Functional cross-talk between the promoter and terminator of a gene has long been noted. Promoters and terminators are juxtaposed to form gene loops in several organisms, and gene looping is thought to be involved in transcriptional regulation. The general transcription factor IIB (TFIIB) and the C-terminal domain phosphatase Ssu72, essential factors of the transcription preinitiation complex and the mRNA processing and polyadenylation complex, respectively, are important for gene loop formation. TFIIB and Ssu72 interact both genetically and physically, but the molecular basis of this interaction is not known. Here we present a crystal structure of the core domain of TFIIB in two new conformations that differ in the relative distance and orientation of the two cyclin-like domains. The observed extraordinary conformational plasticity may underlie the binding of TFIIB to multiple transcription factors and promoter DNAs that occur in distinct stages of transcription, including initiation, reinitiation, and gene looping. We mapped the binding interface of the TFIIB-Ssu72 complex using a series of systematic, structure-guided in vitro binding and site-specific photocross-linking assays. Our results indicate that Ssu72 competes with acidic activators for TFIIB binding and that Ssu72 disrupts an intramolecular TFIIB complex known to impede transcription initiation. We also show that the TFIIB-binding site on Ssu72 overlaps with the binding site of symplekin, a component of the mRNA processing and polyadenylation complex. We propose a hand-off model in which Ssu72 mediates a conformational transition in TFIIB, accounting for the role of Ssu72 in transcription reinitiation, gene looping, and promoter-terminator cross-talk.

This work was supported by Welch Foundation Grant I-1790, CPRIT Grant R1119, Rita Allen Foundation Grant, the UT Southwestern Medical Center Endowed Scholar fund, and National Institutes of Health Grants GM114576 and GM121662 (to X. L.). This work also received support from the Cecil H. and Ida Green Center for Reproductive Biology Sciences and Division of Basic Research, Department of Obstetrics and Gynecology, and the Department of Biophysics, University of Texas Southwestern Medical Center, Dallas, Texas 75390 and the Department of Chemistry and Center of Systems Biology and Human Health, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, China.

Edited by Patrick Sung
protein, symplekin, which stimulates the phosphatase activity of Ssu72 (28). Both Ssu72 and symplekin are involved in pre-mRNA 3’-end cleavage and polyadenylation in yeast and human (24, 29).

Transcription-coupled functional cross-talk between promoters and terminators has long been noted. The general transcription factor TFIIID, consisting of TBP and TBP-associated factors, was shown to recruit cleavage and polyadenylation specificity factor CPSF, to which both Ssu72 and symplekin are also bound on terminators (30, 31). In addition to Ssu72, TFIIIB also interacts with the CstF64 (homolog of Rna15 in yeast) component of the cleavage stimulation factor (CstF). In particular, TFIIIB phosphorylation on serine 65 is crucial for this interaction and for CstF occupancy on promoters as well (32, 33). Furthermore, both TFIIIB and the general transcription factor TFIIF were found to be associated with and modulate the enzymatic activity of another pol II CTD phosphatase, Fcp1, which is involved in dephosphorylation of serine 2 on CTD (34, 35). Deletion of the transcription termination factor Pcf11 depletes TFIIIB and TBP from promoters (36). Notably, a point mutation of the polyadenylation site causes defects in transcription initiation (36).

These functional connections between promoters and terminators may be understood in the framework of chromatin looping. Chromatin looping is a phenomenon in which distal genomic elements are brought within close spatial proximity of each other to influence transcription. Looping occurs between enhancer elements and promoters, intergenic regions and promoters, and promoters and terminators. Promoter-terminator gene looping was first discovered in Saccharomyces cerevisiae and was later shown to be present in other organisms, such as human and plants, as well as in mitochondria and HIV (37–44). Promoter-terminator gene looping is believed to facilitate pol II recycling during transcription reinitiation (45, 46). It is also involved in sustaining transcriptional memory and enhancing transcription directionality (47–49).

TFIIB and Ssu72 occupy both promoters and terminators and are required for gene loop formation (49, 50). Mutation of TFIIIB or Ssu72 abolishes gene looping in yeast, as demonstrated by chromatin conformation capture (3C) assays (49, 50). Loss of Ssu72 leads to defects in TFIIIB recruitment to promoters and elimination of TFIIIB from terminators (50) (data not shown). To begin to reveal the mechanism of transcription-dependent promoter-terminator cross-talk that may be achieved at least in part through the interaction between TFIIIB and Ssu72, we first conducted structural analysis of the core domain of human TFIIIB captured in two new conformations. We next carried out systematic, structure-guided mapping of the binding interface of the TFIIIB-Ssu72 complex using a series of in vitro binding assays and photocross-linking analysis based on site-specifically incorporated unnatural amino acids. We also generated a knowledge-based structural model for the TFIIIB-Ssu72 complex. We propose a handoff mechanism to explain the role of the TFIIIB-Ssu72 interaction in facilitating pol II reinitiation.

### Results

**The core domain of TFIIIB displays a large degree of conformational plasticity**

We solved the crystal structure of human apo-TFIIIBc to 3.4 Á resolution by molecular replacement, with four molecules per asymmetric unit (Table 1 and Fig. S1, labeled as Molecules A–D). TFIIIBc forms two distinct, previously uncharacterized conformations (Fig. 1B). Molecules A and B are found in one conformation, and C and D are in the other conformation; for simplicity, we will only discuss molecules A and D further (Fig. 1B). Notably, each conformation is characterized by prominent translational and rotational movement of the two cyclin-like domains along the flexible C-linker between them (Fig. 1, A and B). Whereas the secondary structure of each individual cyclin-like domain remains the same, their orientation and distance with respect to each other differ (Fig. S2, A and B).

Our structure reveals an unprecedented degree of plasticity of TFIIIBc, although we also note that the conformations of TFIIIBc from our structure, like all crystal structures, are partially influenced by crystal packing. Although we cannot say for certain the extent to which these conformations occur in vivo, in the following, we interpret how this plasticity would influence the biological roles of TFIIIB. Two separate conserved regions that would contact the TATA box (e.g. residues Arg189 and Lys193 on helix H5), the BREu (e.g. residues Ala281, Val283, Arg286, and Arg290 on helix H5’), and BREd (e.g. residues 152–154 between helices H2 and H3) are solvent-accessible in both molecules A and D (Fig. 24). However, they are not spatially aligned and hence in an inactive conformation compared with

### Table 1

Data collection and refinement statistics (molecular replacement)

Data were collected from a single crystal.

| Crystal | TFIIIBc |
|---------|---------|
| **Data collection** | |
| Space group | P2₁,2₁ |
| Cell dimensions | |
| a, b, c (Å) | 108.2, 126.5, 158.4 |
| α, β, γ (degrees) | 90, 90, 90 |
| Resolution (Å) | 50.00–3.40 (3.46–3.40)¹ |
| Rmerge | 0.238 (0.00)² |
| Rfree | 0.144 (0.960) |
| I/σI | 9.15 (1.20) |
| Completeness (%) | 100.0 (100.0) |
| Redundancy | 7.2 (7.2) |
| Resolution (Å) | 50.00–3.40 (3.46–3.40) |

| **Refinement** | |
| Resolution (Å) | 50.00–3.40 (3.46–3.40) |
| No. of reflections | 30,767 |
| Rmerge | 0.194/0.214 |
| No. of atoms | 6640 |
| Protein | 6510 |
| Sulfate | 130 |
| Water | 0 |
| Average B-factors (Å²) | |
| Protein | 103.28 |
| Sulfate | 155.14 |
| Water | 0 |
| Root mean square deviations | |
| Bond lengths (Å) | 0.010 |
| Bond angles (degrees) | 1.24 |
| Ramachandran plot | |
| Favored | 95.4 |
| Allowed | 3.5 |
| Outliers | 1.1 |

¹ Values in parentheses are for highest-resolution shell.
² HKL2000 does not include values >1.
their counterparts in TFIIBc bound to TBP and the TATA box, where the two cyclin-like domains are orientated to allow simultaneous binding of helix H5 to the TATA box, the C-linker to TBP, helix H5’ to the BREu, and the loop between H2 and H3 to the BREd (Fig. 2B and Fig. S3A) (8). Interestingly, apo-TFIIBc from *Trypanosoma brucei* follows essentially the same conformation as human TFIIBc in the active ternary complex, with the exception that the C-linker is replaced by a rigid bent helix (51), implying that *T. brucei* TFIIB may always assume an active conformation regardless of whether TBP and the TATA box are bound (Fig. 2C and Fig. S3B).

Molecules A and D are also distinct from the inactive conformation of a human apo-TFIIBc determined previously by NMR (Fig. 2D and Fig. S3C) (20, 52). In the latter case, the two cyclin-like domains are stably packed against each other such that about 1900 Å² of solvent-accessible surface area is buried. In stark contrast, the same two domains in the current structures barely contact one another, displaying an extraordinary rotational and translational plasticity that is probably subjected to cellular regulation. In line with this proposal, residues Arg286 and Arg290 from helix H5/H11032 of the second cyclin-like domain change conformation when interacting with the archetypal acidic activator VP16 (20, 53). Thus, it appears possible that TFIIBc can assume multiple flexible, inactive conformations.

The C-linker of TFIIB is involved in ternary complex formation (7, 8). Interestingly, this region varies in length between species (Fig. 2A). To test whether the relative position of the two cyclin-like domains in TFIIBc is controlled by the C-linker, we appended a GSGS linker directly C-terminal of residue Leu208 located in the middle of the C-linker. This TFIIBc variant was found to bind TBP and the TATA box DNA to the same extent as wild-type protein (Fig. 2E, WT and GSGS-C), which agrees with the evolutionary divergence of TFIIB, in that *S. cerevisiae* TFIIB has an extended C-linker beyond residue Leu208 compared with human and fly TFIIB but still retains the ability to bind TBP and the TATA box (Fig. 2A). In contrast, the precise arrangement of the first part of the C-linker is critical for formation of the TFIIBc-TBP-TATA ternary complex (8), and replacing residues 200–208 with a GSG linker or adding a GSGS linker directly N-terminal of residue Lys200 significantly decreased formation of the TFIIBc-TBP-TATA box ternary complex (Fig. 2E, GSGS-C and GSGS-N). Overall, our results indicate that the semiconservative C-linker is not the only determinant of the active TFIIBc conformation.

**Ssu72 binds to the first cyclin-like domain of TFIIBc**

Previous studies have implicated a direct physical interaction between TFIIB and Ssu72 (23, 24). To understand the molecular
TFIIB-Ssu72 interaction and promoter-terminator cross-talk

A

H1

H2

H3

H4

H5

H1'

H2'

H3'

H4'

H5'

H6'

B

C

D

E

BREu Binding

BRE Ed Binding

TFIIB FL-TBPc-DNA

DNA
lar basis of this interaction, we first measured the binding affinity between TFIIB and Ssu72 by isothermal titration calorimetry (ITC), which was estimated to be around 20.2 μM (Fig. 3A).

Next, we conducted pulldown assays of C-terminal FLAG-tagged human TFIIB with His₆-tagged Ssu72. Human TFIIB constructs used for these assays included full-length TFIIB (TFIIB₈₁), TFIIB₉ (residues 107–316), and the isolated cyclin-like domain 1 (TFIIB₉₁, residues 107–206) and 2 (TFIIB₂₂, residues 207–316). Whereas TFIIB₉₁, TFIIB₂₂, and TFIIB₉ all bound to Ssu72, TFIIB₂₂ could not do so (Fig. 3B), indicating that the first cyclin-like domain of TFIIB serves as the minimal region of TFIIB for Ssu72 binding. TFIIB₂₂ is required for the full Ssu72-binding activity of TFIIB; in its absence, the binding of TFIIB₂₂ to Ssu72 became too weak to be detected by ITC (Fig. S4).

Interaction of TFIIB with Ssu72 did not appear stoichiometric in the pulldown analysis, probably due to the relatively weak interaction, as indicated by the ITC results. However, the interaction is specific, as is detailed in the studies that follow.

Charge reversal mutation of any of the positively charged residues Arg₁₈₅, Lys₁₈₉, Arg₁₉₃, and Lys₂₀₀ from helix H5 or the linker between helices H₄ and H₅ in TFIIB₁ decreased binding to Ssu72 compared with wild-type TFIIB₁, and the K₁₈₉E/R₁₉₃E mutation essentially abolished the interaction (Fig. 3C). All TFIIB₁ variants were monomeric by gel filtration chromatography analysis, indicating that the mutations did not significantly perturb protein structure (Fig. S5). Thus, a positively charged surface of TFIIB₁ proves to be necessary for Ssu72 binding. Intriguingly, an overlapping set of residues from this region of TFIIB₁ is also implicated in promoter DNA and acidic activator binding (8, 20).

In addition, we tested whether the binding interface between TFIIB₁ and Ssu72 was conserved across species using purified proteins from S. cerevisiae (yTFIIB₁ and ySsu72) and Drosophila melanogaster (dTFIIB₁ and dSsu72). In both cases, TFIIB₁ and Ssu72 were found to interact directly (Fig. 3D). Notably, whereas critical interacting residues Lys₁₈₉ and Arg₁₉₃ of human TFIIB₁ are fairly conserved in yeast (as Lys₂₀₁ and Lys₂₀₅, respectively; Fig. 2A), residues Arg₁₈₅ and Lys₂₀₀ are not conserved (His₁₉₇ and Asn₁₂₂ in yeast; Fig. 2A). In the pulldown assays, the conserved K₂₀₁E/K₂₀₅E mutation of yTFIIB₁ abolished the interaction with ySsu72 in a manner similar to the human TFIIB. All yTFIIB₁c mutants were monomeric by gel filtration analysis, indicating that mutation did not disrupt normal protein structure (Fig. S5). Surprisingly, the H₁₉₇E/N₂₁₂E mutation of yTFIIB₁ also significantly reduced binding to ySsu72 (Fig. 3D), suggesting that the mode of interaction between TFIIB₁ and Ssu72 is highly conserved from yeast to humans despite differences in protein sequence.

TFIIB₁c interacts with an Ssu72 surface formed by three discrete sequence patches

To probe Ssu72 residues responsible for TFIIB₁c binding, we used a site-specific photocross-linking approach based on incorporation of the unnatural amino acid p-benzoyl-l-phenylalanine (pBpa) via suppression of an amber stop codon (54, 55). Under UV illumination, pBpa can cross-link to residues on a binding protein within 2.4–4 Å, allowing for a clear readout of protein-protein interactions (56). Based on analysis of the human Ssu72-symplekin structure (28), several residues of Ssu72 with surface-exposed side chains were chosen for substitution. A cluster of negatively charged residues (positions 171–181) were chosen under the assumption that they may interact with positively charged residues of TFIIB₁c important to interaction (Fig. 3C). Residues around the interface with symplekin (28) were also chosen under the hypothesis that they may form a similar binding interface with TFIIB₁.

Ssu72 surface residues that were substituted for pBpa included Glu₁₂₇, Gln₁₂₈, Thr₁₃₀, Cys₁₃₁, Pro₁₃₃, Asp₁₆₉, Glu₁₇₃, Leu₁₈₈, Phe₁₉₃, and Tyr₁₉₄. Strong UV-dependent cross-linking to TFIIB₁ was observed for all of the pBpa substitutions except E₁₂₇pBpa, L₁₈₈pBpa, and Y₁₉₄pBpa (Fig. 4A). In agreement with the results from the pulldown assays, no cross-linking was observed with the K₁₈₉E/R₁₉₃E mutant of TFIIB₁, that loses the Ssu72-binding capacity (Fig. 4B). The photocross-linkable substitutions are clustered on an Ssu72 surface on the back of its substrate-binding site, outlining an elongated TFIIB₁ binding surface (Fig. 4C, green dashed outline).

Residues 128–133 of Ssu72 form a loop that is partially buried in the Ssu72-symplekin complex (28) (Fig. 4C and Fig. S5). Residue Asp₁₆₉ sticks out of a helix opposite to this loop and approaches residue Pro₁₃₃. Residue Glu₁₇₃ is further down in the same helix and both Asp₁₆₉ and Glu₁₇₃ are part of a negatively charged, surface-exposed patch that also includes residues Glu₁₇₁, Glu₁₇₄, and Glu₁₈₁ (Fig. 4C, green sticks). When any of these residues are mutated to pBpa, cross-linking to TFIIB₁c was observed (Fig. 4D). Residue Phe₁₉₃ was known to be critical for symplekin binding, and it now appears to contribute to the TFIIB₁-binding surface as well (Fig. 4C and Fig. S6) (28). Residues Glu₁₂₇ and Leu₁₈₈ are located close to but outside of the TFIIB₁-binding surface, and residue Tyr₁₉₄ is partially buried during Ssu72 folding (Fig. 4C, yellow sticks). These facts not only account for the negative photocross-linking results upon UV illumination for the corresponding pBpa substitutions but also support the remarkable specificity of photocross-linking offered by pBpa. Thus, three discrete and spatially proximate patches (i.e., regions centered on residues 128–133, residues

Figure 2. Sequence alignment and structural comparison of TFIIB₁c. A, sequence alignment of TFIIBc from Homo sapiens (Hs), S. cerevisiae (Sc), D. melanogaster (Dm), and T. brucei (Tb). Secondary structural domain organization is labeled based on the human TFIIBc structures determined in this report. Residues mutated in experiments are labeled by blue triangles. B, alignment of TFIIBc molecule D (color-coded according to Fig. 1A) with human TFIIBc from the TFIIB₁c-TBP-DNA structure (PDB code 1C9B, TFIIBc in gray schematic, DNA in multicolor schematic, and TBP omitted). Cyclin-like domain 1 of molecule D was aligned with that of TFIIBc from 1C9B. In the subsequent panels, alignment follows the same scheme, except that the second TFIIBc molecule was first aligned with TFIIBc from 1C9B to allow for modeling of the DNA. C, alignment of molecule D with TFIIBc from T. brucei (PDB code 3H4C, magenta), D, alignment of molecule D with human TFIIBc determined by NMR (PDB code 1TFB; cyan). The same TATA box DNA molecule from B is shown in C and D as well to guide the view. The purpose of the structural alignment in B, C, and D is to illustrate the conformational plasticity but not the actual new conformations of TFIIBc₁, E, electromobility shift assay with E4 promoter DNA, TBP₁c, and TFIIB variants.
Figure 3. Physical interaction of TFIIBc and Ssu72. A, ITC experiment of Ssu72 and TFIIBc. The best fit $K_d$ value (95% confidence interval) is 20.2 (16.8–24.4) $\mu$M. B, FLAG pulldown assay with TFIIB-FLAG variants and His$_6$-Ssu72, as visualized by Coomassie staining and Western blotting with a His$_6$ antibody. *, IgG light chain that leached from the resin upon elution with glycine. C, FLAG pulldown assay with TFIIBc-FLAG mutants and His$_6$-Ssu72. D, FLAG pulldown assay with TFIIBc-FLAG WT and mutants from various species (top labels) with His$_6$-Ssu72 from the corresponding species as visualized by Western blotting with a His$_6$ antibody. In the figure, h represents human, d is D. melanogaster, and y is S. cerevisiae. Lanes labeled with FLAG indicate controls without TFIIBc added.
169–181, and residue 193) of Ssu72 form the binding surface for TFIIB (Fig. 4C).

To support the conservation of Ssu72-TFIIB binding, we further showed that the pBpa substitutions of human Ssu72 within the three TFIIB binding patches cross-linked to yeast TFIIBc. The cross-linking was abolished when the K201E/K205E mutant of yeast TFIIBc was attempted (Fig. 4E), indicating that the binding mode between TFIIB and Ssu72 is highly conserved across species, although one of the three patches exhibits poor sequence conservation (Fig. 4F).

**TFIIB and symplekin bind to overlapped Ssu72 surfaces**

Ssu72 is known to form a stable complex with the N-terminal region (residues 30–360) of symplekin (referred to below as Symp-N) (28). In our pulldown assays, TFIIB was unable to interact with Ssu72 in the presence of Symp-N, indicating that the binding sites for TFIIB and Symp-N overlap on the Ssu72 surface (Fig. 5A). To further delineate the competitive binding of TFIIB and Symp-N to Ssu72, we utilized a competition photocross-linking assay. Ssu72 with the pBpa substitutions were prebound with an excess of either TFIIBc or Symp-N, incubated with the competitor protein, and UV-cross-linked. Symp-N strongly cross-linked to Ssu72 C131pBpa (Fig. 5B, lane 2). When cross-linked in competition with TFIIBc, only the cross-linked Ssu72–Symp-N but not Ssu72-TFIIB complex was visible (Fig. 5B, lane 6), which indicates that Symp-N outcompetes TFIIB in Ssu72 binding, probably due to a higher binding affinity (Fig. 5A). Notably, Ssu72 F193pBpa cross-linked to Symp-N in a much weaker manner compared with Ssu72 C131pBpa (Fig. 5B, compare lanes 2 and 8), possibly because the former disturbs Symp-N binding (28). The weaker interaction between Ssu72 F193pBpa and Symp-N allowed TFIIBc to cross-link to Ssu72 to a similar degree in both the presence and absence of an excess of Symp-N (Fig. 5B, compare lanes 10 and 12). In a similar experiment, the K185A mutant of Symp-N that weakens...
Ssu72 binding also allowed photocross-linking between Ssu72 and TFIIBc (Fig. 5C, lane 6) (28). Overall, our results indicate that Ssu72 contains an overlapping yet not identical binding surface for TFIIBc and symplekin. Although binding to an overlapped Ssu72 surface, unlike symplekin, TFIIB did not enhance the intrinsic phosphatase activity of Ssu72. In addition, TFIIB did not dramatically influence the stimulation of Ssu72 catalysis by symplekin, even in large excess (Fig. S7), confirming that TFIIB cannot outcompete symplekin for Ssu72 binding, at least at the concentration used in our assays.

Ssu72 binding also allowed photocross-linking between Ssu72 and TFIIBc (Fig. 5C, lane 6) (28). Overall, our results indicate that Ssu72 contains an overlapping yet not identical binding surface for TFIIB and symplekin.

Although binding to an overlapped Ssu72 surface, unlike symplekin, TFIIB did not enhance the intrinsic phosphatase activity of Ssu72. In addition, TFIIB did not dramatically influence the stimulation of Ssu72 catalysis by symplekin, even in large excess (Fig. S7), confirming that TFIIB cannot outcompete symplekin for Ssu72 binding, at least at the concentration used in our assays.

Ssu72-binding site on TFIIBc overlaps with the VP16AD- and TFIIBn-binding site

Two additional interactions mediated by TFIIBc were noted to influence transcription initiation, including VP16AD acidic activator binding and TFIIBn binding, in the context of a TFIIB intramolecular complex. Because Ssu72 binds to a similar region of TFIIBc, we further investigated such functional cross-talk with the photocross-linking competition assays.

We first established that TFIIBc could be photocross-linked to an N-terminal SUMO fusion of TFIIBn (residues 1–106) with a pBpa substitution at residue Trp52 located in the finger/reader region (Fig. 6A, compare lanes 1 and 2). In the competition assays, we next observed that unlabeled Ssu72 decreased TFIIBc-TFIIBn cross-linking compared with the GST control, probably due to competitive binding of Ssu72 and TFIIBn to TFIIBc (Fig. 6A, compare lanes 3 and 4 with lanes 6 and 7). In a reciprocal competition cross-linking experiment, GST-tagged TFIIBn could barely disrupt TFIIBc-Ssu72 cross-linking compared with the GST control, probably due to competitive binding of Ssu72 and TFIIBn to TFIIBc (Fig. 6A, compare lanes 3 and 4 with lanes 6 and 7). In a reciprocal competition cross-linking experiment, GST-tagged TFIIBn could barely disrupt TFIIBc-Ssu72 cross-linking compared with the GST control (Fig. 6B, compare lanes 4 and 6 with lane 12), indicating that TFIIBc displays a comparable or a lower binding affinity toward TFIIBc when compared with Ssu72. Conversely, GST-tagged VP16AD efficiently prevented cross-linking between Ssu72 and TFIIBc (Fig. 6B, compare lanes 8 and 10 with lanes 4 and 6). Accordingly, we conclude that Ssu72 may open up the inhibitory TFIIBc-TFIIBn intramolecular complex to facilitate pol II transcription with a similar mechanism used by VP16AD (16–21).

A computational model predicts the interactions of TFIIBc and Ssu72

To further elucidate how the residues identified by our experiments interact on the TFIIBc-Ssu72 interface, we devised a computational approach combining molecular dynamics (MD) simulations and knowledge-based protein-protein docking. Our approach successfully generated an interface model that agrees with all experimental observations. This model suggests a mechanism whereby Ssu72 activates TFIIB for reinitiation. Moreover, the model includes additional, hypothetical TFIIBc and Ssu72 interactions that may represent secondary binding interfaces.
We first performed MD simulations of TFII\textsubscript{Bc} and Ssu72 separately. We observed that whereas Ssu72 is a rigid protein, TFII\textsubscript{Bc} is highly flexible (Fig. 8), in line with our structural observations analyzed above (Fig. 1B). Therefore, we extracted TFII\textsubscript{Bc} structures from the MD simulations that may potentially bind to Ssu72 (example in Fig. S9) and then performed protein-protein docking to generate the binding models. Models with low HADDOCK score (less negative) and experimental violation were disregarded, resulting in a final high-quality model that best fits the experimental results.

In this model, the cross-linked residues on Ssu72 are very close to TFII\textsubscript{Bc1}. A 60-ns MD simulation of this model revealed that the average D2 distance of all cross-linked residues is well below the threshold value (Fig. 7A), indicating that this binding interface is rather stable. Consistent with the mutagenesis experiment on Lys\textsubscript{200}, Arg\textsubscript{189}, and Arg\textsubscript{193} of TFII\textsubscript{Bc1}, four salt bridges are found on the interface: Lys\textsubscript{200}, Glu\textsubscript{121}, Lys\textsubscript{200}, Asp\textsuperscript{117}, Lys\textsubscript{189}, Glu\textsubscript{167}, Arg\textsubscript{193}, Glu\textsubscript{169}, with the TFII\textsubscript{Bc} residues underlined (Fig. 7B, bottom). These salt bridges are mostly mediated by TFII\textsubscript{Bc1}, consistent with its dominant role in Ssu72 binding as shown by our pulldown assay (Fig. 3B). Additional interactions include the salt bridges Lys\textsuperscript{152}, Glu\textsubscript{127}, Lys\textsuperscript{188}, Asp\textsuperscript{172}, Arg\textsuperscript{206}, Glu\textsubscript{181}, Lys\textsuperscript{272}, Glu\textsubscript{174}, and Arg\textsuperscript{206}, Glu\textsubscript{178}. The last three interactions involve residues of TFII\textsubscript{Bc2}. Probably, these interactions are weaker compared with those of TFII\textsubscript{Bc1} because TFII\textsubscript{Bc2} is not sufficient to pull down Ssu72 (Fig. 3B), and mutation of TFII\textsubscript{Bc1} abolishes interaction with Ssu72 (Figs. 3C and 4B).

**Discussion**

Whereas the role of TFII\textsubscript{B} in transcription initiation is well established, its role in other processes such as gene looping is not clear. The current study sheds light on possible new roles of TFII\textsubscript{B} in this process by revealing a structure of human TFII\textsubscript{B} that adopted two new conformations. We also examined interactions between TFII\textsubscript{Bc} and Ssu72 and other factors known to interact with these two proteins, such as VP16\textsubscript{AD} and simplekin, respectively.

The structure of human TFII\textsubscript{B} displays the dynamic nature of the protein and indicates that conformational plasticity of TFII\textsubscript{B} may underlie a mechanism for functional control. Apo-TFII\textsubscript{B} is in an autoinhibited, “closed” conformation where TFII\textsubscript{B} blocks the TBP and the TATA DNA-binding region of TFII\textsubscript{B}, (16–21) (Fig. 8, “Closed” Autoinhibited state). VP16\textsubscript{AD} displaces TFII\textsubscript{B}, by interacting with an overlapping surface on TFII\textsubscript{B}, promoting the structural conversion of TFII\textsubscript{B} to an “open” conformation (20) that is then stabilized by TBP, TATA DNA, and other components of the transcription machinery during initiation. Both conformations of TFII\textsubscript{B}, revealed by our study are essentially transcriptionally inactive because the two cyclin-like domains and C-linker are not aligned to interact with TBP and the TATA DNA due to intrinsic conformational flexibility, although structural elements for these functional interactions are all surface-exposed. The structural transition from an inactive conformation to an active conformation may be subjected to additional cellular regulation.

We showed that Ssu72 predominantly interacts with the TATA-binding region of cyclin-like domain 1 of TFII\textsubscript{B} (TFII\textsubscript{B1}) and that residues Arg\textsubscript{189} and Arg\textsubscript{193} are critical for binding (Fig. 3C). This region is also a binding site of VP16\textsubscript{AD}, suggesting that Ssu72 could function to transition TFII\textsubscript{B} from a closed to an open conformation in a similar manner (Fig. 8). Indeed, similar to VP16\textsubscript{AD}, Ssu72 can easily outcompete the binding of TFII\textsubscript{B}, to TFII\textsubscript{B1}, indicating that it can at least relieve TFII\textsubscript{B} from autoinhibition (Fig. 8). In addition, our computational study provides a possibility that Ssu72 may further form a stable complex with the active form of TFII\textsubscript{B}, (Fig. 8, “Open” Active state). The open active state of TFII\textsubscript{B}, maintained by Ssu72 may facilitate formation of the TFII\textsubscript{B}-TBP-TATA ternary complex through a “hand-off” model. This process may occur during transcription reinitiation in the context of a gene loop, through
TFIIB-Ssu72 interaction and promoter-terminator cross-talk

Figure 8. A schematic model that illustrates a possible role of the TFIIBc-Ssu72 complex in pol II transcription reinitiation. A "hand-off" model may explain the conceivable role of Ssu72 in pol II reinitiation by modulating TFIIB conformation, in a manner similar to the archetypal acidic activator VP16. Arrows, relative orientations of the cyclin domains of TFIIB, Ssu72 may further maintain the relative orientation of the two cyclin-like domains in an active conformation.

a reinitiation scaffold complex on a TATA promoter that contains TBP but not TFIIB (14). The hand-off model in transcriptional regulation has also been proposed previously for transcriptional activators (57).

Interaction between TFIIB and Ssu72 is well conserved from yeast to flies and humans (Fig. 3D), suggesting a likely functional conservation. TFIIB binds to Ssu72 over an extended surface that includes a negatively charged helix, a loop, and a region near the C terminus of the protein. The latter two TFIIB-binding regions overlap with the binding surface of symplekin, which precludes TFIIB from binding to Ssu72. The Ssu72-symplekin interaction is strong, but it can be disrupted by a point mutation of Phe193 of Ssu72 that does not impede TFIIB binding. Interestingly, residue Phe193 is close to several conserved serine, threonine, and tyrosine residues. Although we have not found a known phosphorylation motif in this area, it is possible that phosphorylation or other posttranslational modification of the C terminus of Ssu72 could be used to weaken the interaction with symplekin such that TFIIB binding is allowed.

The fact that TFIIB and Ssu72 physically interact yet are involved in disparate steps of transcription, initiation and termination, respectively, is intriguing. In the context of promoter-terminator gene looping, one can imagine a mechanism in which Ssu72 dephosphorylates the CTD of pol II during transcription termination, recruits TFIIB, and keeps it in an open conformation to allow for reassembly of PIC on the promoter for subsequent rounds of transcription. This hypothetical mechanism would explain why both TFIIB and Ssu72 are required for gene looping, and it can be further tested with separation-of-function mutants in future studies.

Materials and methods

Protein expression and purification

All proteins were produced recombinantly in Escherichia coli. DNA sequences for each protein were cloned into a pET28a vector (containing an N-terminal His6 tag), a SUMO-pET28a vector (containing a His6 SUMO tag), or pGEX4T1 (containing an N-terminal GST tag). All mutations were introduced by site-directed mutagenesis of the parental plasmid. Expression plasmids were transformed into BL21(DE3)-Rosetta II cells and cultured in LB medium supplemented with 50 µg/ml chloramphenicol and either 50 µg/ml kanamycin or 100 µg/ml ampicillin and cultured at 37 °C until reaching an A600 of 0.6, at which point a final concentration of 0.5 mM isopropyl 1-thio-β-D-galactopyranoside was added, and cells were incubated overnight at 16 °C.

All TFIIB proteins, with the exception of GST-TFIIBn, were purified by nickel chromatography and contained a His6-SUMO tag and C-terminal FLAG tag. Cell pellets were resuspended in lysis buffer (20 mM HEPES, pH 8.0, 500 mM NaCl, 10 mM imidazole, and 2 mM β-mercaptoethanol) and lysed by sonication. Following centrifugation, the cleared lysate was applied to HisPure nickel resin (Thermo Scientific), the resin was washed with 20 column volumes of lysis buffer, and proteins were eluted with elution buffer (20 mM HEPES, pH 8.0, 300 mM NaCl, 10 mM imidazole, and 2 mM β-mercaptoethanol). Proteins that contained the N terminus of TFIIBc (i.e. TFIIBcFL and TFIIBc) additionally contained 10 µM zinc acetate in all buffers. The His6-SUMO tag was removed by Ulp1 cleavage overnight during dialysis into low imidazole buffer (20 mM HEPES, pH 8.0, 300 mM NaCl, 10 mM imidazole, and 2 mM β-mercaptoethanol). Proteins were then reapplied to a new nickel column to remove Ulp1 and contaminants, and the flow-through was collected. Proteins were dialyzed into a low-salt buffer (HEPES, pH 7.0 or 8.0, 100 mM NaCl, 100 mM imidazole, and 2 mM β-mercaptoethanol) and then applied to either a HiTrap heparin column (human TFIIBcFL wild type and mutants) or a MonoS 5/50 column (human TFIIBc, TFIIB cyclin-like 1, and TBPc) and eluted with a gradient of NaCl (both columns were from GE Healthcare). The crystallized human TFIIBc that lacked a C-terminal FLAG tag was concentrated following the MonoS 5/50 column and then loaded onto a Superdex 200 10/300 GL column (GE Healthcare) in 20 mM HEPES, pH 7.0 or 8.0, 100 mM NaCl, 10 mM imidazole, and 2 mM DTT. Because TFIIB cyclin-like 2 and TFIIBc mutants (human K189E/K200E, R185E/R193E, R185E/K200E, K189E/R193E; yeast H197E/N212E and K201E/K205E) failed to bind to heparin, they were passed through a MonoQ 5/50 column (GE Healthcare) to remove contaminants, and the flow-through was collected.

All Ssu72 and N-terminal symplekin (residues 30–360) proteins contained an N-terminal His6 tag and were purified by nickel chromatography followed by dialysis into a low-salt buffer.
buffer (HEPES, pH 8.0, 100 mM NaCl, and 2 mM β-mercaptoethanol) and a final MonoQ purification step. To assemble the Ssu72-symplekin complex, Ssu72 was mixed with symplekin in a 2:1 molar ratio, incubated on ice for 1 h, and then separated on a Superdex 200 10/300 GL column in 20 mM Tris, pH 8.5, 200 mM NaCl, and 10 mM DTT. Fractions containing the complex were pooled, concentrated, and stored at −80 °C until use.

Cells containing GST fusion proteins (TFIIBn and VP16AD) were lysed by sonication in GST lysis buffer (50 mM Tris, pH 8.0, 150 mM NaCl, and 5 mM DTT), and clarified lysate was bound to glutathione-agarose (Thermo Scientific). The resin was washed with 20 column volumes of GST lysis buffer and eluted with lysis buffer containing 20 mM reduced glutathione. The proteins were then applied to a MonoQ 5/50 column and eluted over a gradient of NaCl. Following purification, all proteins were concentrated, flash-frozen in liquid nitrogen, and stored at −80 °C until needed.

For Ssu72 proteins containing pBpA mutations, the Ssu72 plasmid containing the amber mutation of interest was transformed into BL21* (DE3) cells containing the pEVOL-pBpA plasmid that encodes for an evolved aminoacyl tRNA synthetase and tRNA that incorporate pBpA into the amber suppressor site (55). Cells were cultured at 37 °C in 100 ml of 2XYT medium supplemented with 50 μg/ml chloramphenicol and 50 μg/ml kanamycin. At an A₆₀₀ of 0.3, a final concentration of 0.2% arabinose was added to the culture. At an A₆₀₀ of 0.6, a final concentration of 1 mM isopropyl-β-D-galactopyranoside and 1 mM pBpA were added to the culture, and the culture was then allowed to incubate overnight at 30 °C. Proteins were purified in a manner similar to that described for Ssu72 by nickel and Q Sepharose Fast Flow (GE Healthcare) chromatography. SUMO-TFIIBn-W52pBpA, containing an N-terminal His₆ tag, was expressed in the same manner as Ssu72-pBpA proteins. It was purified by nickel chromatography in a manner similar to other His-tagged proteins, with the addition of 10 μM zinc acetate in all purification buffers.

**Protein crystallization, data collection, and structure solution and refinement**

Human TFIIBc was crystallized at a protein concentration of 7 mg/ml in the hanging drop format, over a well solution containing 100 mM sodium acetate trihydrate, pH 4.5, and 2 M ammonium sulfate. Sucrose at a concentration of 3% (w/v) was used as an additive to improve crystal quality. Crystals were cryoprotected in mother solution supplemented with a final concentration of 30% glycerol and plunged into liquid nitrogen. Diffraction data were collected at beamline 19ID of the Advanced Photon Source at Argonne National Laboratory. Data were collected over a total of 180° with oscillation of 1° and an exposure time of 2 s/degree. Diffraction data were processed with HKL2000 (58), and TFIIBc was found to be in space group P2₁2₁2₁ (R₁, for the lowest-resolution shell (50.00–9.21) is 6.0%). The structure of TFIIBc was solved by using the separate cyclin-like 1 and cyclin-like 2 domains from the TFIIBc·TBPc-DNA structure (8) using MOLREP (59). Manual building of connecting loops was completed with Coot (60), followed by refinement with PHENIX (61). Figures were created with PyMOL (PyMOL Molecular Graphics System, Version 1.8.6.2., Schrödinger, LLC, New York).

**Isothermal titration calorimetry (ITC)**

ITC was performed on a VP-ITC instrument (GE Healthcare) at 25 °C in 25 mM HEPES, pH 8.0, 150 mM NaCl, 0.5 mM tris(2-carboxyethyl)phosphine. Ssu72 (1 mM) in the syringe was titrated into TFIIBc (50 μM) in the cell over 30 injections, and data were background-corrected using a control titration where buffer was injected into Ssu72. Data were analyzed with NITPIC and Sedphat software (62). We noticed slight precipitation after the titration, which might cause an inaccuracy of the protein concentrations and the estimate of K₆₄.

**Pulldown assays**

For FLAG pulldown assays, an amount of 75 μg of TFIIBc variant with a C-terminal FLAG tag was mixed with 150 μg of N-terminal tagged Ssu72 in binding buffer (50 mM Tris, pH 8.0, 150 mM NaCl, and 0.1% Tween) and bound to 50 μl of anti-FLAG-agarose (Sigma-Aldrich) at room temperature for 1 h with rotation. For the assay conducted in Fig. 5A, TFIIBc·FLAG was prebound to the resin, followed by extensive washing before the addition of other proteins. The resin was then centrifuged for 3 min at 500 rpm at 4 °C, the supernatant was removed, and the resin was washed three times with 1 ml of binding buffer. Proteins were then eluted with 100 mM glycine, pH 3.5, neutralized with 1 μl of 1 M Tris, pH 8.0, and run on SDS-PAGE. Protein interactions were analyzed by staining with Coomassie Brilliant Blue and/or Western blotting with an anti-His₆ antibody (Abcam).

**Site-specific UV-cross-linking analysis**

Amounts of 20 μg of TFIIBc·FLAG and 40 μg of His₆-Ssu72-pBpA mutant were mixed in assay buffer (20 mM Tris, pH 8.0, 150 mM NaCl, 1 mM DTT) in a final volume of 50 μl and incubated on ice for 20 min. An amount of 25 μl was removed from each sample as a no UV control. Samples were then UV-cross-linked on ice for four cycles at energy level 9999 in a Spectrolinker cross-linker (Krackeler Scientific), with 5 min of incubation on ice between cycles. 6.25 μl of 6× SDS dye was added, samples were boiled, and 12.5 μl was run on SDS-PAGE. Cross-linking was analyzed by separate Western blots with an anti-FLAG antibody and an anti-His₆ antibody (Abcam). For competition cross-linking with TFIIBn and VP16AD (Fig. 4), TFIIBc was preincubated with the molar excess of competitor, as indicated in the figure, for 15 min on ice. Samples were then processed as described above. For competition cross-linking between TFIIBc, Ssu72, and symplekin (Fig. 5), 40 μg of Ssu72 was preincubated with 30 μg of symplekin for 20 min on ice. An amount of 60 μg of TFIIBc was then added, and samples were then processed as described above.

**EMSA**

The template strand DNA oligonucleotide containing the E4 promoter sequence (TTTTTAGTCCTATATATACTCGC TCT) was annealed to the reverse complement strand containing a 3’-end Cy5 label (Sigma-Aldrich). TFIIBFl, TBPc, and DNA were mixed at equal molar concentrations (300 nM) in
assay buffer (20 mM Tris, pH 8.0, 60 mM KCl, 5 mM MgCl₂, 4% glycerol, and 10 μg bovine serum albumin) in a final volume of 20 μl and incubated at 30 °C for 40 min. Samples were then subjected to electrophoresis on a 6%, 37.5:1 (acylamide/bisacrylamide) gel at 4 °C in 25 mM Tris, pH 8.3, 190 mM glycine running buffer. Fluorescence was visualized with a Pharos imager (Bio-Rad).

Molecular dynamics simulation of TFIIB₆

The two available crystal structures of TFIIB₆ were used as initial structures for MD simulation. The first one is TFIIB₆ in complex with TBP and DNA promoter (PDB code 1C9B) (8). The other one is the crystal structure of apo-TFIIB₆ (this report).

For TFIIB₆ from PDB entry 1C9B, only the TFIIB₆ structure was used. For apo-TFIIB₆ crystal structure, only molecule D was used. Both structures were solvated in a dodecahedron box with 17,779 TIP3P (63) water molecules. The system also contains 51 Na⁺ and 61 Cl⁻ ions to create a neutral environment with 0.15 mol/liter salt concentration. The GROMACS version 4.5 (64) simulation package and AMBER 99SB-ILDN (65) force field were used to carry out the succeeding steps. Long-range electrostatic interactions were treated by the particle mesh–Ewald method (66). Electrostatic and van der Waals short-range interactions were calculated using a cutoff of 10 Å.

Energy minimization was performed using the steepest-descent method, followed by NVT equilibration for 100 ps with position restraints on the heavy atom of TFIIB₆ at 300 K using V-rescale thermostat (67), which is used in all MD simulations in succeeding steps. At this step, the positions of heavy atoms were restrained with a force constant of 1000 kJ mol⁻¹ nm⁻², and all bonds in protein molecules and water were constrained using the LINCS algorithm (68). NPT equilibration was then carried out using the Parrinello–Rahman barostat (69), to a reference pressure of 1 bar until the average pressure reached around 1 bar.

Eight production MD simulations starting from TFIIB₆ structure from PDB entry 1C9B were performed in the NVT ensemble with the length of each simulation ranging from 50 to 300 ns, with an accumulated time of 1600 ns for TFIIB₆ structure from PDB entry 1C9B. Five production MD simulations starting from apo-TFIIB₆ were performed in the NVT ensemble with the length of each simulation ranging from 100 to 400 ns, giving an accumulated time of 1300 ns. In the succeeding MD simulations, we used the same simulation package, force field, non-bonded interaction calculation method, barostat, thermostat, and bond constraint.

TFIIB₆ is an inherently flexible protein, as shown by the large root mean square fluctuation values for most of the backbone atoms (Fig. S8A). Because we obtained many TFIIB₆ structures other than the crystal structures, we selected the starting models based on the following criteria. In TFIIB₆, residues 189, 193, and 200 are likely to interact with Ssu72 (Fig. 3C). Given that residue Ala281 in TFIIB₆ was shown previously to bind VP16AD and that Ssu72 and VP16AD bind to an overlapping surface of TFIIB₆ (Fig. 6B), we assumed that residues 281–284 of TFIIB₆ are located on the binding interface of the TFIIB₆-Ssu72 complex. Due to the flexibility of TFIIB₆, these two clusters of interacting residues on TFIIB₁₆ or TFIIB₂₆ can be on the same plane (coplanar) (Fig. S9A) or not (Fig. S9B). To choose TFIIB₆ structures, clustering for simulations from each crystal structure was performed by K-center clustering (70) to 50 states. Eighteen structures were chosen based on 1) the coplanarity of the interacting residues that mimics the active conformation of TFIIB₆ when it binds to TBP and the TATA DNA and 2) the distance between the center of mass of TFIIB₁₆ and TFIIB₂₆ being less than 40 Å. The crystal structure of TFIIB₆ from PDB entry 1C9B was also used for protein-protein docking, which resulted in a total of 19 starting structural models of TFIIB₆.

Molecular dynamics simulation of Ssu72

The starting structure for MD simulation was taken from PDB entry 3Q2Q (28). It was solvated in a dodecahedron box with 11,321 TIP3P water molecules. The system also contains 44 Na⁺ ions and 34 Cl⁻ ions to neutralize the system and creates a 0.15 mol/liter salt environment. Energy minimization was performed using the steepest-descent method, followed by NVT equilibration for 100 ps with position restraints (force constant = 1000 kJ mol⁻¹ nm⁻²) on the heavy atom of Ssu72 at T = 300 K with coupling constant = 0.1 ps. NPT equilibration was then carried out to a reference pressure of 1 bar with coupling constant = 0.2 ps until the average pressure reached around 1 bar. Six production NVT MD simulations with different random initial velocities were performed for 50 ns each, giving an accumulated time of 300 ns. Because Ssu72 is a rigid protein (as shown by the small root mean square fluctuation value of all of the backbone atoms; Fig. S8B), we used the crystal structure to do protein-protein docking.

Protein-protein docking using HADDOCK

The 19 selected structures from MD simulation of TFIIB₆ were used as input for protein-protein docking, along with the crystal structure of Ssu72 (PDB code 3Q2Q). The cross-linked residues on Ssu72 were used as active residues input into the HADDOCK web server (71) (i.e. Pro¹³³, Phe¹⁹⁵, Gln¹₂⁸, Thr¹₃₀, Cys¹₃¹, Glu¹⁶⁹, Asp¹₇₃, Glu¹⁷⁴, Glu¹⁸₁). The interacting residues of TFIIB₆, which were either determined by mutagenesis or prior knowledge of interaction with VP16AD, were also used as active residues (i.e. Lys¹⁸⁹, Arg¹₉₃, Lys²₀₀, Ala²₈₁, Asp²₈₂, Val²₈₃, Thr²₈₄). The intermolecular and intramolecular interactions were determined by CNS 1.3 (72, 73) along with PARALLET-HDG5.4 (74, 75) as force field for protein. For each docking simulation, the two input molecules were separated for 25 Å and rotated randomly around their centers of mass, followed by rigid body minimization and docking. Two hundred best structures from rigid body docking were chosen for semiflexible refinement, based on the weighted sum of several of energy terms (i.e. electrostatic, van der Waals, desolvation, ambiguous interaction restraints, and total buried surface area). Of 1000 structures generated by rigid body docking, 200 structures were used for final solvated refinement and clustering. Clustering was done based on fraction of native contacts using 0.75 similarity as the threshold. For each cluster of every docking simulation, we picked the four models with the lowest HADDOCK scores for the following analysis.
Selection of docking model

HADDOCK score and experimental results were considered to select the best model. The HADDOCK score used in the preceding steps excluded the $E_{	ext{AIR}}$ part of the score. For consideration of the cross-linking experiment, we defined a distance that measured the likelihood of cross-linking reaction (D2 distance, shown in Fig. S10). The D2 distance is defined as the distance between the C-atom of cross-linked residues on Ssu72 and the C-atom of the nearest residue on TFIIß that can potentially form a cross-linking bond. The cross-linking reaction only occurs between bPpa and a non-quaternary carbon atom (i.e. a carbon atom that binds to at least one hydrogen atom). The threshold for the D2 distance was determined to be 15 Å. Any cross-linked residue that has a D2 distance violating the threshold is considered as 1 violation count. The contacts between residues 189, 193, 200, 281, 282, 283, and 284 on TFIIß and Ssu72 residues were determined by using the g_contacts module in GROMACS (76), with a threshold of 3 Å. If any of these residues was not in contact, the violation count was 2 for residues 189, 193, and 200, giving priority to the mutagenesis experiment, and 1 for residues 281, 282, 283, and 284. The total experimental violation count was the sum of all of the violations. The plot of HADDOCK score versus experimental violation count is shown in Fig. S11A. All of the models with violation less than 1 were analyzed further. According to the pulldown data, the interaction between TFIIß$_{c2}$ and Ssu72 is stronger than that between TFIIß$_{c1}$ and Ssu72. Considering this result, we calculated the ratio of contacts between TFIIß$_{c1}$ and Ssu72 and contacts between TFIIß$_{c2}$ and Ssu72. The plot in Fig. S11B clearly shows one model having the largest ratio and a low HADDOCK score (represented as a star), hence fitting best to the pulldown assay experiment (Table S1).

MD simulation of TFIIß-Ssu72 complex

We solvated the TFIIß$_{c2}$-Ssu72 model in a dodecahedron box with a 1-nm distance between the closest point of the protein and the walls of the box. The protein was then solvated in TIP3P water molecules and an equal number of sodium and chloride ions to create a 0.15 mol/liter salt concentration.

The prepared system was then energy-minimized by the steepest-descent method, followed by NVT equilibration with $T = 300$ K and coupling constant $= 0.1$ ps as well as position restraints on heavy atoms of the protein (force constant $= 1000$ kJ mol$^{-1}$ nm$^{-1}$) for 100 ps. Next, NPT equilibration was performed to reference pressure $= 1$ bar and coupling constant $= 0.2$ ps, until the average pressure reached 1 bar. The frame from MD simulation that has a volume equal to the average volume was chosen for a 200-ns NVT MD simulation. The last frame of this short MD simulation was the starting ensemble for the production MD simulations. Finally, MD simulation in an NVT environment was performed for 60 ns, with the first 10 ns excluded in analysis.

References

1. Bushnell, D. A., Westover, K. D., Davis, R. E., and Kornberg, R. D. (2004) Structural basis of transcription: an RNA polymerase II-TFIIB cocrystal at 4.5 Angstroms. Science 303, 983–988 CrossRef Medline
2. Kostrewa, D., Zeller, M. E., Armache, K. J., Seizl, M., Leike, K., Thomm, M., and Cramer, P. (2009) RNA polymerase II-TFIIF structure and mechanism of transcription initiation. Nature 462, 323–330 CrossRef Medline
3. Liu, X., Bushnell, D. A., Wang, D., Calero, G., and Kornberg, R. D. (2010) Structure of an RNA polymerase II-TFIIIF complex and the transcription initiation mechanism. Science 327, 206–209 CrossRef Medline
4. Sainsbury, S., Niessler, J., and Cramer, P. (2013) Structure and function of the initially transcribing RNA polymerase II-TFIIIF complex. Nature 493, 437–440 CrossRef Medline
5. Pinto, J., Wu, W. H., Na, J. G., and Hampsey, M. (1994) Characterization of sua7 mutations defines a domain of TFIIH involved in transcription start site selection in yeast. J. Biol. Chem. 269, 30569–30573 CrossRef Medline
6. Čabart, P., Úvári, A. P., and M., and Luse, D. S. (2011) Transcription factor TFIIH is not required for initiation by RNA polymerase II, but it is essential to stabilize transcription factor TFIIH in early elongation complexes. Proc. Natl. Acad. Sci. U.S.A. 108, 15786–15791 CrossRef Medline
7. Nikolov, D. B., Chen, H., Halay, E. D., Usheva, A. A., Hisatake, K., Lee, D. K., Roeder, R. G., and Burley, S. K. (1995) Crystal structure of a TFIIH-TBP-TATA-element ternary complex. Nature 377, 119–128 CrossRef Medline
8. Tsai, F. T., and Sigler, P. B. (2000) Structural basis of preinitiation complex assembly on human pol II promoters. EMBO J. 19, 25–36 CrossRef Medline
9. Lagrange, T., Kapanidis, A. N., Tang, H., Reinberg, D., and Ebright, R. H. (1998) New core promoter element in RNA polymerase II-dependent transcription: sequence-specific DNA binding by transcription factor IIB. Genes Dev. 12, 34–44 CrossRef Medline
10. Deng, W., and Roberts, S. G. (2005) A core promoter element downstream of the TATA box that is recognized by TFIIH. Genes Dev. 19, 2418–2423 CrossRef Medline
11. Cox, J. M., Hayward, M. M., Sanchez, J. F., Gegnas, L. D., van der Zee, S., Dennis, J. H., Sigler, P. B., and Schepartz, A. (1997) Bidirectional binding of the TATA box binding protein to the TATA box. Proc. Natl. Acad. Sci. U.S.A. 94, 13475–13480 CrossRef Medline
12. Roberts, S. G., Choy, B., Walker, S. S., Lin, Y. S., and Green, M. R. (1995) A role for activator-mediated TFIIIB recruitment in diverse aspects of transcriptional regulation. Curr. Biol. 5, 508–516 CrossRef Medline

Acknowledgments—We thank Shih-Chia Tso, Thomas Scheuermann, and Chad Brautigam for assistance with ITC experiments and Diana Tomchick and Zhe Chen for help with X-ray data collection. We thank Haydn Ball of the LIT Southwestern Protein Technology Core for peptide synthesis. We thank Qiong Wu and Jose Rizo-Rey for preliminary NMR analysis. We thank Michael Carey (Department of Biological Chemistry, UCLA, Los Angeles, CA) for generously providing the VP16AD expression plasmid. We thank Peter Schulze (Scripps Research Institute) for generously providing the pEVL-pBPα plasmid. This research used resources of the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under contract DE-AC02-06CH11357. The Advanced Light Source is supported by the Director, Office of Science, Office of Basic Energy Sciences, of the DOE under contract DE-AC02-05CH11231. Use of the Stanford Synchrotron Radiation Lightsource, SLAC National Accelerator Laboratory, is supported by the DOE, Office of Science, Office of Basic Energy Sciences under contract DE-AC02-76SF00515. The SSRL Structural Molecular Biology Program is supported by the DOE Office of Biological and Environmental Research and by NIGMS, National Institutes of Health, Grant P41GM103393.
31. Calvo, O., and Manley, J. L. (2003) Strange bedfellows: polyadenylation

25. Krishnamurthy, S., He, X., Reyes-Reyes, M., Moore, C., and Hampsey, M.

24. Dichtl, B., Blank, D., Ohnacker, M., Friedlein, A., Roeder, D., Langen, H.,

21. Roberts, S. G., Ha, I., Maldonado, E., Reinberg, D., and Green, M. R. (1993)

20. Hayashi, F., Ishima, R., Liu, D., Tong, K. I., Kim, S., Reinberg, D., Bagby, S.,

18. Zhang, D. Y., Dorsey, M. J., Voth, W. P., Carson, D. J., Zeng, X., Stillman,

16. Roberts, S. G., and Green, M. R. (1994) Activator-induced conformational

change of initiation, and roles of activators and coactivators. Genetics 189,

705–736 CrossRef Medline

17. C.L. (2003) Functional interactions between the transcription and mRNA

end processing machineries mediated by Ssu72 and Sub1. Nat. Genet. 30, 1479–1490 CrossRef Medline

16, 1557–1566 CrossRef Medline

13. Zawel, L., Kumar, K. P., and Reinberg, D. (1995) Recycling of the general

transcription factors during RNA polymerase II transcription. Genes Dev. 9,

1497–1490 CrossRef Medline

12. Palenchar, J. B., Bellofatto, V., and Wah, D. A. (2009) Structure of the

human CD68 gene is associated with changes in Pol II phosphorylation

and short-range intrachromosomal gene looping. Genomics 90, 407–415 CrossRef Medline

11. Martin, M., Cho, J., Cesare, A. J., Griffith, J. D., and Attardi, G. (2005)

Termination-factor-mediated DNA loop between termination and initiation

sites drives mitochondrial tRNA synthesis. Cell 123, 1227–1240 CrossRef Medline

10. Tan-Wong, S. M., French, J. D., Proudfoot, N. J., and Brown, M. A. (2008)

Dynamic interactions between the promoter and terminator regions of the mammalian BRCA1 gene. Proc. Natl. Acad. Sci. U.S.A. 105,

5160–5165 CrossRef Medline

9. O’Reilly, D., and Greaves, D. R. (2007) Cell-type-specific expression of the

human CD68 gene is associated with changes in Pol II phosphorylation

and short-range intrachromosomal gene looping. Genomics 90, 407–415 CrossRef Medline

8. Le May, N., Fradin, D., Ihls, I., Bougnères, P., and Egly, J. M. (2012) XPG

and XPF endonucleases trigger chromatin looping and DNA demethylation

for accurate expression of activated genes. Mol. Cell 47, 622–632 CrossRef Medline

7. Crevilla, P., Sonmez, C., Wu, Z., and Dean, C. (2013) A gene loop containing

the floral repressor FLC is disrupted in the early phase of vernalization.

EMBO J. 32, 140–148 Medline

6. Martin, M., Cho, J., Cesare, A. J., Griffith, J. D., and Attardi, G. (2005)

Termination-factor-mediated DNA loop between termination and initiation

sites drives mitochondrial tRNA synthesis. Cell 123, 1227–1240 CrossRef Medline

5. Yao, J., Ardehali, M. B., Fecko, C. J., Webb, W. W., and Lis, I. T. (2007)

Intracellular distribution and local dynamics of RNA polymerase II during

transcription activation. Mol. Cell 28, 978–990 CrossRef Medline

4. Tan-Wong, S. M., Vwijayatilleke, H. D., and Proudfoot, N. J. (2009) Gene

loops function to maintain transcriptional memory through interaction

with the nuclear pore complex. Genes Dev. 23, 6610–6624 CrossRef Medline

3. Tan-Wong, S. M., Zaug, J. B., Camblong, J., Xu, Z., Zhang, D. W., Mischo,

H. E., Ansari, A. Z., Lascombe, N. M., Steinmetz, L. M., and Proudfoot,

N. J. (2012) Gene loops enhance transcriptional directionality. Science 338,

671–675 CrossRef Medline

2. Singh, B. N., and Hampsey, M. (2007) A transcription-independent role for

TFII-B in gene looping. Mol. Cell 27, 806–816 CrossRef Medline

1. Ibrahim, B. S., Kanneganti, N., Rieckhof, G. E., Das, A., Laurents, D. V.,

Palenchar, J. B., Bellofatto, V., and Wah, D. A. (2009) Structure of the C-terminal domain of transcription factor IIb from Trypanosoma brucei.

Proc. Natl. Acad. Sci. U.S.A. 106, 13242–13247 CrossRef Medline

Bagby, S., Kim, S., Maldonado, E., Tong, K. I., Reinberg, D., and Ikura, M.

(1995) Solution structure of the C-terminal core domain of human TFII-B:

"TFIIB-Ssu72 interaction and promoter-terminator cross-talk"
TFIIB-Ssu72 interaction and promoter-terminator cross-talk

64. Pronk, S., Páll, S., Schulz, R., Larsson, P., Bjelkmar, P., Apostolov, R., Shirts, M. R., Smith, J. C., Kasson, P. M., van der Spoel, D., Hess, B., and Lindahl, E. (2013) GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. Bioinformatics 29, 845–854

65. Lindorff-Larsen, K., Piana, S., Palmo, K., Maragakis, P., Klepeis, J. L., Dror, R. O., and Shaw, D. E. (2010) Improved side-chain torsion potentials for the Amber ff99SB protein force field. Proteins 78, 1950–1958

66. 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

67. Bussi, G., Donadio, D., and Parrinello, M. (2007) Canonical sampling through velocity rescaling. J. Chem. Phys. 126, 014101

68. Hess, B., Bekker, H., Berendsen, H. J. C., Fraaije, J. G. E. M. (1997) LINCS: A linear constraint solver for molecular simulations. J. Comput. Chem. 18, 1463–1472

69. Parrinello, M., and Rahman, A. (1981) Polymorphic transitions in single crystals: a new molecular dynamics method. J. Appl. Phys. 52, 7182–7190

70. Gonzalez, T. F. (1985) Clustering to minimize the maximum intercluster distance. Theor. Comput. Sci. 38, 293–306

71. de Vries, S. J., van Dijk, M., Bonvin, A. M. J. J. (2010) The HADDOCK web server for data-driven biomolecular docking. Nat. Protoc. 5, 883–897

72. Brunger, A. T. (2007) Version 1.2 of the Crystallography and NMR system. Nat. Protoc. 2, 2728–2733

73. Brunger, A. T., Adams, P. D., Clore, G. M., Delano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, F., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR System: a new software suite for macromolecular structure determination. Acta Crystallogr. D Biol. Crystallogr. 54, 905–921

74. Linge, J. P., and Nilges, M. (1999) Influence of non-bonded parameters on the quality of NMR structures: a new force field for NMR structure calculation. J. Biomol. NMR 13, 51–59

75. Linge, J. P., Williams, M. A., Spronk, C. A. E. M., Bonvin, A. M. J. J., and Nilges, M. (2003) Refinement of protein structures in explicit solvent. Proteins 50, 496–506

76. Blau, C., and Grubmüller, H. (2013) G-contacts: fast contact search in bio-molecular ensemble data. Comput. Phys. Commun. 184, 2856–2859