/5 complex and basis of universal transcription processivity. *EMBO J.* 30, 1302–1310. doi: 10.1038/emboj.2011.64
- Okazaki, K., and Takada, S. (2008). Dynamic energy landscape view of coupled binding and protein conformational change: induced-fit versus population-shift mechanisms. *Proc. Natl. Acad. Sci. U.S.A.* 105, 11182–11187. doi: 10.1073/pnas.0802524105
- Pande, V. S., and Rokhsar, D. S. (1999). Molecular dynamics simulations of unfolding and refolding of a beta-hairpin fragment of protein G. *Proc. Natl. Acad. Sci. U.S.A.* 96, 9062–9067. doi: 10.1073/pnas.96.16.9062
- Qin, F., Ye, W., Chen, Y., Chen, X., Li, Y., Zhang, J., et al. (2012). Specific recognition between intrinsically disordered LEF and DNA. *Phys. Chem. Chem. Phys.* 14, 538–545. doi: 10.1039/C1CP22610J
- Reines, D., Ghannoui, P., Gu, W., Mote, J. Jr., and Powell, W. (1993). Transcription elongation by RNA polymerase II: mechanism of SII activation. *Cell. Mol. Biol. Res.* 39, 331–338.

- Ryckaert, J. P., Ciccotti, G., and Berendsen, H. J. C. (1977). Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* 23, 327–341. doi: 10.1016/0021-9991(77)90098-5
- Shaevitz, J. W., Abbondanzieri, E. A., Landick, R., and Block, S. M. (2003). Backtracking by single RNA polymerase molecules observed at near-base-pair resolution. *Nature* 426, 684–687. doi: 10.1038/nature02191
- Shea, J. E., and Brooks, C. L. III. (2001). From folding theories to folding proteins: a review and assessment of simulation studies of protein folding and unfolding. *Annu. Rev. Phys. Chem.* 52, 499–535. doi: 10.1146/annurev.physchem.52.1.499
- Shilatifard, A., Conaway, R. C., and Conaway, J. W. (2003). The RNA polymerase II elongation complex. *Annu. Rev. Biochem.* 72, 693–715. doi: 10.1146/annurev.biochem.72.121801.161551
- Silva, D. A., Bowman, G. R., Sosa-Peinado, A., and Huang, X. (2011). A role for both conformational selection and induced fit in ligand binding by the LAO protein. *PLoS Comput. Biol.* 7:e1002054. doi: 10.1371/journal.pcbi.1002054
- Sims, R. J. III., Belotserkovskaya, R., and Reinberg, D. (2004). Elongation by RNA polymerase II: the short and long of it. *Genes Dev.* 18, 2437–2468. doi: 10.1101/gad.1235904
- Suenaga, A., Okimoto, N., Futatsugi, N., Hirano, Y., Narumi, T., Ohno, Y., et al. (2006). Structure and dynamics of RNA polymerase II elongation complex. *Biochem. Biophys. Res. Commun.* 343, 90–98. doi: 10.1016/j.bbrc.2006.02.124
- Thomas, M. J., Platas, A. A., and Hawley, D. K. (1998). Transcriptional fidelity and proofreading by RNA polymerase II. *Cell* 93, 627–637. doi: 10.1016/S0092-8674(00)81191-5
- Toulmé, F., Guérin, M., Robichon, N., Leng, M., and Rahmouni, A. R. (1999). *In vivo* evidence for back and forth oscillations of the transcription elongation complex. *EMBO J.* 18, 5052–5060. doi: 10.1093/emboj/18.18.5052
- Tsai, C. J., Ma, B., Sham, Y. Y., Kumar, S., and Nussinov, R. (2001). Structured disorder and conformational selection. *Proteins* 44, 418–427. doi: 10.1002/prot.1107
- Tsai, C. J., and Nussinov, R. (2011). Gene-specific transcription activation via long-range allosteric shape-shifting. *Biochem. J.* 439, 15–25. doi: 10.1042/BJ20110972
- Turjanski, A. G., Gutkind, J. S., Best, R. B., and Hummer, G. (2008). Binding-induced folding of a natively unstructured transcription factor. *PLoS Comput. Biol.* 4:e1000060. doi: 10.1371/journal.pcbi.1000060
- Vendruscolo, M., Paci, E., Dobson, C. M., and Karplus, M. (2001). Three key residues form a critical contact network in a protein folding transition state. *Nature* 409, 641–645. doi: 10.1038/35054591
- Wang, D., Bushnell, D. A., Huang, X., Westover, K. D., Levitt, M., and Kornberg, R. D. (2009). Structural basis of transcription: backtracked RNA polymerase II at 3.4 angstrom resolution. *Science* 324, 1203–1206. doi: 10.1126/science.1168729
- Wang, J. Y. (1998). Cellular responses to DNA damage. *Curr. Opin. Cell Biol.* 10, 240–247. doi: 10.1016/S0955-0674(98)80146-4
- Wind, M., and Reines, D. (2000). Transcription elongation factor SII. *Bioessays* 22, 327–336. doi: 10.1002/(SICI)1521-1878(200004)22:4<327::AID-BIES3>3.0.CO;2-4
- Wlodarski, T., and Zagrovic, B. (2009). Conformational selection and induced fit mechanism underlie specificity in noncovalent interactions with ubiquitin. *Proc. Natl. Acad. Sci. U.S.A.* 106, 19346–19351. doi: 10.1073/pnas.0906966106
- Ye, W., Chen, Y., Wang, W., Yu, Q., Li, Y., Zhang, J., et al. (2012a). Insight into the stability of cross-beta amyloid fibril from VEALYL short peptide with molecular dynamics simulation. *PLoS ONE* 7:e36382. doi: 10.1371/journal.pone.0036382
- Ye, W., Qin, F., Zhang, J., Luo, R., and Chen, H. F. (2012b). Atomistic mechanism of microRNA translation upregulation via molecular dynamics simulations. *PLoS ONE* 7:e43788. doi: 10.1371/journal.pone.0043788

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Wu, Ye, Yang and Chen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.