

Nano positioning system reveals the course of upstream and nontemplate DNA within the RNA polymerase II elongation complex

Joanna Andrecka¹, Barbara Treutlein¹, Maria Angeles Izquierdo Arcusa², Adam Muschielok¹, Robert Lewis¹, Alan C. M. Cheung^{1,3}, Patrick Cramer^{1,3} and Jens Michaelis^{1,*}

¹Department of Chemistry and Biochemistry and Center for Integrated Protein Science München, Ludwig-Maximilians-Universität München, Butenandtstr.11, 81377 München, Germany, ²Departamento de Química Inorgánica y Orgánica, Universidad Jaume I de Castellón, 12071 Castellón, Spain and ³Gene Center Munich, Ludwig-Maximilians-Universität München, Feodor Lynen Str. 25, 81377 München, Germany

Received May 15, 2009; Revised June 30, 2009; Accepted July 01, 2009

ABSTRACT

Crystallographic studies of the RNA polymerase II (Pol II) elongation complex (EC) revealed the locations of downstream DNA and the DNA-RNA hybrid, but not the course of the nontemplate DNA strand in the transcription bubble and the upstream DNA duplex. Here we used single-molecule Fluorescence Resonance Energy Transfer (smFRET) experiments to locate nontemplate and upstream DNA with our recently developed Nano Positioning System (NPS). In the resulting complete model of the Pol II EC, separation of the nontemplate from the template strand at position +2 involves interaction with fork loop 2. The nontemplate strand passes loop β 10- β 11 on the Pol II lobe, and then turns to the other side of the cleft above the rudder. The upstream DNA duplex exits at an approximately right angle from the incoming downstream DNA, and emanates from the cleft between the protrusion and clamp. Comparison with published data suggests that the architecture of the complete EC is conserved from bacteria to eukaryotes and that upstream DNA is relocated during the initiation-elongation transition.

INTRODUCTION

RNA polymerase II (Pol II) is the key enzyme that produces all mRNA in eukaryotic cells. Over the last years, the Pol II structure and functional mechanism have been extensively studied. Major breakthroughs were the solution of the crystal structures of Pol II in

free form (1–4) and in form of active elongation complexes (ECs) with bound DNA and RNA (5–9). These studies revealed the course of the downstream DNA duplex (register +3 to +15), the DNA template single strand within the transcription bubble (positions +2 to –10) and the nascent RNA (positions +1 to –10) within the Pol II EC (Figure 1, register +1 corresponds to the nucleotide addition site; positive and negative numbers refer to downstream and upstream positions, respectively). However, the location of the upstream DNA duplex, the nontemplate DNA strand within the transcription bubble, and the exiting nascent RNA beyond register –10 were not resolved in Pol II EC structures, despite their presence in at least one crystal (6), indicating their mobility.

The structure and dynamics of such mobile regions in macromolecules can be analyzed by single molecule Fluorescence Resonance Energy Transfer (smFRET) measurements (10,11). This method requires that dye molecules are attached to the mobile region and to other, static regions within the macromolecule or the macromolecular complex. If at least three different distances are probed by a smFRET experiment, the data can be used for determining the desired relative position of a mobile region with respect to known positions within a structure by triangulation (12–16). Previously, we used smFRET-based triangulation to determine the course of the RNA exiting from Pol II (13,17). These results were recently confirmed by an independent group that used a similar approach (18). There are, however, uncertainties associated with a smFRET measurement that must be taken into account by computing not only the most likely position but the three dimensional probability density function (PDF) of the position. To determine such PDFs, we developed a Nano Positioning System (NPS) that uses X-ray crystallographic information, smFRET

*To whom correspondence should be addressed. Tel: +49 89 2180 77561; Fax: +49 89 2180 77560; Email: michaelis@lmu.de

data, and Bayesian parameter estimation (17). We used NPS to determine the position of the 5'-end of a 29-nt-long RNA and showed how transcription factor IIB (TFIIB) influences this position (17).

Despite this progress, current structural studies of the Pol II EC did not arrive at a complete picture, since the course of the upstream DNA and the nontemplate strand in the transcription bubble remain unknown. The upstream DNA duplex contributes to EC stability, as shown for a bacterial EC (19). The nontemplate DNA strand is also required for EC stability, and for maintaining the upstream edge of the bubble (20,21). Nontemplate DNA also maintains the lateral stability of the polymerase by reducing the probability of backtracking (22) and influences polymerase pausing and arrest (23). Furthermore, several transcription factors apparently interact with the nontemplate strand in the bubble, including the elongation factor RfaH (24), and the eukaryotic initiation factor TFIIE (25).

Here, we used NPS to determine the position of upstream DNA and the nontemplate strand in the Pol II EC. We measured smFRET efficiencies between 'antenna dye molecules' (ADMs) attached to the nontemplate DNA strand and several 'satellite dye molecules' (SDMs) attached to positions within the EC that are known from crystallographic studies. The obtained data allowed us to build a model for the complete Pol II EC containing all nucleic acid regions.

MATERIALS AND METHODS

Nucleic acids scaffolds

Nucleic acid scaffolds which were used to build artificial ECs were constructed from RNA and DNA oligomers using partially mismatched template and nontemplate DNA strands (Figure 1). The DNA and RNA strands were purchased from IBA (Göttingen, Germany). This construct has been used previously for solving the Pol II EC structure (6). DNA and RNA molecules were annealed as described (13).

Preparation of Pol II EC

Rpb4/7 wild-type, as well as Rpb7-C150 and Rpb4-S73C mutants were expressed and purified as described (17). Dye labeling of the single-cysteine mutants was conducted using 8–10-fold molar excess of Alexa647-C2-Maleimide (Molecular Probes) in assembly buffer (50 mM HEPES, 40 mM (NH₄)₂SO₄, 5 μM ZnCl₂, 5% glycerol) at 37°C for 1 h. Free dye was removed using G-50 spin-columns (Amersham Biosciences, Freiburg, Germany) with assembly buffer containing 10 mM DTT. The Pol II–DNA–RNA–Rpb4/7 complexes was assembled as described (4).

Determination of the Förster distances

For each donor-acceptor pair the isotropic Förster distance R_0^{iso} was determined using standard procedures (26). First, the donor quantum yield was determined by comparing its fluorescence to that of Rhodamine 101 in ethanol, a well established fluorescence standard.

Second, overlap integrals were calculated using donor emission spectra from 530–700 nm (excitation wavelength 528 nm) and acceptor excitation spectra from 400–700 nm (detection wavelength 705 nm) recorded in solution using a steady state fluorescence spectrometer (Edinburgh Instruments F900). Moreover, $n = 1.35$ and $\kappa^2 = 2/3$ were used to calculate the isotropic Förster distance R_0^{iso} . In order to account for uncertainties in the Förster distance due to orientation effects we first measured the anisotropies of the donor and acceptor dyes for all attachment sites (Supplementary Table 1). Assuming no additional rotational movement beyond the timescale of the fluorescence lifetime, we then performed Monte Carlo Simulation to calculate the Förster distance PDF assuming an isotropic distribution of the average dye molecule orientation (17). For all donor-acceptor pairs the calculated PDF was approximated by 10 Gaussians for use in the NPS analysis as described recently (17).

Experimental setup for sp-FRET, data collection and analysis

Single-pair FRET experiments were performed on an upgraded version of the custom-built prism-based total internal reflection fluorescence microscope (TIRFM) described previously (13). For the excitation of donor and FRET-pairs a frequency doubled Nd-YAG laser at 532 nm was used (Spectra Physics), while for direct excitation of acceptor molecules a diode laser at 637 nm (Coherent) was employed. The fluorescence signal of donor and acceptor were isolated in the detection path by the use of a dichroic mirror (Chroma 645DCXR) an emission band-pass filters centered at 580 nm and 610 nm (3RD Millennium, Omega Optical), and collected on an EM-CCD camera (DU897BV, Andor). Pol II ECs were attached to the glass surface via a PEG-Biotin/Neutravidin/Biotin layer system. The acquired data were analyzed using custom software written in MATLAB. We used a fully automated routine to find FRET pairs, to calculate and subtract the local background and to compute fluorescence trajectories. For the calculation of FRET efficiency of the individual FRET pairs, we used the following formula:

$$E = \frac{I_A - \beta I_D}{I_A + \gamma I_D}, \quad 1$$

where $\gamma = \frac{I_A - I'_A}{I_D - I'_D}$, and $\beta = \frac{I'_A}{I'_D}$. I_A and I_D are the background corrected intensities from the acceptor and donor channels and I and I' are the intensities before and after acceptor photobleaching, respectively. β and γ are correction factors; β accounts for the leakage of the donor emission into the acceptor channel, while γ is a factor that includes the quantum yields of the fluorophores and the detection efficiencies of the two channels. We determined the correction factors for all FRET pairs individually by time averaging the intensities I and I' . FRET pairs where no acceptor bleaching was observed were discarded from the analysis. Examples of typical single-molecule fluorescence signals and calculated FRET values are shown in Supplementary Figure 1. For all measured data direct

excitation of the acceptor was negligible and was therefore omitted from the analysis for practical reasons.

The histograms of single molecule events were fitted using single Gaussians extracting the center FRET efficiency and its standard error. The standard error was always extremely low in comparison to the estimated systematic errors of ~2% for $E < 0.9$ and ~4% for $E \geq 0.9$ and therefore the systematic errors were used as an experimental uncertainty. The FRET efficiencies and errors were then used for further analysis with NPS (17). Only for the position -18 we observed a subpopulation of ~30% of the complexes with a different FRET value leading to a second peak in the histogram. The obtained smFRET histograms were therefore fitted using two Gaussians and the center of the two Gaussians were used in the NPS analysis. Thus the ADM attached to the -18 position can be located at two distinct positions and dynamic switching between the two positions was observed at the single-molecule level (Supplementary Figure 2).

Uncertainty in the position of dye molecules attached to known positions

Dye molecules were attached to known positions within the Pol II EC using flexible six or 12 carbon-atom linkers (SDMs). Thus, while the attachment point is known from the structural model, the precise location of the dye molecule is not known. We therefore calculated the volume that is sterically accessible to the dye molecules, given the point of attachment, size of the molecule and the linker length (17). We assume each SDM position within this accessible volume equally probable and for calculation purposes approximate the resulting PDF by 15 Gaussians (Supplementary Figure 3). These Gaussians are used in the NPS analysis to describe the uncertainty of the SDM position (17).

Determination of the nontemplate DNA position

The X-ray structure of the EC (6) was used as a reference frame for the position calculation. Moreover, the volume occupied in the crystal structure was used as a restriction for the possible positions of the dye molecules. We assumed zero probability density within an already occupied volume and equal probability density elsewhere in order to calculate the ADM prior. Moreover, for the positions -15 and -18 the volume accessible to the ADM can be restricted further by using the length of the DNA molecule. The accessible volume was calculated using a flexible chain (with a radius of 10 Å), which is allowed to sample all possible pathways, with a length of 35 Å for the -15 position and 45 Å for the -18 position starting at the last known position (template DNA -10; (17)). In order to compute the desired positions of the ADMs along the nontemplate DNA strand we applied the recently developed NPS method separately for each ADM using its corresponding FRET efficiencies and the information about the possible SDM positions and Förster distances as described above [for details about the method see ref. (17)]. As a result we obtain the PDF for the ADM position, $p(x|\{E_i\}, I)$. Here, $\{E_i\}$ denotes the mean FRET

efficiencies of the measurements used for a particular calculation and I the background information such as attachment points and isotropic Förster distances. The ADM position PDF evaluated on a cubic lattice (spacing 1 Å) was saved as XPLOR file (Supplementary Data) in order to display the credibility volume of the ADM positions as iso-surfaces in Pymol (DeLano Scientific). The complete NPS analyses was performed using a custom written software which is available free of charge at www.cup.uni-muenchen.de/pc/michaelis/software (17).

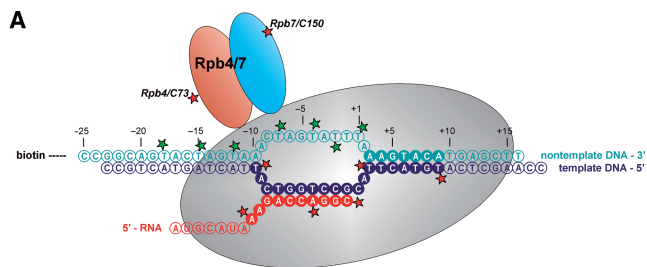
Modeling

The calculated ADM PDFs were used to construct a model of the nontemplate strand from register +2 to -23 and of the template strand from -11 to -23. This represents the single-stranded region of the nontemplate strand and the upstream duplex DNA, neither of which are visible in the crystal structure of the Pol II EC 1Y1W (6). The modeled nontemplate DNA was positioned such that an ADM labeled base passed adjacent to the corresponding ADM position (but not entirely within it), since the ADM position describes the position of the dye molecule and not that of the actual base. Base stacking was maintained throughout the single-stranded part of the model. It was assumed that the point of rewinding occurs at register -11 and a straight B-form DNA helix was inserted at this point to represent the upstream duplex. The model was built using coot (27) and energy minimised with phenix.refine (28).

RESULTS

To determine the positions of the nontemplate and upstream DNA, we attached a fluorescence donor that acted as the ADM to the nontemplate DNA strand at positions +1, -2, -4, -7, -12, -15 or -18 (Figure 1). Fluorescence acceptors that acted as the SDMs were attached to known positions within the EC, including positions in the template DNA strand, the RNA strand, and the Pol II subcomplex Rpb4/7 ('Materials and Methods' section and Figure 1). ECs were assembled, purified, immobilized, and analyzed in a custom-built total internal reflection fluorescence (TIRF) microscope as described (13). For each SDM-ADM pair, hundreds of smFRET time-traces were recorded. The smFRET data were then entered into histograms which were fitted using Gaussians to extract the mean FRET efficiency between ADM and SDM (Supplementary Figures 4-10). For each ADM we used 6-8 different SDMs yielding in total 50 ADM-SDM pairs (Figure 1B and Supplementary Table 1). For each pair we experimentally determined the isotropic Förster distance R_0^{iso} ('Materials and Methods' section). We then computed the three-dimensional probability density for the position of the ADM using NPS (17). The densities were displayed as credibility volumes enclosing a certain probability of finding the dye molecule which reflect the localization accuracy of the experiment (Figure 2).

NPS revealed the positions of all ADMs in the DNA nontemplate strand (Figure 2A). The close spacing of the



| Antenna (DONOR) | Satellites (ACCEPTOR = alexa 647) |
|--------------------|---|
| <i>NT-DNA(+1)</i> | <i>T-DNA(-10)*</i> , <i>T-DNA(+9)*</i> , <i>RNA1*</i> , <i>RNA4*</i> , <i>RNA10*</i> , <i>Rpb7/C150**</i> |
| <i>NT-DNA(-2)</i> | <i>T-DNA(-10)*</i> , <i>T-DNA(+3)*</i> , <i>T-DNA(+9)*</i> , <i>RNA1*</i> , <i>RNA4*</i> , <i>RNA10*</i> , <i>Rpb7/C150**</i> |
| <i>NT-DNA(-4)</i> | <i>T-DNA(-10)*</i> , <i>T-DNA(+3)*</i> , <i>T-DNA(+9)*</i> , <i>RNA1*</i> , <i>RNA4*</i> , <i>RNA10*</i> , <i>Rpb7/C150**</i> |
| <i>NT-DNA(-7)</i> | <i>T-DNA(-10)*</i> , <i>T-DNA(+3)*</i> , <i>T-DNA(+9)*</i> , <i>RNA1*</i> , <i>RNA4*</i> , <i>RNA10*</i> , <i>Rpb7/C150**</i> |
| <i>NT-DNA(-12)</i> | <i>T-DNA(+3)*</i> , <i>T-DNA(+9)*</i> , <i>RNA4*</i> , <i>RNA10*</i> , <i>Rpb7/C150**</i> , <i>Rpb4/C73**</i> |
| <i>NT-DNA(-15)</i> | <i>T-DNA(-10)*</i> , <i>T-DNA(+3)**</i> , <i>T-DNA(+9)**</i> , <i>RNA1*</i> , <i>RNA4*</i> , <i>RNA10*</i> , <i>Rpb7/C150**</i> , <i>Rpb4/C73**</i> |
| <i>NT-DNA(-18)</i> | <i>T-DNA(-10)*</i> , <i>T-DNA(+3)**</i> , <i>T-DNA(+9)**</i> , <i>RNA1*</i> , <i>RNA4*</i> , <i>RNA10*</i> , <i>Rpb7/C150**</i> , <i>Rpb4/C73**</i> |

* alexa 555 used as donor

** TMR used as donor

Figure 1. Experimental design. (A) Labeling positions. Elongation complexes were formed using mismatched nucleic acid scaffolds. The template DNA strand, nontemplate DNA strand, and product RNA are colored blue, cyan, and red, respectively. Positions of attached dye molecules are indicated by green stars for donor and red stars for acceptor dye molecules. Bases whose positions were determined by crystallographic studies (6) are shown as solid circles. Those positions that could not be determined previously are shown as open circles. Pol II core (gray) and Rpb4/7 (red/blue) are shown schematically. A fluorescent dye molecule (Alexa 555 or TMR) that acted as the fluorescence donor in the smFRET experiments was attached to position +1, -2, -4, -7 -12, -15 or -18 of the nontemplate DNA. Since these are the positions that are to be determined the molecule is called the 'antenna'. The acceptor dye molecule (Alexa 647) was used like a 'satellite' at a known position and was either attached to the template DNA (at positions -10, +3 or +9) or RNA (at -1, -4 or -10) or at one of two positions on the heterodimer Rpb4/7 (Rpb7-C150 and Rpb4-C73). (B) Overview of satellite positions. For each antenna position only those satellites whose distance to the antenna was expected to fall in the sensitive range for FRET measurements were used. Moreover, depending on the expected distance either Alexa 555 or TMR were used as fluorescence donor.

ADM positions defined the overall course of the DNA. The ADM PDFs were used to build a model for the nontemplate DNA and the upstream DNA duplex, taking into account that the volumes define the locations of the dye molecules attached to the DNA nucleotides via a linker, not the nucleotides *per se* (Figure 2B). The model was built manually and its stereochemistry was optimized ('Materials and Methods' section). The accessible volume for the ADMs calculated from the modeled nontemplate

DNA overlapped well with the experimentally determined credibility volumes for all ADM positions (Figure 3).

In our model, the strands of the downstream DNA duplex are separated at register +2 at fork loop 2 (Figure 4), passing near the conserved residue R504, which is important for transcription *in vitro* (29). From there, the nontemplate strand is directed towards the lobe of Rpb2 (Figure 4). At register -4, the nontemplate strand approaches a loop on top of the lobe domain (loop β 10- β 11, Rpb2 residues 272-278) on the Rpb2 side of the cleft (Figure 4). This loop is invariant between yeast and human Pol II and includes an exposed lysine residue (K277) that likely interacts with the nontemplate strand. The nontemplate strand then turns towards the Rpb1 side of the cleft and continues to the rudder, an extension from the clamp domain (Figure 4). The backbone of the nontemplate DNA at register -7 to -9 is close to the Rpb1 residues 309-315 in the rudder as well as to conserved Lys 471 of fork loop 1 at position -10 (Figure 4). Nontemplate and template strands re-anneal at register -11 to form the upstream DNA duplex. Upstream DNA exits from the Pol II cleft between the Rpb1 clamp and the Rpb2 protrusion domains. The axis of the upstream DNA duplex is parallel to the axes of helices α 8 in the clamp and α 11 in the protrusion (Figure 4).

As the upstream DNA exits from polymerase, its course becomes less well defined, indicating increasing mobility. At the most upstream ADM (register -18), smFRET time traces were not constant but rather showed dynamic switching between two values (Supplementary Figure 2). The resulting histograms revealed two peaks, which were fitted by two different Gaussians (Supplementary Table 1). Using NPS we determined two different positions of the ADM. The dominant position, which accounts for ~70% of the data, is shown in Figure 2A and was used for building the DNA model. The alternative position is shifted towards the protrusion by about 10 Å (Supplementary Figure 11). Apparently, the upstream DNA assumes another conformation and leaves the polymerase at a slightly different angle.

DISCUSSION

Here, we used smFRET measurements and NPS analysis to reveal the course of the nontemplate and upstream DNA within the complete Pol II EC. We arrived at a complete model for the Pol II EC that is consistent with published biochemical, biophysical and genetic data. First, the angle between the downstream and upstream DNA duplexes in our model is $\sim 80^\circ$, in agreement with previous studies using atomic force microscopy (30). Second, the bases in the upstream part of the nontemplate DNA within the bubble (positions -5 to -10) are exposed, explaining why limited micrococcal nuclease (MN) cleavage degrades DNA from the upstream end up to position -5 in the bacterial EC (22). Downstream of position -5 the bases of the nontemplate strand point inside the cleft, preventing MN cleavage. Third, our model is consistent with site-specific DNA-protein cross-linking within the bacterial EC (31,32) (Supplementary Figure 12).

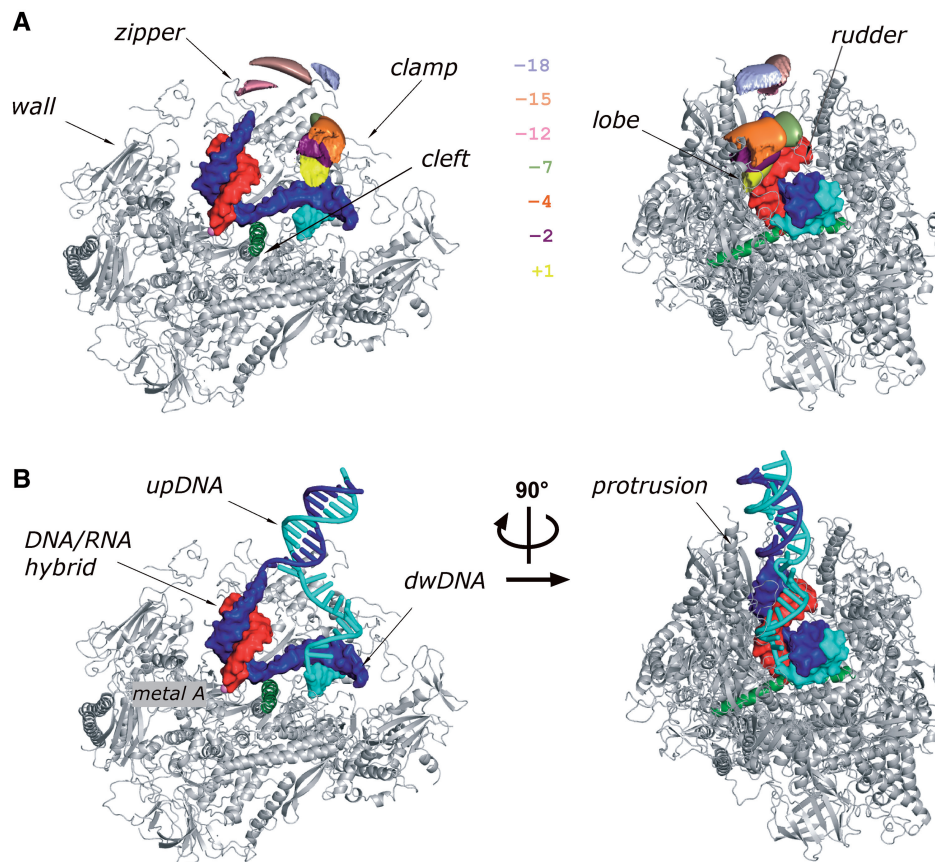


Figure 2. Position of nontemplate and upstream DNA in Pol II elongation complexes. **(A)** NPS results calculated from measured FRET values. Thirty-eight percent credibility volumes obtained for nontemplate DNA at register +1 (yellow), -2 (purple), -4 (orange), -7 (green), -12 (pink), -15 (dark salmon) and -18 (light blue) are displayed using surface representation. (Left) Side view of Pol II core enzyme shown in cartoon representation (gray). Rpb2 was omitted for clarity to reveal the nucleic acids. (Right) Alternate view of the polymerase (rotation by 90° as indicated). Here, all 10 polypeptides of the core enzyme are shown. The previously determined parts of the template DNA (dark blue), nontemplate DNA (light blue) and RNA (red) (6) are shown using surface representation. **(B)** Complete picture of the Pol II elongation complex. Obtained probability densities were used to build a model of nontemplate and upstream DNA. Modeled DNA is shown in cartoon representation. Both panels show the same orientations as in (A).

These previous data and our results suggest that the architecture of the EC is conserved from bacteria to eukaryotes for all cellular RNA polymerases.

However, the locations inferred from chemical cross-linking deviate in some respects from our experimentally determined locations of the nontemplate and upstream DNA, likely due to a lack of cross-links to the nontemplate strand between positions -4 and -14 in the previous study (31). Whereas the nontemplate strand in the bubble was previously modeled to penetrate the groove between the lobe and protrusion (31), our results show that this region passes loop $\beta 10$ - $\beta 11$ close to the four-stranded β -sheet of the lobe. Deletions in this domain cause defects in the formation of the bacterial open promoter complex (33,34) indicating that interactions between loop $\beta 10$ - $\beta 11$ and the nontemplate strand may contribute to DNA melting or stabilization of the bubble. The upstream DNA duplex in our model also exits at a different angle from Pol II when compared to that previously suggested (31).

Our complete EC model suggests that the rudder, a loop protruding from the clamp into the cleft, plays crucial

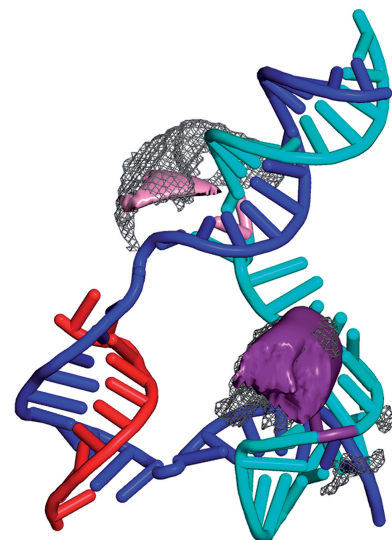


Figure 3. Consistency check of built model. A 38% credibility volume for dye molecule attached to nontemplate DNA at register -2 (purple) and -12 (pink) determined using NPS (surface representation) and computed accessible volume for a dye molecule attached to the respective position of the modeled nontemplate DNA (gray meshes).

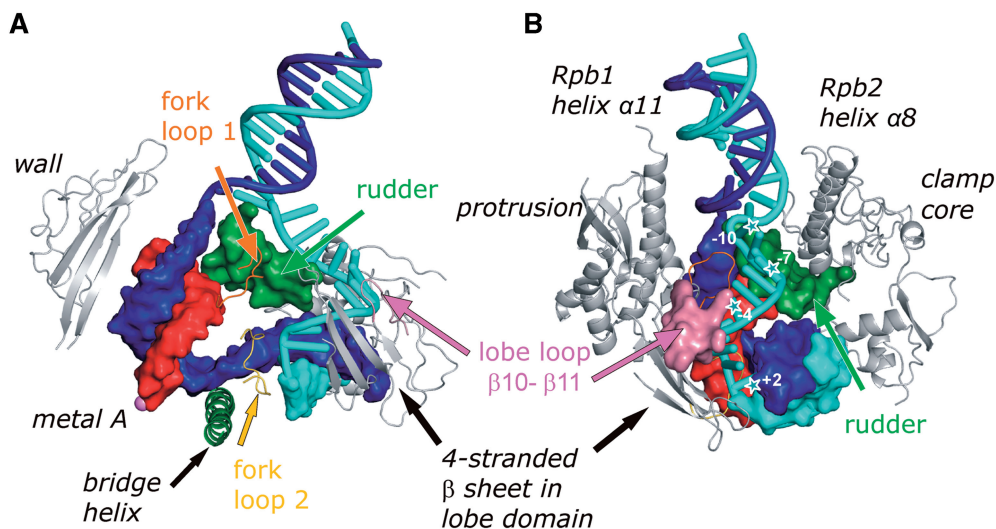


Figure 4. Interaction network. Zoom into the elongation complex showing important structural features that define the path of the nontemplate DNA. Previously known DNA and RNA positions are shown in surface representation and modeled position of nontemplate and upstream DNA using cartoons. (A) Side view, polymerase elements important for the pathway of the nontemplate DNA are highlighted: rudder (green), fork loop 1 (orange), fork loop 2 (yellow) and lobe loop (residues 272–278, pink). (B) A 90° rotated view (same orientation as Figure 2, right).

roles during transcription. The DNA nontemplate strand passes above Rpb1 residues 309–315 in the rudder, and re-anneals with the template strand to form the exiting upstream DNA duplex, suggesting that the rudder is important for proper DNA rewinding and maintenance of the upstream end of the bubble. The model also explains why the rudder is critical for promoter opening and EC stability *in vitro* (29,35), and for transcription elongation *in vivo* (36). Our results argue against the original proposal that two other loops extending from the clamp, the zipper and lid, separate DNA strands at the upstream bubble (1,5). The lid instead is involved in RNA separation from DNA at the end of the hybrid (6,7). However, a bacterial polymerase mutant that lacks the lid maintains the correct hybrid length, but only when nontemplate DNA is present (37,38), suggesting that maintenance of the upstream end of the bubble and maintenance of the hybrid length are coupled. In addition, other transcription factors which bind to the nontemplate strand during initiation (25) and elongation (24) will contribute to the interaction network and further stabilize the upstream end of the bubble.

Our results further suggest that the location of upstream DNA changes dramatically during the transition from initiation to elongation. In the closed complex, upstream DNA is located to the polymerase upstream face outside the cleft (39), whereas in the EC it is located above the cleft between the polymerase protrusion and clamp. This relocation of upstream DNA may occur during open complex formation, when the DNA is melted and the template strand slips into the active center, or during promoter escape, when the initial transcription bubble collapses and initiation factors are released (40). These and other dynamic structural transitions during the transcription cycle can be addressed using NPS in the future.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We would like to thank Julia Nagy for help with initial measurements.

FUNDING

The Deutsche Forschungsgemeinschaft, SFB 646; the Center for Nanoscience (CeNS); the Nanosystems Initiative Munich (NIM); and the Fundacio Caixa Castello-Bancaixa and Universidad Jaume I (Project 07I003.15/1 to M.A.I.). Funding for open access charge: Deutsche Forschungsgemeinschaft, SFB 646.

Conflict of interest statement. None declared.

REFERENCES

1. 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.
2. Armache,K.-J., Mitterweiger,S., Meinhart,A. and Cramer,P. (2005) Structures of complete RNA polymerase II and its subcomplex Rpb4/7. *J. Biol. Chem.*, **280**, 7131–7134.
3. Bushnell,D.A. and Kornberg,R.D. (2003) Complete RNA polymerase II at 4.1 Å resolution: implications for the initiation of transcription. *Proc. Natl Acad. Sci. USA*, **100**, 6969–6972.
4. Armache,K.-J., Kettenberger,H. and Cramer,P. (2003) Architecture of the initiation-competent 12-subunit RNA polymerase II. *Proc. Natl Acad. Sci. USA*, **100**, 6964–6968.
5. Gnatt,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.
6. 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.

7. Westover, K.D., Bushnell, D.A. and Kornberg, R.D. (2004) Structural basis of transcription: separation of RNA from DNA by RNA polymerase II. *Science*, **303**, 1014–1016.
8. Westover, K.D., Bushnell, D.A. and Kornberg, R.D. (2004) Structural basis of transcription: nucleotide selection by rotation in the RNA polymerase II active center. *Cell*, **119**, 481–489.
9. Brueckner, F. and Cramer, P. (2008) Structural basis of transcription inhibition by [alpha]-amanitin and implications for RNA polymerase II translocation. *Nat. Struct. Mol. Biol.*, **15**, 811–818.
10. Ha, T., Enderle, T., Ogle, D.F., Chemla, D.S., Selvin, P.R. and Weiss, S. (1996) Probing the interaction between two single molecules: fluorescence resonance energy transfer between a single donor and a single acceptor. *Proc. Natl Acad. Sci. USA*, **93**, 6264–6268.
11. Joo, C., Balci, H., Ishitsuka, Y., Buranachai, C. and Ha, T. (2008) Advances in single-molecule fluorescence methods for molecular biology. *Annu. Rev. Biochem.*, **77**, 51–76.
12. Mekler, V., Kortkhonjia, E., Mukhopadhyay, J., Knight, J., Revyakin, A., Kapanidis, A.N., Niu, W., Ebright, Y.W., Levy, R. and Ebright, R.H. (2002) Structural organization of bacterial RNA polymerase holoenzyme and the RNA polymerase-promoter open complex. *Cell*, **108**, 599–614.
13. Andrecka, J., Lewis, R., Bruckner, F., Lehmann, E., Cramer, P. and Michaelis, J. (2008) Single-molecule tracking of mRNA exiting from RNA polymerase II. *Proc. Natl Acad. Sci. USA*, **105**, 135–140.
14. Knight, J.L., Mekler, V., Mukhopadhyay, J., Ebright, R.H. and Levy, R.M. (2005) Distance-restrained docking of rifampicin and rifamycin SV to RNA polymerase using systematic FRET measurements: developing benchmarks of model quality and reliability. *Biophys. J.*, **88**, 925–938.
15. Rasnik, I., Myong, S., Cheng, W., Lohman, T.M. and Ha, T. (2004) DNA-binding orientation and domain conformation of the E. coli rep helicase monomer bound to a partial duplex junction: single-molecule studies of fluorescently labeled enzymes. *J. Mol. Biol.*, **336**, 395–408.
16. Margittai, M., Widengren, J., Schweinberger, E., Schroder, G.F., Felekyan, S., Haustein, E., Konig, M., Fasshauer, D., Grubmuller, H., Jahn, R. et al. (2003) Single-molecule fluorescence resonance energy transfer reveals a dynamic equilibrium between closed and open conformations of syntaxin I. *Proc. Natl Acad. Sci. USA*, **100**, 15516–15521.
17. Muschiok, A., Andrecka, J., Jawhari, A., Bruckner, F., Cramer, P. and Michaelis, J. (2008) A nano-positioning system for macromolecular structural analysis. *Nat. Methods*, **5**, 965–971.
18. Chen, C.Y., Chang, C.C., Yen, C.F., Chiu, M.T. and Chang, W.H. (2009) Mapping RNA exit channel on transcribing RNA polymerase II by FRET analysis. *Proc. Natl Acad. Sci. USA*, **106**, 127–132.
19. Sidorenkov, I., Komissarova, N. and Kashlev, M. (1998) Crucial role of the RNA:DNA hybrid in the processivity of transcription. *Mol. Cell*, **2**, 55–64.
20. Kireeva, M.L., Komissarova, N. and Kashlev, M. (2000) Overextended RNA:DNA hybrid as a negative regulator of RNA polymerase II processivity. *J. Mol. Biol.*, **299**, 325–335.
21. Kireeva, M.L., Komissarova, N., Waugh, D.S. and Kashlev, M. (2000) The 8-nucleotide-long RNA:DNA hybrid is a primary stability determinant of the RNA polymerase II elongation complex. *J. Biol. Chem.*, **275**, 6530–6536.
22. Wang, D. and Landick, R. (1997) Nuclease cleavage of the upstream half of the nontemplate strand DNA in an Escherichia coli transcription elongation complex causes upstream translocation and transcriptional arrest. *J. Biol. Chem.*, **272**, 5989–5994.
23. Nudler, E., Goldfarb, A. and Kashlev, M. (1994) Discontinuous mechanism of transcription elongation. *Science*, **265**, 793–796.
24. Sevostyanova, A., Svetlov, V., Vassilyev, D.G. and Artsimovitch, I. (2008) The elongation factor RfaH and the initiation factor σ^{70} bind to the same site on the transcription elongation complex. *Proc. Natl Acad. Sci. USA*, **105**, 865–870.
25. Chen, H.-T., Warfield, L. and Hahn, S. (2007) The positions of TFIIF and TFIIE in the RNA polymerase II transcription initiation complex. *Nat. Struct. Mol. Biol.*, **8**, 696–703.
26. Vamosi, G., Gohlke, C. and Clegg, R.M. (1996) Fluorescence characteristics of 5-carboxytetramethylrhodamine linked covalently to the 5' end of oligonucleotides: multiple conformers of single-stranded and double-stranded dye-DNA complexes. *Biophys. J.*, **71**, 972–994.
27. Emsley, P. and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.*, **60**, 2126–2132.
28. Afonine, P.V., Grosse-Kunstleve, R.W. and Adams, P.D. (2005) A robust bulk-solvent correction and anisotropic scaling procedure. *Acta Crystallogr. D Biol. Crystallogr.*, **61**, 850–855.
29. Naji, S., Bertero, M.G., Spitalny, P., Cramer, P. and Thomm, M. (2008) Structure-function analysis of the RNA polymerase cleft loops elucidates initial transcription, DNA unwinding and RNA displacement. *Nucleic Acids Res.*, **36**, 676–687.
30. Rees, W., Keller, R.W., Vesenka, J.P., Yang, G. and Bustamante, C. (1993) Evidence of DNA bending in transcription complexes imaged by scanning force microscopy. *Science*, **260**, 1646–1649.
31. Korzheva, N., Mustaev, A., Kozlov, M., Malhotra, A., Nikiforov, V., Goldfarb, A. and Darst, S.A. (2000) A structural model of transcription elongation. *Science*, **289**, 619–625.
32. Naryshkin, N., Revyakin, A., Kim, Y., Mekler, V. and Ebright, R.H. (2000) Structural organization of the RNA polymerase-promoter open complex. *Cell*, **101**, 601–611.
33. Nechaev, S., Chlenov, M. and Severinov, K. (2000) Dissection of two hallmarks of the open promoter complex by mutation in an RNA polymerase core subunit. *J. Biol. Chem.*, **275**, 25516–25522.
34. Severinov, K. and Darst, S.A. (1997) A mutant RNA polymerase that forms unusual open promoter complexes. *Proc. Natl Acad. Sci. USA*, **94**, 13481–13486.
35. Kuznedelov, K., Korzheva, N., Mustaev, A. and Severinov, K. (2002) 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**, 1369–1378.
36. Malagon, F., Kireeva, M.L., Shafer, B.K., Lubkowska, L., Kashlev, M. and Strathern, J.N. (2006) Mutations in the Saccharomyces cerevisiae RPB1 gene conferring hypersensitivity to 6-azauracil. *Genetics*, **172**, 2201–2209.
37. Touloukhonov, I. and Landick, R. (2006) The role of the lid element in transcription by E. coli RNA polymerase. *J. Mol. Biol.*, **361**, 644–658.
38. Naryshkina, T., Kuznedelov, K. and Severinov, K. (2006) The role of the largest RNA polymerase subunit lid element in preventing the formation of extended RNA-DNA hybrid. *J. Mol. Biol.*, **361**, 634–643.
39. Chen, H.T. and Hahn, S. (2004) Mapping the location of TFIIF within the RNA polymerase II transcription preinitiation complex: a model for the structure of the PIC. *Cell*, **119**, 169–180.
40. Ujvari, A. and Luse, D.S. (2006) RNA emerging from the active site of RNA polymerase II interacts with the Rpb7 subunit. *Nat. Struct. Mol. Biol.*, **13**, 49–54.