



## Transcription Starts Here: Structural Models of a “Minimal” Preinitiation Complex

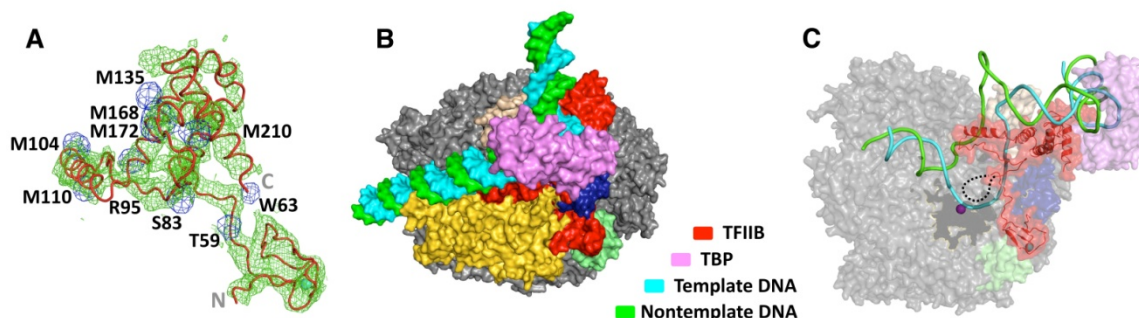
RNA polymerase II (pol II) plays a central role in the regulation of gene expression. Pol II is the enzyme responsible for synthesizing all the messenger RNA (mRNA) and most of the small nuclear RNA (snRNA) in eukaryotes. One of the key questions for transcription is how pol II decides where to start on the genomic DNA to specifically and precisely turn on a gene. This is achieved during transcription initiation by concerted actions of the core enzyme pol II and a myriad of transcription factors including five general transcription factors, known as TFIIB, -D, -E, -F, -H, which together form a giant transcription preinitiation complex on a promoter prior to transcription.

One of the most prominent core promoter DNA elements is the TATA box, usually directing transcription of tissue-specific genes. TATA-box binding protein (TBP), a key component of TFIID, recognizes the TATA DNA sequence. Based on the previous crystallographic studies, the TATA box DNA is bent by nearly 90 degree through the binding of TBP (1). This striking structural feature is thought to serve as a physical landmark for the location of active genes on the genome. In addition, the location of the TATA box at least in part determines the transcription start site (TSS) in most eukaryotes, including humans. The distance between the TATA box and the TSS is conserved at around 30 base pairs. TBP does not contact pol II directly and the TATA-containing promoter must be directed to the core enzyme through another essential transcription factor TFIIB. TFIIB intimately associates with TBP, pol II and the promoter DNA bridging the gap between the TBP bound promoter and pol II. Moreover, TFIIB plays a role in determining the directionality of transcription by sequence-specific interaction with the DNA sequences termed B recognition elements (BRE) both upstream and downstream of the TATA box (2, 3). In an *in vitro* transcription assay on a negatively supercoiled DNA template pol II, TBP and TFIIB form a “minimal” preinitiation complex capable of mediating accurate transcription (4). Similarly, archaea utilizes a simpler version of eukaryotic transcription apparatus, where only archaeal homologs of TBP and TFIIB are required (5, 6).

With diffraction data collected at SSRL, the structure of a pol II – TFIIB complex was solved at a 3.8 Å resolution (7). Anomalous peaks originated from selenomethionine substitutions on TFIIB were employed to trace the amino acid register (**Fig. 1A**). In addition to the previously determined TFIIB N-terminal zinc ribbon domain, two additional TFIIB regions termed the “linker” and “core” were revealed in the current structure. The TFIIB core is located above the polymerase active site cleft and the linker snakes down from above the cleft toward the active center. A previous pol II-TFIIB crystal structure revealed a region termed “B-finger” between the zinc ribbon and the linker, which reaches into the active site (8), however, the current crystal structure obtained under different solution conditions lacks ordered density for the B-finger. The current structure, however, is complementary to the previous structure, adding two new domains of TFIIB including the core domain which interacts directly with TBP. The current structure, taken together with others previously obtained, dispels longstanding mysteries of the transcription initiation process.

Specifically, two major advances were achieved that correlate well with previous structural and biochemical studies. First, a structural model of a closed promoter complex of the minimal preinitiation complex was constructed revealing how the TATA box directs the genome DNA to the pol II surface (**Fig. 1B**). The double-stranded genomic DNA downstream of the TATA box and harboring the TSS was placed above the active site cleft of pol II. Second, a structural model of an open promoter complex of the minimal preinitiation complex was constructed accounting for the selection of the template strand in

a transcription bubble and the conserved spacing between the TSS and the upstream TATA box (**Fig. 1C**). TFIIB linker and core are directly involved in the formation of the template strand tunnel, which leads the template strand to the pol II active site, buried some 30 Å inside of the pol II active site cleft. In addition, the overall structural organization of TFIIB on the pol II surface characterized in this and previous studies couples the length of the RNA transcript to the stability of the open promoter complex. The growing RNA competes with TFIIB for RNA exit channel and results in the displacement of TFIIB and hence the disruption of the preinitiation complex during promoter escape, a critical transition from the transcription initiation phase to the elongation phase.



**Figure 1. A.** The  $F_0 - F_c$  difference electron density corresponding to TFIIB is shown in green and anomalous peaks corresponding to the selenomethionine derivatizations are indicated in blue; **B.** A structural model of a “minimal” closed promoter complex; **C.** A structural model of a “minimal” open promoter complex.

This research was supported by NIH grants GM049985 and AI21144 to R.D.K. X.L. was supported by the Jane Coffin Childs Memorial Fund fellowship. D.W. was supported by the NIH Pathway to Independence Award (K99 GM085136).

### Primary Citation

Liu, X., Bushnell, D. A., Wang, D., Calero, G., and Kornberg, R. D. (2010) Structure of an RNA Polymerase II–TFIIB Complex and the Transcription Initiation Mechanism, *Science*, **327**, 206-209.

### References

1. Y. Kim, J. H. Geiger, S. Hahn, P. B. Sigler, *Nature* **365**, 512 (1993).
2. T. Lagrange, A. N. Kapanidis, H. Tang, D. Reinberg, R. H. Eubright, *Genes Dev* **12**, 34 (1998).
3. W. Deng, S. G. Roberts, *Genes Dev* **19**, 2418 (2005).
4. J. D. Parvin, P. A. Sharp, *Cell* **73**, 533 (1993).
5. W. Hausner, J. Wettach, C. Hethke, M. Thomm, *J Biol Chem* **271**, 30144 (1996).
6. S. A. Qureshi, S. D. Bell, S. P. Jackson, *EMBO J* **16**, 2927 (1997).
7. X. Liu, D. A. Bushnell, D. Wang, G. Calero, R. D. Kornberg, *Science* **327**, 206 (2010).
8. D. A. Bushnell, K. D. Westover, R. E. Davis, R. D. Kornberg, *Science* **303**, 983 (2004).

SSRL is primarily supported by the DOE Offices of Basic Energy Sciences and Biological and Environmental Research, with additional support from the National Institutes of Health, National Center for Research Resources, Biomedical Technology Program, and the National Institute of General Medical Sciences.