Molecular Cell, *Volume 46*

Supplemental Information

Dynamic Architecture of a Minimal RNA Polymerase II Open Promoter Complex

Barbara Treutlein, Adam Muschielok, Joanna Andrecka, Anass Jawhari, Claudia Buchen, Dirk Kostrewa, Friederike Hög, Patrick Cramer, and Jens Michaelis

Supplemental Experimental Procedures

Protein Expression, Purification, and Labeling

TBP

Saccharomyces cerevisiae TBP core with an N-terminal His-tag (amino acids 61-240, gift from Z. S. Juo) and mutant, full-length TBP lacking all native cysteines (C78A, C164A) and holding a single cysteine mutation at residue 128 (S128C) were expressed for 15 h at 20 °C in *E. coli* BL21-Codon plus (DE3) RIL (Stratagene). For mutant TBP-S128C, a plasmid for TBP-C61 (gift from M. Meisterernst) lacking all native cysteines (C78A, C164A) and holding a single cysteine mutation at residue 61 (S61C) was subjected to site-directed mutagenesis to change C61 back to a serine and to introduce a new cysteine at residue 128 (S128C). Cells were harvested by centrifugation, resuspended in buffer A (25 mM Tris-HCl, pH 8.0 (25 °C), 500 mM NaCl, 50 mM ammonium acetate, 10% glycerol) and lyzed by sonication. After centrifugation, the supernatant was loaded onto a 2 ml Ni-NTA column (Qiagen), equilibrated with buffer A. After washing, bound protein was eluted with 250 mM imidazole in buffer A. The protein was then purified by a heparin sepharose column (HiTrap Heparin HP, 5ml, GE Healthcare) using buffer A and a linear gradient from 250 mM to 1 M NaCl. After concentration, TBP core was further purified by application to a Superose 6 HR gel filtration column (GE Healthcare) equilibrated with Pol II buffer (50 mM HEPES pH 7.3, 40 mM ammonium sulphate, 10 µM $ZnCl_2$, 20% glycerol, 2 mM dithiothreitol (DTT)). Mutant TBP-S128C was directly subjected to dye labeling using a 10-15 fold molar excess of Alexa 647-C2-Maleimide (Molecular Probes) in buffer A containing 600mM NaCl at room temperature for 1h. Free dye was removed using G-50 spin-columns (GE Healthcare) with assembly buffer (50 mM HEPES, 40 mM ammonium sulfate, 10 µM $ZnCl_2$, 20% glycerol, 10 mM DTT).

TFIIB

Full-length *S. cerevisiae* TFIIB with a C-terminal His-tag was expressed for 15 h at 20 °C in *E. coli* BL21 (DE3) RIL. Cells were harvested by centrifugation, resuspended in buffer B (50 mM Tris-HCl, pH 7.5 (25 °C), 500 mM NaCl, 10 mM imidazole, 5% glycerol, 20 µM PMSF, 5 mM -mercaptoethanol) and lyzed by sonication. After centrifugation, the supernatant was loaded onto a 2 ml Ni-NTA column (Qiagen), equilibrated with buffer B. After washing, bound protein was eluted with buffer B containing 1 M imidazole. The protein was diluted 1:3 with buffer C (50 mM Tris-HCl pH 7.5 (25 °C), 5 mM DTT) and was further purified by cation exchange chromatography (MonoS 5/50 GL, GE Healthcare) using buffer C and a linear gradient from 50mM to 1M NaCl. After concentration, a final gelfiltration purification step was performed using a Superose 6 HR column (GE Healthcare) equilibrated with Pol II buffer containing 5% glycerol.

TFIIB-ybbR

Site specific labeling of TFIIB was achieved by the Sfp phosphopantetheinyl transferase-catalyzed ybbR-tag (YSLEFIASKLA, labeling occurs at underlined serine) protein labeling technique (Yin et al., 2006). Two different TFIIB-ybbR constructs were engineered in order to introduce a labeling site close to the cyclin folds. The first one was prepared with the ybbR-tag just N-terminal of the first cyclin fold of TFIIB between residue 122 and 123 (TFIIB-(122)ybbR), and the second one with the ybbR-tag in the TFIIB C-terminal tail (TFIIB-(C-term)ybbR). TFIIB-ybbR was expressed and purified as described above for wt TFIIB. Expression and purification of Sfp (plasmid gift from C. Kaiser) and synthesis of an Alexa 647-CoA conjugate from Coenzyme A (Sigma Aldrich) and Alexa 647-C2-Maleimide (Molecular Probes) was done as

described (Yin et al., 2006). TFIIB labeling was performed at room temperature for 30 min using 5 µM TFIIB-ybbR, 1 µM Sfp and 10-20 µM Alexa 647-CoA. Free dye was removed using Bio-Spin 6 columns (Biorad) equilibrated with assembly buffer containing 5% glycerol (50 mM HEPES, 40 mM ammonium sulfate, 10 µM $ZnCl_2$, 5% glycerol, 10 mM DTT).

Nucleic Acids Scaffolds

Nucleic acids scaffolds used to build OCs for smFRET measurements were constructed from 55 nt long DNA oligomers containing an 11 nt mismatch region around the TSS from register +2 to -9 (Figure 1A). The sequence contained at the 5'-terminus a TFIIB response element followed by a TATA box. For surface immobilization of the OCs in the measurement chamber, the non-template DNA strand had biotin attached at the 5'-end via a C6-amino linker. DNA oligomers were annealed as described (Andrecka et al., 2008). The FRET network used for NPS analysis of the OC is schematically represented in Figure 1C: Template DNA strands were labeled at register -10, +3, +7 or +12 with Tamra (donor) or Alexa 647 (acceptor) and used as satellites in order to map the position of antenna dyes on the upstream non-template DNA at register -30, -23 or -18, on TBP or on TFIIB. For localization of the antennas on the upstream non-template DNA, Tamra labeled non-template DNA strands were combined with Alexa 647 labeled template strands. For localization of the antennas on TBP and TFIIB (Alexa647 labeled), Tamra labeled template strands were annealed with unlabeled non-template DNA strands. For control measurements, bubbles containing 17 nt RNA with the sequence 5'-AUGCAUAAAGACCAGGC-3' were used and annealed as described (Andrecka et al., 2008). All DNA and RNA strands were purchased from IBA (Göttingen, Germany).

OC Assembly

S. cerevisiae Pol II-TFIIF as well as 10- and 12-subunit *S. cerevisiae* Pol II were expressed and purified as described (Chen et al., 2010). For OC assembly, TFIIB (60 pmol) was first added to the nucleic acid scaffold (15 pmol), followed by TBP (60 pmol) and the Pol II-TFIIF complex (10 pmol), yielding OCs in a final volume of 50 µl assembly buffer (50 mM HEPES, 40 mM ammonium sulfate, 10 µM $ZnCl_2$, 5% glycerol) containing 10 mM DTT. The mixture was incubated with slight agitation at room temperature for 1 h, followed by gelfiltration purification using a Superose 6 PC3.2 column (GE Healthcare). The fractions containing the complete OC were combined and diluted 100-1000 times in assembly buffer for sm-FRET measurements. In some experiments, labeled Rpb7-C150 was used as a satellite. Expression and purification of the single-cysteine mutant of Rpb4/7 as well as labeling of Rpb7-C150 was performed as described (Andrecka et al., 2008). Labeled Rpb4/7 was then added in 15-20 fold molar excess to Pol II-TFIIF and exchange of unlabeled for labeled Rpb4/7 was allowed for 10-14 h at 4 °C. Subsequently, OC assembly was conducted as described above. OCs without TFIIF were assembled using the same assembly procedure. For OCs without TFIIF but with labeled Rpb4/7, 10 subunit core Pol II was used and pre-assembled with a 5 fold molar excess of labeled Rpb4/7 for 15 min prior to OC assembly.

Electrophoretic Mobility Shift Assay

Typically, in a total volume of 15 µl assembly buffer containing 10 mM DTT, 20 pmol of TBP, 20 pmol of TFIIB-ybbR labeled with Alexa 647 and 5 pmol of the Pol II-TFIIF complex were sequentially added to 5 pmol of Tamra-labeled nucleic acids and incubated for 1h at 25 °C. The samples were loaded on a 6-20 % gradient native

polyacrylamide gel (Biorad) in 1x TBE buffer. The gel was prerun at 80 V for 45 min and the pockets were rinsed thoroughly before sample loading. Electrophoresis was carried out at 4 °C at a voltage of 120 V for 2 h and fluorescent bands were visualized using a Typhoon Imager (GE Healthcare) at 532 nm and 647 nm.

In vitro Transcription Assay

In an *in vitro* transcription assay, *HIS4* promoter activity was tested using wild type (wt) nuclear extract (NE) from *S. cerevisiae* BY4741, as well as NE from a temperature-sensitive *S. cerevisiae* strain SHY245 carrying a point mutation in the TFIIB gene *SUA7*. To this temperature-sensitive TFIIB NE, either protein storage buffer (negative control) or 5 pmol of wt TFIIB, TFIIB-(122)ybbR or TFIIB-(C-term)ybbR was added prior to start of *in vitro* transcription. NEs were prepared from 3 l of either BY4741 or SHY245 yeast culture as described previously (Seizl et al., 2011). In vitro transcription and analysis by primer extension were performed as described (Seizl et al., 2011).

smFRET Measurements: Experimental Setup and Data Analysis

smFRET experiments were performed on a homebuilt prism-based total internal reflection fluorescence microscope (TRIFM) as described previously (Andrecka et al., 2008). Briefly, a frequency-doubled Nd:YAG laser (532 nm, Spectra-Physics) was used for the excitation of donor molecules and a He-Ne laser (637 nm, Coherent) for the direct excitation of the acceptor molecules. The fluorescence signal of donor and acceptor was split by the use of a dichroic mirror. Fluorescence intensity was collected through a water-immersion objective (Plan Apo 60X, NA 1.2, Nikon) and directed to an EMCCD camera (iXon DU-897E-CS0-BV, Andor). Microfluidic chambers were constructed similar to a procedure published previously (Andrecka et

al., 2008). Pol II OCs were immobilized onto the chamber surface via PEG-Biotin/Neutravidin/Biotin attachment. All measurements were recorded with an integration time of 100 ms per frame for a total duration time of 40 s. The acquired data was analyzed using custom-written MATLAB (The MathWorks) software. We used a fully automated routine to find FRET pairs using an intensity threshold for the acceptor signal during the FRET measurement. The algorithm then calculates and subtracts the local background and computes fluorescence trajectories (Andrecka et al., 2008). For the calculation of the FRET efficiency of each individual FRET pair, we used the following formula:

$E = \frac{I_A - \beta I_D}{I_A + \gamma I_D}$, 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 acceptor and donor channels, and I and I' are intensities before and after acceptor photobleaching, respectively. and are correction factors, where accounts for the leakage of the donor emission into the acceptor channel, and 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. Direct excitation of the acceptor was so low that including a correction factor did not alter the observed FRET efficiencies and therefore, it was disregarded in the analysis. The resulting histograms were computed from every frame of the time traces smoothed by a 5 point moving average (frame-wise histogram). The histograms were fitted with one or more Gaussian distributions and the mean FRET efficiency as well as its standard error was determined from the fit. The overall error of the FRET efficiency was estimated by performing error propagation using the

standard error from the fit and systematic errors and yielded to about 2 % and in a few cases to 3 % when only a low number of molecules with a broad distribution of FRET efficiencies could be obtained. The mean FRET efficiency and its overall error were used for further analysis by NPS (Muschielok et al., 2008; Muschielok and Michaelis, 2011). An overview of all smFRET data is given in Table S1 and frame-wise FRET efficiency histograms are shown in figure S3-S5.

Nano-Positioning System

Dye Position Priors

Satellite dye molecules were attached via flexible linkers to positions on the template DNA and on Pol II within OCs. The attachment points were assumed to be unchanged compared to their position within the Pol II elongation complex (EC), of which a structure has been solved by x-ray crystallography (pdb ID: 1Y1W) (Kettenberger et al., 2004). This assumption was validated by distance measurements discussed in the main text. Hence, the attachment points are known from the x-ray structure, whereas the precise location and the orientation of the dye molecules are not. For NPS analysis, we therefore calculated the volume that is sterically accessible to the dye molecules, given the point of attachment, the size of the dye molecule, the linker length and the diameter of the linker (Muschielok et al., 2008). Briefly, the satellites were approximated by a sphere of a diameter d_{dye} = 7 Å, linked to template DNA and protein by flexible linkers with a diameter of d_{linker} = 4.5 Å and a length of L_{linker} = 13 Å (internal DNA labels) and 7 Å (protein labels), respectively (Table S2). Additionally, we assumed vanishing probability density within the volume already occupied by the reference structure of the Pol II EC shrunk by 2 Å to account for uncertainties in the x-ray structure. Each satellite position within the accessible

volume was given equal probability (uniform prior) (Muschielok et al., 2008). The position and size of the computed satellite priors are displayed in Figure S1.

Antenna dye molecules were attached via flexible linkers to positions on the upstream non-template DNA (NT(-18)), the TATA box (NT(-23), NT(-30)), TBP (C128) and TFIIB ((C-term)ybbR and (122)ybbR). All attachment points were unknown in the coordinate system of the Pol II EC and were to be determined within, by application of the NPS. However, we used information available about the local macromolecule structure - given that this structure is not subject to conformational change when incorporated into an OC - and simulated accessible volumes relative to these structures in a similar way like for the satellite dyes (Figure 1C). By displacing the structures also the antenna dyes were moved to other positions, so that the structures served as frames of reference to each antenna dye. To simulate the accessible volume of the antenna on the upstream non-template DNA we used the structure of a 12 nt B-DNA (pdb ID: 1BNA) (Drew et al., 1981). For the TATA box and TBP positions, the simulation was based on a structure of a B-core-TBP-TATA-element complex (pdb ID: 1VOL without TFIIB) (Nikolov et al., 1995). The final position and orientation of the antennas was calculated from their position and orientation relative to the local structure and from the position and orientation of the reference frames attached roughly to the center of mass of each local structure. The volume accessible to the origin of the docked reference frames was set to a cube of 200 Å x 290 Å x 220 Å with the origin at position (x,y,z) = (0, 105, -90) Å, while their orientation remained completely free. The antennas attached to TFIIB were treated in a different way because the conformation TFIIB adopts in the OC is likely to deviate from its conformation in the TFIIB-Pol II structure (Kostrewa et al., 2009) or in the B-core-TBP-TATA structure (Nikolov et al., 1995). Therefore, we restricted the TFIIB

antenna priors to a volume not occupied by Pol II in the EC structure. To this end, we assumed zero probability density within the volume occupied by Pol II shrunk by 2 Å to account for uncertainties in the x-ray structure, and equal probability densities elsewhere in a cubic box around Pol II with a side length of 300 Å. The downstream DNA in the OC was observed to adopt two conformations, the conformation expected from the EC structure and an alternate conformation (see main text for details). We examined the alternate conformation by global NPS. To this aim, we treated the dye molecules attached to the downstream DNA as antennas with unknown position in combination with the FRET efficiencies of the side peaks present in all histograms of measurements between sites on the downstream DNA and sites on the upstream DNA or Rpb7-C150. Yet, for the main peaks in the FRET efficiency histograms, the dye molecules on the downstream DNA served as satellites with known position. No reference frame was assigned to the antenna dyes on the downstream DNA because the DNA conformation is likely to deviate from the structure of a straight B-DNA in the alternate conformation, especially close to the melted region at register (+3). Therefore, we restricted the downstream DNA antenna priors to a volume not occupied by Pol II in the EC structure (Kettenberger et al., 2004), in a similar way than for the TFIIB antenna priors.

Global Determination of the Positions and Orientations of Antenna Dyes and Reference Frames by NPS Analysis

The global NPS analysis emanates from a static model (Muschielok and Michaelis, 2011). Hence, each fluorophore was assumed to capture a static position $\boldsymbol{x}_i$. Its transition dipole moment was assumed to perform orientation fluctuations much faster than the time-scale of the fluorescence lifetime and the average transition dipole moment orientation of each fluorophore was assumed to be static and equal

for both absorption and emission. Moreover, the docked reference frames were treated as rigid bodies and interactions of fluorophores attached to one reference frame with parts of macromolecules within other reference frames were neglected.

Bayesian parameter estimation was applied using all measured average FRET efficiencies $\{E_{ij}\}$, where $\{E_{ij}\} = E_{i1}, E_{i2}, E_{i3},\ldots, E_{iNsat}$, are the FRET efficiencies between the ith antenna (i = 1, 2,…, N_{ant}) and each of N_{sat} satellites, as well as all measured fluorescence anisotropies $\{r_i,r_j\}$ of the fluorophores (i = 1, 2,…, N_{ant} and j = 1, 2,…, N_{sat}) (Table 1), to simultaneously infer position $\boldsymbol{x}_i$ and orientation $_i$ of all fluorophores within the "laboratory" coordinate system of the Pol II EC. Moreover, the information of all fluorophores linked to a reference frame could be used to also position and orient the reference frame itself with respect to the Pol II EC coordinate system. As a result, we obtain as posterior the three dimensional probability density function (PDF) $p(\{x_i, {}_i\} \mid \{E_{ij}\}, \{r_i,r_j\}, I)$ of the antenna dyes that represent the inferred information given the measured data, $\{E_{ij}\}$ and $\{r_i,r_j\}$, and the background information I such as attachments points and isotropic Förster distances. As a feature of the docking NPS analysis, the marginal position PDF of any desired point in the docked reference frames relative to the Pol II EC coordinate system can be extracted in addition to the antenna dye position PDFs. The complete calculation was performed using a custom implementation of the nested sampling algorithm (Skilling, 2006) based on Markov chain Monte-Carlo in C and MATLAB (The MathWorks). The NPS software is available free of charge at www.uni-ulm.de/nawi/nawi-biophys.html (Muschielok and Michaelis, 2011).

<u>Determination of Isotropic Förster Distances and Anisotropies</u>

For each donor-acceptor pair the isotropic Förster radius R_0^{iso} was determined using standard procedures (Vamosi et al., 1996). First, the quantum yield of the donor sample was determined by comparing its fluorescence to that of the standard Rhodamine 101 in ethanol (QY = 100 %). Second, overlap integrals were calculated from recorded donor fluorescence emission spectra (emission wavelength 528-700 nm, excitation wavelength 523 nm) and acceptor absorption spectra (400-700 nm). Finally, n = 1.35 and ^{2} = 2/3 were used to calculate isotropic Förster distances R_0^{iso}. In order to account for uncertainties in the Förster distance due to orientation effects, we measured the steady state fluorescence anisotropies of the donor and acceptor dyes for all attachment sites using a steady state fluorescence spectrometer (Edinburgh Instruments F900) (Table S1). Both, isotropic Förster distances as well as fluorescence anisotropies were used as prior information in the global NPS analysis (Muschielok and Michaelis, 2011).

Supplemental References

Seizl, M., Lariviere, L., Pfaffeneder, T., Wenzeck, L., and Cramer, P. (2011). Mediator head subcomplex Med11/22 contains a common helix bundle building block with a specific function in transcription initiation complex stabilization. Nucleic Acids Res.

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.

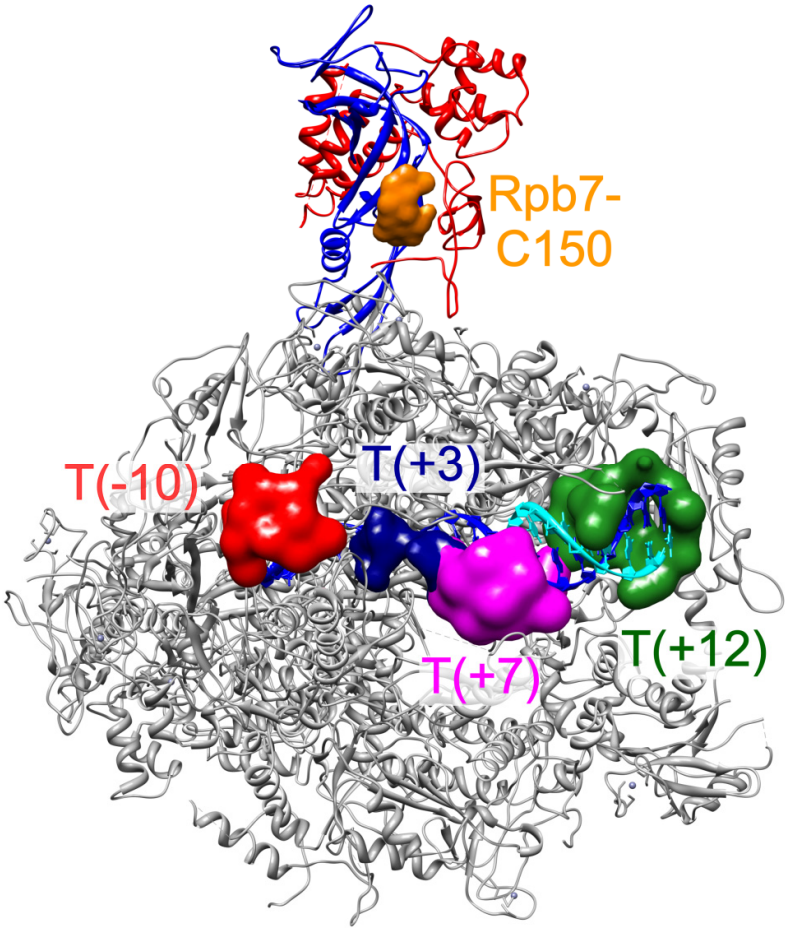

Rpb7-
C150
T(-10)
T(+3)
T(+7)
T(+12)

Figure S1. Satellite Position Priors within Pol II EC Structure Lacking RNA, Related to Figure 1

As a model for the OC we used the EC structure (Kettenberger et al., 2004) (pdb ID: 1Y1W) without the RNA (see main text for discussion). Pol II core is shown in grey, Rpb4 in blue, Rpb7 in red, template DNA strand in blue and non-template DNA strand in cyan. Each solid density represents the accessible volume of a satellite revealed by modeling using a flexible linker model. Each satellite prior is shown in the color that is used throughout the paper for this satellite dye position. The image was prepared using Chimera (Pettersen et al., 2004).

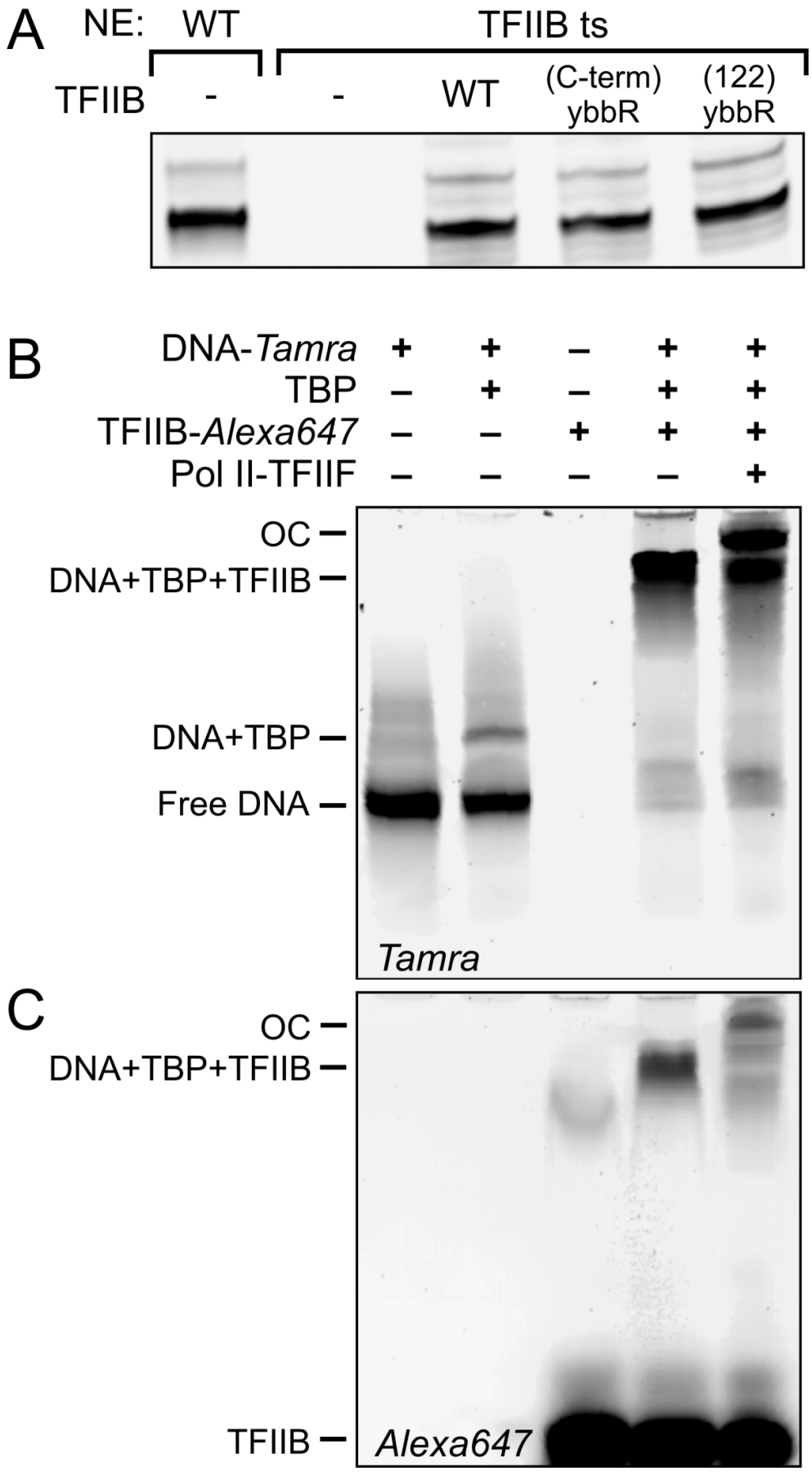

A
NE:
WT
TFIIB ts
TFIIB
-
-
WT
(C-term) ybbR
(122) ybbR
B
DNA-Tamra + + – + +
TBP – + – + +
TFIIB-Alexa647 – – + + +
Pol II-TFIIF – – – – +
OC
DNA+TBP+TFIIB
DNA+TBP
Free DNA
Tamra
C
OC
DNA+TBP+TFIIB
TFIIB
Alexa647

Figure S2. In Vitro Transcription of TFIIB-ybbR Mutants and Electrophoretic Mobility Shift Assay of OC Assembly, Related to Figure 1

(A) *In vitro* transcription assay testing *HIS4* promoter activity of TFIIB-ybbR mutants. Analysis of transcription products was performed by primer extension using fluorescently labeled primers (Cy5, excitation: 633 nm, detection: 655-685 nm). Nuclear extract (NE) from wild type (WT) yeast strain shows full *in vitro* transcription activity (lane 1), whereas NE from a TFIIB temperature-sensitive yeast strain is essentially inactive (lane 2). Transcriptional activity can be restored when recombinant WT TFIIB and the mutants TFIIB-(122)ybbR and TFIIB-(C-term)ybbR are supplied (lane 3-5).

(B and C) Fluorescent gel images of OCs prepared from pre-melted DNA bubble, TBP, TFIIB-(C-term)ybbR and Pol II-TFIIF. (B) Tamra labeled DNA was detected in the green channel (excitation: 532 nm, detection: 565-595 nm). To the DNA bubble (lane 1), protein factors were sequentially added: first TBP (lane 2), then labeled TFIIB-(C-term)ybbR (lane 4) and finally Pol II in complex with TFIIF (lane 5). Each additional protein factor produced an additional upshift of the nucleic acid scaffold in the gel showing its binding to the DNA. (C) Alexa 647 labeled TFIIB-(C-term)ybbR was detected in the red channel (excitation: 633 nm, detection: 655-685 nm). Lane 3 shows the signal of TFIIB alone. The fluorescent band is shifted up due to the binding of TFIIB to the DNA-TBP assembly (lane 4) and to Pol II in the OC (lane 5).

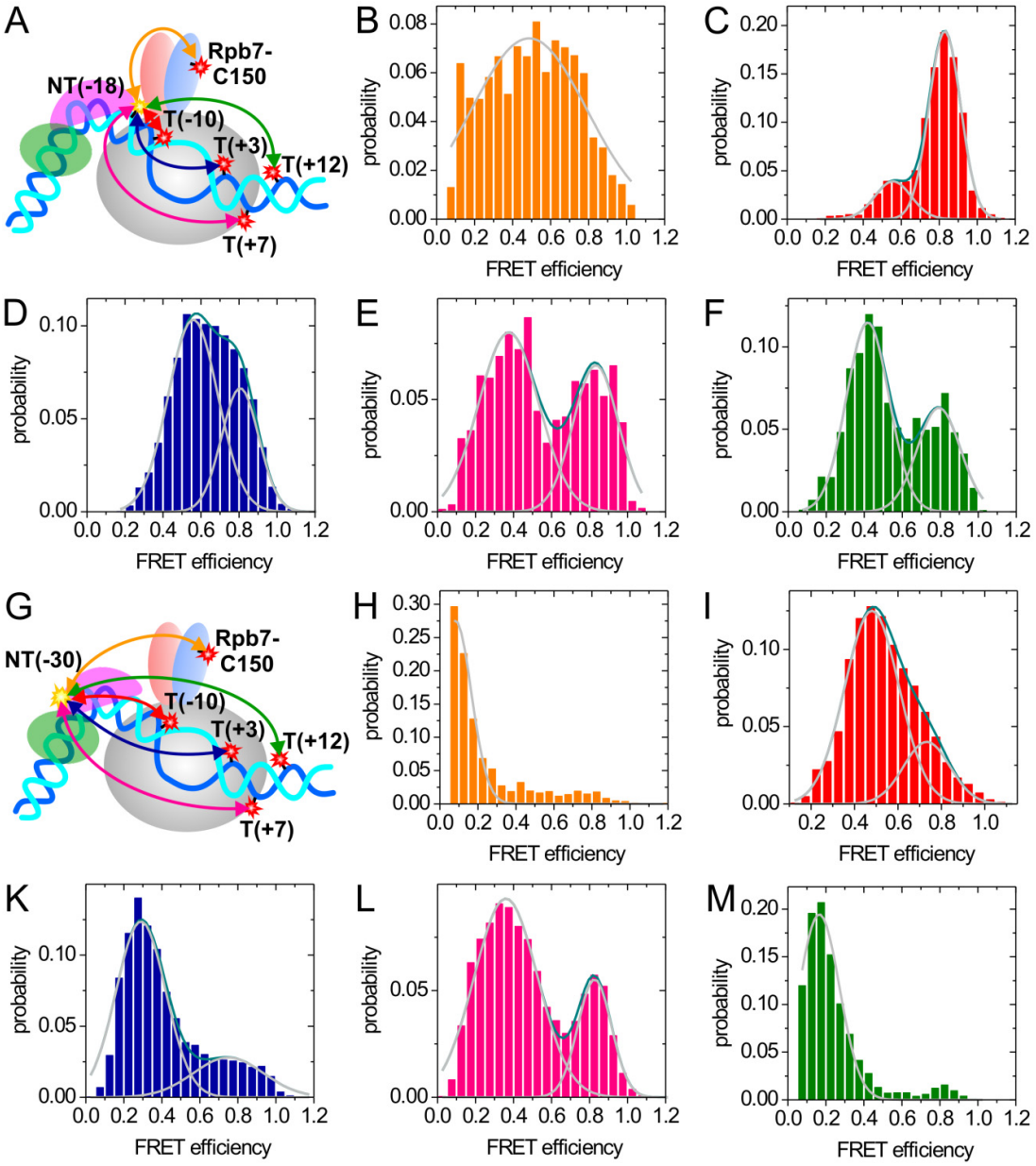

A
Rpb7-
C150
NT(-18)
T(-10)
T(+3)
T(+12)
T(+7)
B
C
D
E
F
G
NT(-30)
H
I
K
L
M
probability
FRET efficiency

Figure S3. smFRET Measurements for Localization of NT-DNA(-18) and NT-DNA(-30), Related to Figures 3 and 5

(A) Cartoon illustrating the set of smFRET measurements for the localization of antenna NT-DNA(-18) by NPS. The measurements are indicated by double-headed arrows colored orange for satellite Rpb7-C150, red for T-DNA(-10), blue for T-DNA(+3), pink for T-DNA(+7) and green for T-DNA(+12). An OC is presented schematically using the same colors as in Figure 1 A. The same color coding will be used in the following figures.

(B–F) Sm FRET measurements from the five satellites to antenna NT-DNA(-18) on the upstream DNA. Framewise FRET efficiency histograms are shown in the color corresponding to the color of the respective satellite dye position (see Figure S1 and color of arrows in panel A): (B) satellite Rpb7-C150, (C) T-DNA(-10), (D) T-DNA(+3), (E) T-DNA(+7), (F) T-DNA(+12). The same color coding will be used in all following histograms. The FRET efficiency histograms were fitted with one or two Gaussians (gray, individual Gaussian fit; dark cyan, combined fit; Same color coding will be used in all following histograms). From the fit, the mean FRET efficiency and its standard error were determined and are summarized in table S1.

(G) Cartoon illustrating the set of smFRET measurements for the localization of antenna NT-DNA(-30) by NPS.

(H–M) Framewise FRET efficiency histograms of the measurements from the five satellites to antenna NT-DNA(-30) on the TATA box: (H) satellite Rpb7-C150, (I) T-DNA(-10), (K) T-DNA(+3), (L) T-DNA(+7), (M) T-DNA(+12).

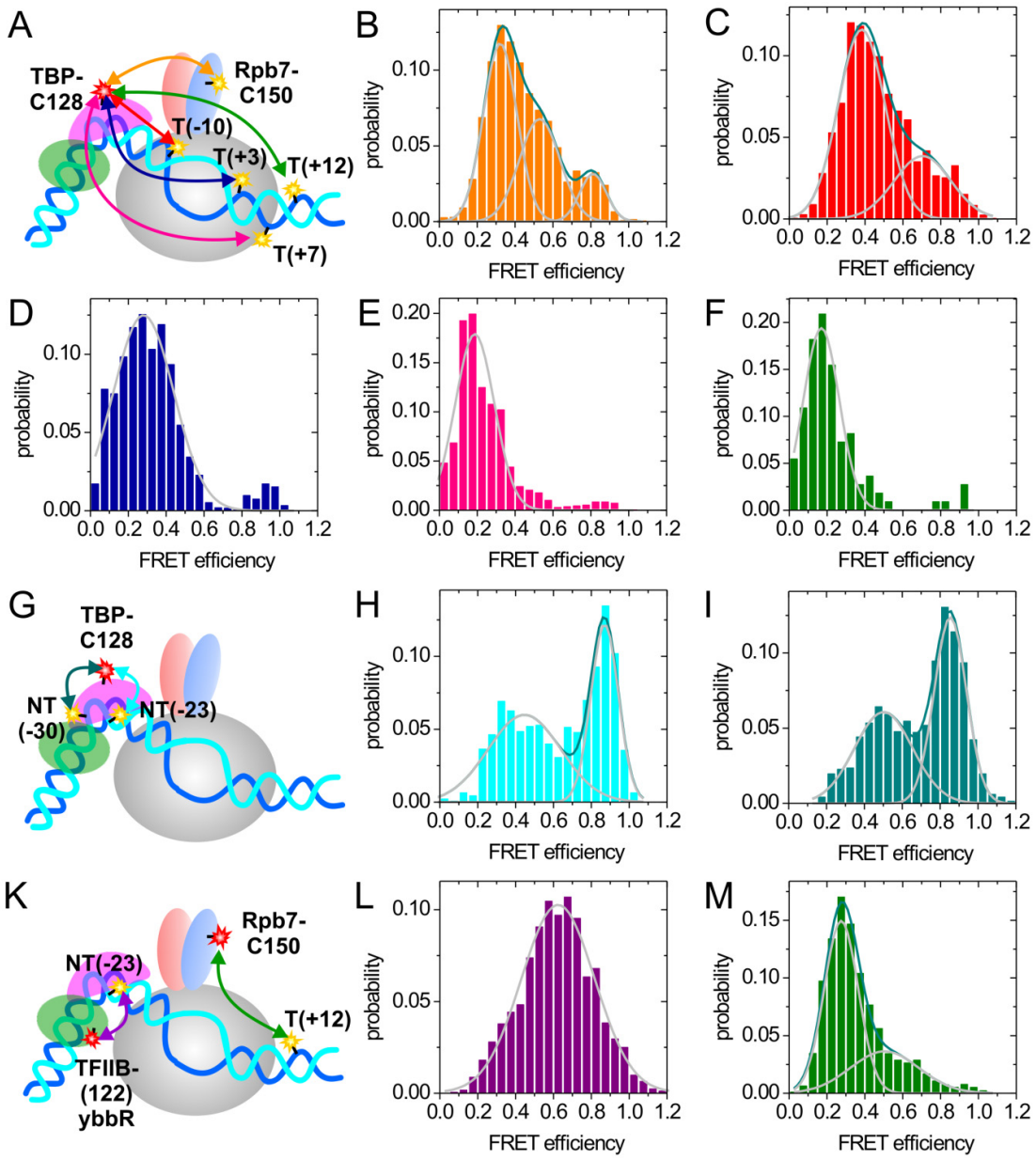

A
TBP-
C128
Rpb7-
C150
T(-10)
T(+3)
T(+12)
T(+7)
B
C
D
E
F
G
TBP-
C128
NT
(-30)
NT(-23)
H
I
K
Rpb7-
C150
NT(-23)
T(+12)
TFIIB-
(122)
ybbR
L
M
probability
FRET efficiency

Figure S4. smFRET Measurements for Localization of TBP-C128 and Distance Tests between TBP-C128–TATA, TFIIB-(122)ybbR–TATA, and T-DNA(+12)–Rpb7-C150, Related to Figures 3 and 5

(A) Cartoon illustrating the set of smFRET measurements for the localization of antenna TBP-C128 by NPS.

(B–F) Framewise FRET efficiency histograms of the measurements from the five satellites to antenna TBP-C128: (B) satellite Rpb7-C150, (C) T-DNA(-10), (D) T-DNA(+3), (E) T-DNA(+7), (F) T-DNA(+12).

(G) Cartoon illustrating the smFRET measurements between antenna TBP-C128 and the antennas on the TATA box.

(H and I) Framewise FRET efficiency histograms of the measurements illustrated in panel (G): (H) between TBP-C128 and NT-DNA(-23) and (I) between TBP-C128 and NT-DNA(-30).

(K) Cartoon illustrating the smFRET measurements between antenna TFIIB-(122)ybbR and antenna NT-DNA(-23) on the TATA box and between the satellites T-DNA(+12) and Rpb7-C150.

(L) Framewise FRET efficiency histograms of the measurement between the antennas TFIIB-(122)ybbR and NT-DNA(-23) illustrated in panel (K).

(M) Framewise FRET efficiency histograms of the measurement between the satellites T-DNA(+12) and Rpb7-C150 illustrated in panel (K).

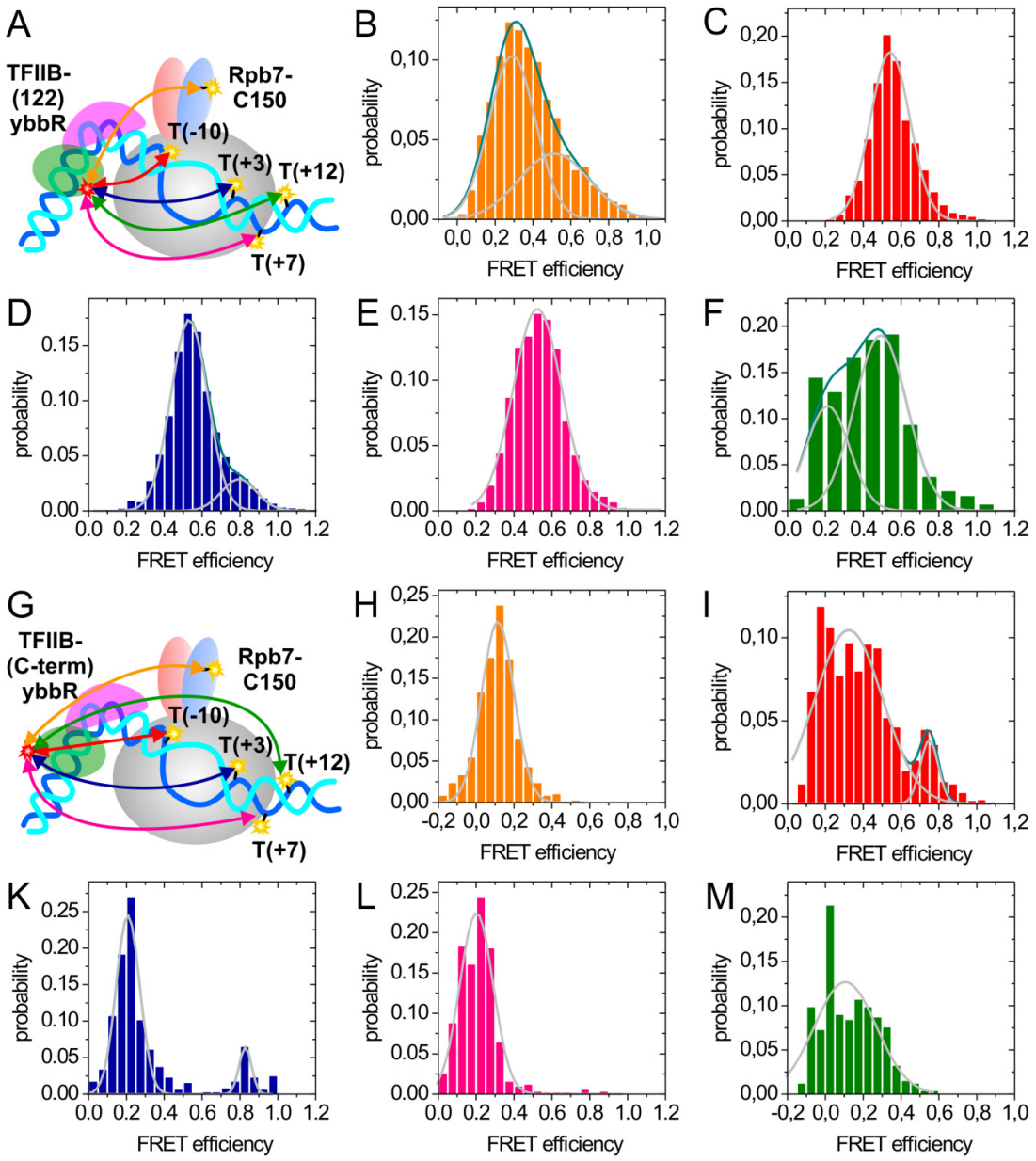

A
TFIIB-
(122)
ybbR
Rpb7-
C150
T(-10)
T(+3)
T(+12)
T(+7)
B
C
D
E
F
G
TFIIB-
(C-term)
ybbR
Rpb7-
C150
T(-10)
T(+3)
T(+12)
T(+7)
H
I
K
L
M
probability
FRET efficiency

Figure S5. smFRET Measurements for Localization of TFIIB-(C-term)ybbR and TFIIB-(122)ybbR, Related to Figure 6.

(A) Cartoon illustrating the set of smFRET measurements for the localization of antenna TFIIB-(122)ybbR by NPS.

(B–F) Framewise FRET efficiency histograms of the measurements from the five satellites to antenna TFIIB-(122)ybbR: (B) satellite Rpb7-C150, (C) T-DNA(-10), (D) T-DNA(+3), (E) T-DNA(+7), (F) T-DNA(+12).

(G) Cartoon illustrating the set of smFRET measurements for the localization of antenna TFIIB-(C-term)ybbR by NPS.

(H–M) Framewise FRET efficiency histograms of the measurements from the five satellites to antenna TFIIB-(C-term)ybbR: (H) satellite Rpb7-C150, (I) T-DNA(-10), (K) T-DNA(+3), (L) T-DNA(+7), (M) T-DNA(+12).

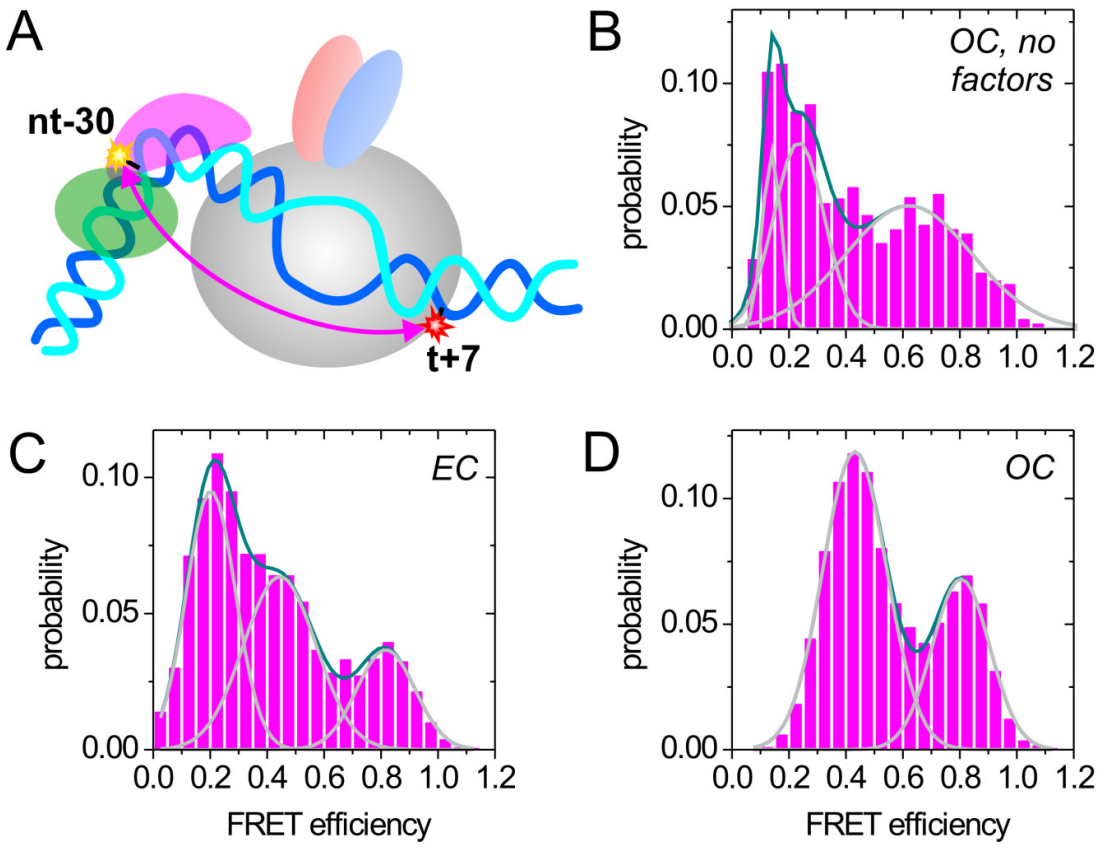

A
nt-30
t+7
B
OC, no factors
probability
0.10
0.05
0.00
0.0 0.2 0.4 0.6 0.8 1.0 1.2
C
EC
probability
0.10
0.05
0.00
0.0 0.2 0.4 0.6 0.8 1.0 1.2
FRET efficiency
D
OC
probability
0.10
0.05
0.00
0.0 0.2 0.4 0.6 0.8 1.0 1.2
FRET efficiency

Figure S6. Position of Upstream DNA in OC Is Stabilized by Initiation Factors

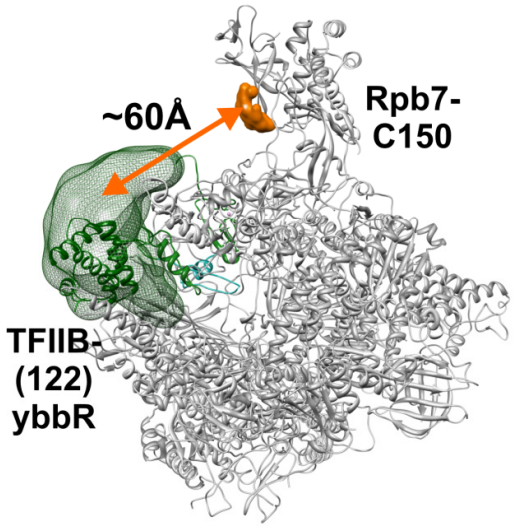
(A) Cartoon illustrating the control measurements. smFRET was measured between the satellite dye T-DNA(+7)-Alexa 647 and the antenna dye NT-DNA(-30)-Tamra within the EC, the OC lacking all initiation factors and the OC (shown here).

(B) Framewise FRET efficiency histogram for EC composed of a DNA-RNA bubble with a 17 nt RNA transcript and Pol II. The data covers a broad range of FRET efficiencies and has to be fitted using at least 3 Gaussians (white, individual Gaussian fits and blue, combined fit).

(C) Framewise FRET efficiency histogram for a factorless OC composed only of DNA bubble and Pol II. Similarly to the data for the EC, the factorless OC yields a broad FRET efficiency histogram that is best described with three Gaussians.

(D) Framewise FRET efficiency histogram for the OC. The observed histogram is different than the histograms for EC and OC lacking all initiation factors and shows two distinct FRET efficiency distributions. The side peak is caused by the conformational dynamics of the downstream DNA in the OC (see main text for details). A second distance was investigated in the same way as presented here and yielded similar results (data not shown).

Pol II-TFIIB
X-ray crystal structure



smFRET experiment
of Pol II-TFIIB complex

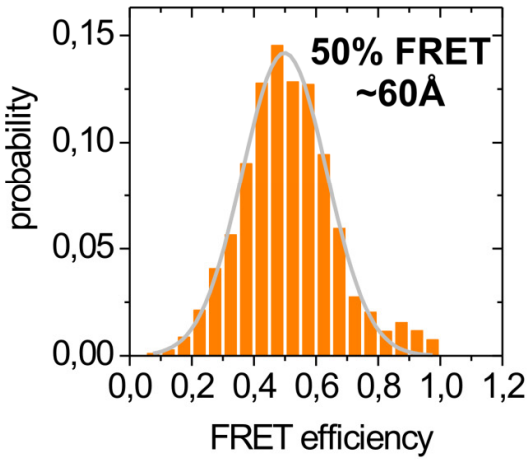


Figure S7. Pol II-TFIIB Complex Captures a Similar Structure in smFRET Measurement Conditions and in X-ray Crystallography, Related to Figure 7

smFRET was measured between Rpb7-C150 and TFIIB-(122)ybbR in Pol II-TFIIB complexes and the distance resulting from the FRET efficiency histogram with a mean FRET efficiency of 50% (right) was compared to the distance in the X-ray structure (left). Both methods revealed a similar distance of about 60Å (assuming the isotropic R_0 value) showing that Pol II-TFIIB complexes capture the same structure in smFRET measurement conditions as in the crystal structure. Note that the corresponding smFRET histogram of the open complex (Figure S5b) looks dramatically different and shows a main peak with a FRET efficiency of about 30%.

Table S1. Overview of Experimental Data Measured for OCs

Antenna	Satellite	Anisotropy (Donor)	Anisotropy (Acceptor)	R_0 [Å]	# of molecules	FRET efficiency	% of probability
Localization of NT-DNA(-18)							
NT-DNA(-18)-T	Rpb7-C150-A	0.26	0.25	59	128	49	-
NT-DNA(-18)-T	T-DNA(-10)-A	0.26	0.3	58	186	56*/83	19/81
NT-DNA(-18)-T	T-DNA(+3)-A	0.26	0.32	58	480	56/81*	67/33
NT-DNA(-18)-T	T-DNA(+7)-A	0.26	0.33	59	192	37/83*	62/38
NT-DNA(-18)-T	T-DNA(+12)-A	0.26	0.34	57	172	42/79*	64/36
Localization of NT-DNA(-23)							
NT-DNA(-23)-T	Rpb7-C150-A	0.28	0.25	65	49	33	-
NT-DNA(-23)-T	T-DNA(-10)-A	0.28	0.3	60	447	66	-
NT-DNA(-23)-T	T-DNA(+3)-A	0.28	0.32	60	382	31/64*	84/16
NT-DNA(-23)-T	T-DNA(+7)-A	0.28	0.33	61	346	33/72*	64/36
NT-DNA(-23)-T	T-DNA(+12)-A	0.28	0.34	61	576	33/74*	81/19
Localization of NT-DNA(-30)							
NT-DNA(-30)-T	Rpb7-C150-A	0.3	0.25	62	70	12	-
NT-DNA(-30)-T	T-DNA(-10)-A	0.3	0.3	57	377	48/73*	79/21
NT-DNA(-30)-T	T-DNA(+3)-A	0.3	0.32	57	456	29/75*	61/39
NT-DNA(-30)-T	T-DNA(+7)-A	0.3	0.33	58	349	43/80*	66/34
NT-DNA(-30)-T	T-DNA(+12)-A	0.3	0.34	56	587	16/81*	89/11
Localization of TBP-C128							
TBP-C128-A	Rpb7-C150-C	0.30	0.25	57	139	32/53*/81*	53/36/11
TBP-C128-A	T-DNA(-10)-T	0.32	0.25	60	199	38/70*	71/29
TBP-C128-A	T-DNA(+3)-T	0.30	0.25	62	159	32	-
TBP-C128-A	T-DNA(+7)-T	0.31	0.25	61	129	19	-
TBP-C128-A	T-DNA(+12)-T	0.22	0.25	61	110	17	-
TBP-C128-A	NT-DNA(-23)-T	0.28	0.25	62	91	45*/87	44/56
TBP-C128-A	NT-DNA(-30)-T	0.30	0.25	59	179	50*/85	43/57
Localization of TFIIB-(122)ybbR							
TFIIB-(122)ybbR-A	Rpb7-C150-C	0.30	0.25	55	104	29/51*	61/39
TFIIB-(122)ybbR-A	T-DNA(-10)-T	0.32	0.25	58	127	53	-
TFIIB-(122)ybbR-A	T-DNA(+3)-T	0.30	0.25	60	247	53/79*	86/14
TFIIB-(122)ybbR-A	T-DNA(+7)-T	0.31	0.25	59	85	52	-
TFIIB-(122)ybbR-A	T-DNA(+12)-T	0.22	0.25	59	99	49/21*	67/33
TFIIB-(122)ybbR-A	NT-DNA(-23)-T	0.28	0.25	60	53	55	-
Localization of TFIIB-(C-term)ybbR							
TFIIB-(C-term)ybbR-A	Rpb7-C150-C	0.30	0.23	55	22	11	-
TFIIB-(C-term)ybbR-A	T-DNA(-10)-T	0.32	0.23	58	45	34/75*	91/09
TFIIB-(C-term)ybbR-A	T-DNA(+3)-T	0.30	0.23	60	24	21/83*	88/12
TFIIB-(C-term)ybbR-A	T-DNA(+7)-T	0.31	0.23	59	24	20	-
TFIIB-(C-term)ybbR-A	T-DNA(+12)-T	0.22	0.23	59	16	11	-
Localization of satellites within OC							
T-DNA(-10)-T	Rpb7-C150-A	0.32	0.25	63	34	44	-
T-DNA(+3)-T	Rpb7-C150-A	0.30	0.25	65	114	44/ 65*	57/43
T-DNA(+7)-T	Rpb7-C150-A	0.31	0.25	64	135	32/ 57*	57/43
T-DNA(+12)-T	Rpb7-C150-A	0.22	0.25	61	213	28/ 50*	67/23

*Donor dyes: T = Tamra, C = Cy3b, acceptor dye: A = Alexa 647. * side peak*

Table S2. Overview of Experimental Data Measured for ECs

Antenna	Satellite	Anisotropy (Donor)	Anisotropy (Acceptor)	R_0 [Å]	# of molecules	FRET efficiency
T-DNA(-10)-T	Rpb7-C150-A	0.32	0.25	63	70	41
T-DNA(+3)-T	Rpb7-C150-A	0.30	0.25	65	23	45
T-DNA(+7)-T	Rpb7-C150-A	0.31	0.25	64	176	37

Donor dyes: T = Tamra, acceptor dye: A = Alexa 647.

Table S3. Prior Information of Satellites and Antennas

Satellites							
Dye position	Ref frame	attachment point			L_{linker} [Å]	d_{linker} [Å]	d_{dye} [Å]
		X	y	Z			
Rpb7-C150	1Y1W	117.283	30.008	-64.547	7	4.5	7
T-DNA(-10)	1Y1W	77.878	35.805	-16.395	13	4.5	7
T-DNA(+3)	1Y1W	95.713	59.792	-15.336	13	4.5	7
T-DNA(+7)	1Y1W	88.553	73.813	-19.986	13	4.5	7
T-DNA(+12)	1Y1W	94.929	87.763	-31.497	13	4.5	7
Antennas							
Dye position	Ref frame	attachment point			L_{linker} [Å]	d_{linker} [Å]	d_{dye} [Å]
		Atom #	Atom	Residue #			
NT-DNA(-18)	1BNA moved to (0,0,0)	141	C6	7	13	4.5	7
NT-DNA(-23)	1VOL moved to (0,0,0)	493	C6	24	13	4.5	7
NT-DNA(-30)	1VOL moved to (0,0,0)	342	C6	17	13	4.5	7
TBP-C128	1VOL moved to (0,0,0)	2056	CA	226	7	4.5	7

For each satellite the coordinates *x*, *y* and *z* of the attachment point of the dye are given in Angstroem in the Pol II coordinate system. Further, for each antenna that is assigned to a reference frame the atom number, atom name and residue number of the attachment point is given in the respective reference frame. Additionally, the parameters of the flexible linker and the dye used for modeling the prior are listed: L_{linker}: length of flexible linker, d_{linker}: diameter of the linker, d_{dye}: diameter of the dye.