Cell, Vol. 85, 773–779, May 31, 1996, Copyright ©1996 by Cell Press

Two-Dimensional Crystallography of TFIIB– and IIE–RNA Polymerase II Complexes: Implications for Start Site Selection and Initiation Complex Formation

Kerstin K. Leuther, David A. Bushnell, and Roger D. Kornberg
Department of Structural Biology
Stanford University School of Medicine
Stanford, California 94305–5400

Summary

Transcription factors IIB (TFIIB) and IIE (TFIIE) bound to RNA polymerase II have been revealed by electron crystallography in projection at 15.7 Å resolution. The results lead to simple hypotheses for the roles of these factors in the initiation of transcription. TFIIB is suggested to define the distance from TATA box to transcription start site by bringing TATA DNA in contact with polymerase at that distance from the active center of the enzyme. TFIIE is suggested to participate in a key conformational switch occurring at the active center upon polymerase–DNA interaction.

Introduction

A set of five general transcription factors (TF), TFIIB, -D, -E, -F, and -H, are required for initiation at most RNA polymerase II promoters (Conaway and Conaway, 1993; Maldonado and Reinberg, 1995). These factors exhibit extraordinary conservation across species; a one–to–one correspondence of all 15 essential subunits of the factors, as well as high sequence similarity, has been demonstrated between Saccharomyces cerevisiae and humans (Feaver et al., 1994; Henry et al., 1994; Svejstrup et al., 1996; J. Q. Svejstrup and R. D. K., unpublished data). Promoter DNA sequences targeted by these factors are similarly conserved. A "core" RNA polymerase II promoter usually comprises a TATA box and transcription start site, located about 30 bp apart. In S. cerevisiae, this spacing is a minimum of 30 bp, but more variable, ranging up to about 120 bp (Corden et al., 1980; Benoist and Chambon, 1981; Struhl, 1987). Much current work is directed toward interactions between the conserved general transcription factors and promoter DNA elements that define the distance from TATA box to start site and other aspects of the initiation mechanism.

Studies to date have identified a small subunit of TFIID, termed the TATA binding protein (TBP), as responsible for recognition of the TATA element of a promoter. TFIIB is believed to play a bridging role between the TBP promoter complex and RNA polymerase II, important for transcription start site determination (Buratowski et al., 1989; Tschochner et al., 1992; Pinto et al., 1992; Li et al., 1994). TFIIF, which interacts most tightly with the polymerase, may always remain associated and enter with polymerase into the initiation complex (Sopta et al., 1985; Burton et al., 1988; Buratowski et al., 1991; Killeen and Greenblatt, 1992; Tan et al., 1995; D. A. B. et al., unpublished data). TFIIE and TFIIH interact with one another (Maxon et al., 1994; D. A. B. et al., unpublished data) in a functionally relevant manner (Li et al., 1994) at a later stage of initiation complex assembly.

Further insight into the roles of general initiation factors may come from structural studies. Structures of the two smallest polypeptides, TBP and a C-terminal fragment of TFIIB, have been determined at atomic resolution by X–ray and nuclear magnetic resonance analyses, both individually and in complexes with DNA (Nikolov et al., 1992, 1995; Kim et al., 1993a, Kim et al., 1993b; Chasman et al., 1993; Bagby et al., 1995). TBP constrains DNA in a bent configuration, interacting with the outer (convex) surface of the bend. The C-terminal domain of TFIIB associates with the inner (concave) aspect of the bend, contacting the DNA backbone and the C-terminal stirrup of TBP. These findings shed light on the basis of TATA sequence recognition but leave open the major questions of start site determination and transcription initiation. To address these issues, structural studies of the larger general initiation factors and of very much larger complexes of these factors with RNA polymerase II are required, and such studies lie beyond the reach of X–ray and nuclear magnetic resonance methods, owing to the size (600 kDa) and complexity (12 polypeptide types) of the polymerase. We have therefore taken the alternative approach of electron microscope crystallography, which is particularly facile for revealing the arrangement and surface topography of proteins embedded in heavy atom stains (negative stain).

A general method of forming two-dimensional (2-D) crystals of proteins for structure determination by electron microscopy entails adsorption on lipid layers with appropriate surface properties (Kornberg and Darst, 1991; Jap et al., 1992; Brisson et al., 1994). In the case of several RNA polymerases, lipid layers with positively charged surfaces have been employed (Darst et al., 1989, 1991a; Edwards et al., 1990; Schultz et al., 1990). The 3-D structure of RNA polymerase II embedded in negative stain has been determined to a nominal resolution of 16 Å in this way (Darst et al., 1991b). We report here on extension of this analysis to RNA polymerase II complexed with TFIIB and TFIIE. The results lead to specific proposals concerning the roles of these factors in transcription start site determination and initiation complex assembly.

Results

TFIIB–RNA Polymerase II Cocrystals

TFIIB binds directly to RNA polymerase II, as shown by its specific retention on a column of immobilized polymerase (Tschochner et al., 1992). We sought to form 2-D crystals of the TFIIB–polymerase complex on lipid layers under the same conditions as used previously for the polymerase alone. These experiments employed a mutant form of RNA polymerase II lacking two small subunits, Rpb4 and Rpb7, that forms especially large, well ordered 2-D crystals. Similar results were obtained with 3- and 5-fold molar excesses of TFIIB over polymerase; only those at the 3-fold ratio are described below.

Although 2-D crystals were less abundant in the presence of TFIIB than in its absence, they were sufficient

Table 1. Summary of Crystallographic Data

Data Set	Number of Images	Number of Molecules	Average Phase Error (Degrees)	Resolution (Å)	Completeness of Data (%)
RNA Pol II	17	69,000	26.1	15.7	99.3
RNA Pol II: TFIIB (1:3)	8	24,000	21.3	15.7	94.7
RNA Pol II: TFIIB (1:5)	9	33,000	23.6	15.7	95.4
RNA Pol II: mutant TFIIB (1:3)	10	52,000	26.3	15.7	98.2
RNA Pol II: TFIIE (1:3)	12	32,000	27.2	15.7	95.1

The p1 unit cell (225.0 ± 2.0 Å × 225.2 ± 2.6 Å × 121.3 ± 0.4°) contains two molecules of RNA polymerase II. The average phase residual was estimated according to Henderson et al. (1986) for all data of IQ less than or equal to 7. IQ is a measure of the signal–to–noise ratio for each reflection, such that signal–to–noise equals 7/IQ. The mean figure of merit of the averaged structure factors (Henderson et al., 1986) is greater than 0.89 for all data sets to 15.7 Å resolution. Data to 13 Å resolution has a mean figure of merit of greater than 0.80 for all data sets. Completeness of data refers to the percentage of all reflections expected for the p1 plane group that are observed to 15.7 Å resolution.

for data collection and image processing in negative stain. Crystals with and without TFIIB were evidently isomorphous, allowing difference Fourier analysis. Several electron micrographs of each type of crystal were subjected to image processing, averaging of Fourier components, and Fourier synthesis (Table 1). The two data sets were complete to 15.7 Å resolution, and the resulting projected electron density maps (Figure 1) were essentially the same as that previously reported for the RNA polymerase alone (Darst et al., 1991a). Although not apparent from inspection of the individual maps, appreciable differences were revealed by vector subtraction of Fourier components and Fourier synthesis (Figure 2).

Several observations attest to the significance of peaks seen in the Fourier difference map. First, the two strong peaks were statistically significant at the 99.95% confidence level. Second, the occurrence of two peaks is consistent with biochemical evidence for the monomeric nature of TFIIB. Two difference peaks may reveal two TFIIB monomers interacting with two RNA polymerase II molecules in a unit cell of the 2-D crystal. Third, the size of the difference densities is in keeping with that expected for TFIIB and compatible with direct polymerase binding. A sphere of radius 20–23 Å centered on a TFIIB difference density would just contact the surface of polymerase at the contour level shown and would have a volume corresponding to a protein mass of 39 kDa, compared with a mass of 38 kDa from the deduced amino acid sequence of TFIIB. Fourth, the two

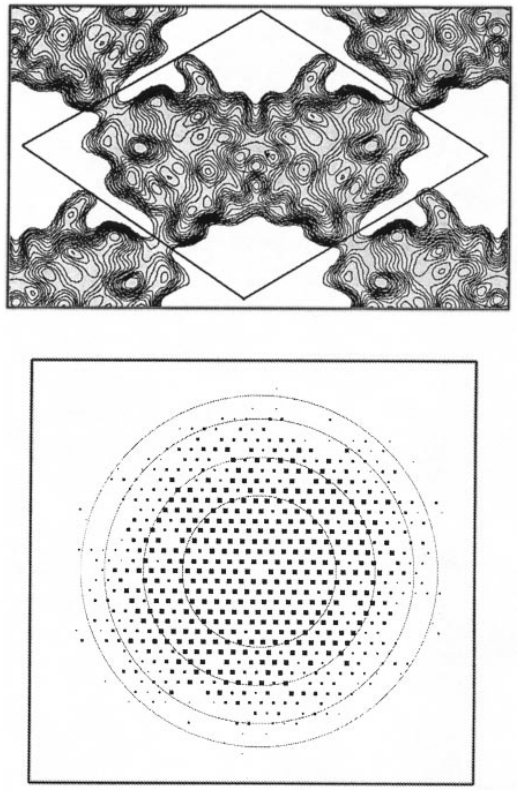

Figure 1. Projected Structure of TFIIB–RNA Polymerase II Complex

(Upper panel) Contour map of TFIIB–RNA polymerase II cocrystals from average of eight images in projection (RNA Pol II:TFIIB [1:3], Table 1). A unit cell (boxed in the contour map) contains two molecules of RNA polymerase II, related by an apparent 2-fold axis in the plane of the figure. (Lower panel) Combined structure factor data displayed as an IQ plot, illustrating data quality and sampling (Henderson et al., 1986). Larger squares indicate a larger signal–to–noise ratio. Circles indicate 30 Å (innermost circle), 20Å, 15Å, and 13Å (outermost circle) resolution.

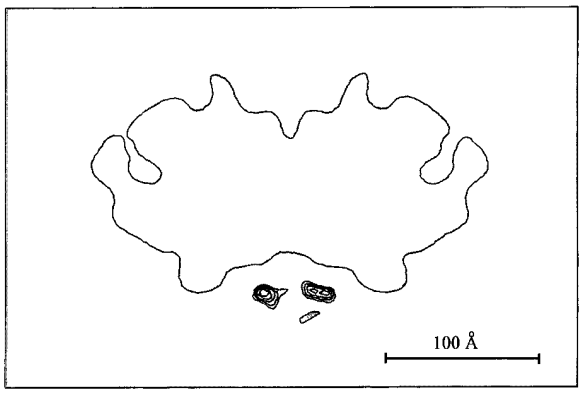

Figure 2. Fourier Difference Between TFIIB–RNA Polymerase II Complex and RNA Polymerase II, Superimposed on Outline of RNA Polymerase II

Difference between TFIIB–RNA polymerase II and RNA polymerase II data sets (RNA Pol II:TFIIB [1:3] and RNA Pol II, Table 1) was determined as described. Difference density greater than two standard deviations above the mean is displayed in contour format (shaded), superimposed on the outermost contour of the RNA polymerase map. The pair of RNA polymerase molecules in the center of a unit cell and associated differences are shown.

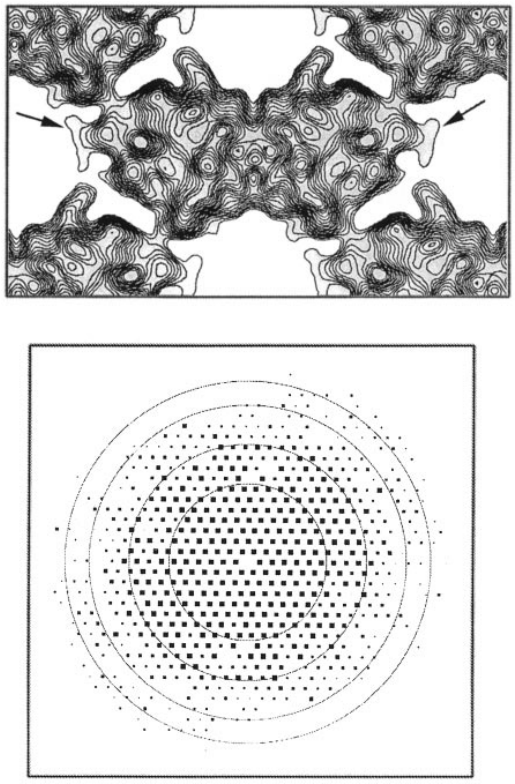

Figure 3. Projected Structure of TFIIE–RNA Polymerase II Complex

(Upper panel) Contour map of TFIIE–RNA polymerase II cocrystals from average of 12 images in projection (RNA Pol II:TFIIE [1:3], Table 1). Arrows indicate additional density due to TFIIE.

(Lower panel) Combined structure factor data displayed as an IQ plot.

strong difference peaks are related by the apparent 2-fold axis between the two RNA polymerase molecules in the unit cell and thus contact the same sites on the surface of the polymerase.

Further evidence for specificity of the TFIIB–RNA polymerase interaction seen in 2-D crystals came from studies with a mutant form of TFIIB in which arginine 78 was replaced by cysteine. TFIIB(R78C) was recovered from a screen for mutations altering the location of transcription start sites (Pinto et al., 1994). It was previously shown that deletion of the amino-terminal region of TFIIB prevented interaction with an RNA polymerase–TFIIF complex (Barberis et al., 1993; Buratowski et al., 1993; Ha et al., 1993; Hisatake et al., 1993; Malik et al., 1993; Yamashita et al., 1993; Lee and Hahn, 1995), and we have found by surface plasmon resonance that TFIIB(R78C) binds pure RNA polymerase II with an affinity approximately 100-fold lower than does wild-type TFIIB (D. A. B. et al., unpublished data). A mixture of TFIIB(R78C) and RNA polymerase was crystallized, and data were collected and processed exactly as for the mixture with wild-type TFIIB described above (Table 1). There were no significant peaks in a difference Fourier map between the TFIIB(R78C)–RNA polymerase crystals and those of RNA polymerase alone. The difference density seen with wild-type TFIIB therefore correlates with TFIIB–RNA polymerase II interaction, arguing against

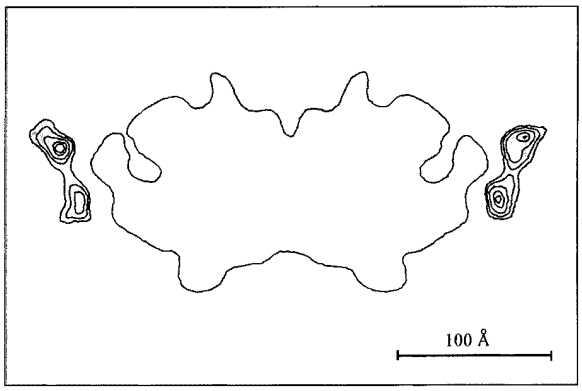

Figure 4. Fourier Difference Between TFIIE–RNA Polymerase II Complex and RNA Polymerase II, Superimposed on Outline of RNA Polymerase II

Difference between TFIIE–RNA polymerase II and RNA polymerase II data sets (RNA Pol II:TFIIE [1:3] and RNA Pol II, Table 1) was determined as described. Difference density greater than two standard deviations above the mean is displayed in contour format (shaded), superimposed on the outermost contour of the RNA polymerase map. The pair of RNA polymerase molecules in the center of a unit cell and associated differences are shown.

such possible artifacts as perturbation of the RNA polymerase II crystal structure or nonspecific binding to an irrelevant site on the polymerase surface.

TFIIE–RNA Polymerase II Cocrystals

In contrast with TFIIB, our previous work failed to show direct interaction of TFIIE with RNA polymerase II by conventional means, such as cochromatography or cosedimentation (Sayre et al., 1992). We have, however, detected TFIIE–RNA polymerase II interaction, by surface plasmon resonance, with an association constant of 3.9×10^7 M^{-1} (D. A. B. et al., unpublished data). This finding prompted us to attempt the formation and analysis of TFIIE–RNA polymerase II cocrystals in a manner analogous to that described for TFIIB. Crystals obtained with a 3-fold molar excess of TFIIE over RNA polymerase II were isomorphous with those of the polymerase alone, and data compiled from 12 electron micrographs were complete to 15.7 Å resolution (Table 1). A projected electron density map of the TFIIE–RNA polymerase II cocrystals displayed a feature not seen in the map of the polymerase alone, a prominent bulge on the surface of the polymerase molecule (Figure 3, arrows). This bulge gave rise to the sole feature in a Fourier difference map between TFIIE–RNA polymerase II and RNA polymerase II, a pair of peaks associated with each of the two polymerase molecules in the unit cell (Figure 4).

As in the case of TFIIB–RNA polymerase II cocrystals, there were several indications of significance of the TFIIE–RNA polymerase II difference densities. The difference peaks were statistically significant at the 99.95% confidence level. The occurrence of one pair of peaks associated with each polymerase molecule, corresponding in size to roughly spherical protein subunits of 64 kDa and 49 kDa (based on similar assumptions to those described above for TFIIB), is consis-

tent with biochemical evidence for a single TFIIE dimer of 66 kDa and 43 kDa subunits interacting with RNA polymerase II (Sayre et al., 1992; Feaver et al., 1994; D. A. B. et al., unpublished data). The two pairs of difference peaks are related by the apparent 2-fold axis between the two RNA polymerase molecules in the unit cell, further attesting to the specificity of TFIIE–RNA polymerase interaction.

Discussion

The central feature of RNA polymerase II promoters, the conserved spacing of TATA box and transcription start sites, may be understood in terms of results from electron crystallography. The argument, which follows, is based on the location of TFIIB bound to RNA polymerase II and the approximate location of the active center of the enzyme. From these locations, we deduce those of the TATA box and start site in a transcription initiation complex and obtain the physical distance between them.

The analysis presented here of TFIIB–RNA polymerase II cocrystals reveals the approximate center of mass of TFIIB, an essentially globular protein about 20 Å in diameter (Nikolov et al., 1995; Bagby et al. 1995). TFIIB, in turn, interacts near the center of the TATA sequence in promoter DNA (Lee and Hahn, 1995; Nikolov et al., 1995). The center of mass of polymerase-bound TFIIB therefore identifies the location of the TATA box in a transcription initiation complex to within about plus or minus 10 Å.

A common feature of polymerase structures identifies the active center region. A set of eight DNA and RNA polymerases so far analyzed by X–ray and electron crystallography exhibit a distribution of protein density around a 25 Å cleft that has been likened to a hand, with fingers, palm, and thumb domains (Ollis et al., 1985; Darst et al., 1989, 1991b; Kohlstaedt et al., 1992; Sousa et al., 1993, 1994; Schultz et al., 1993; Davies et al., 1994; Sawaya et al., 1994; Kim et al., 1995; Korolev et al., 1995). X–ray and electron microscope cocrystal structures show DNA binding in the cleft (Beese et al., 1993; Pelletier et al., 1994; C. Poglitsch et al., unpublished data), and X–ray structures at atomic resolution reveal three conserved catalytic residues in the palm domain forming the floor of the cleft, which places the active center in this region (Kohlstaedt et al., 1992; Sousa et al., 1993; Pelletier et al., 1994; Davies et al., 1994; Sawaya et al., 1994). A close fit of shape and dimensions of density surrounding the 25 Å cleft between the X–ray structure of Escherichia coli DNA polymerase I and electron microscope structures of E. coli RNA polymerase and yeast RNA polymerase II (Ollis et al., 1985; Darst et al., 1989; Polyakov et al., 1995), together with sequence similarities between palm domains of (the largest subunits of) these three enzymes (Allison et al., 1985), leave little doubt as to the location of the active center in the floor of the cleft of yeast RNA polymerase II. The transcription start site in promoter DNA must contact the active center in an initiation complex, so the start site should abut the floor of the cleft.

The locations of the TATA box and transcription start site in an initiation complex deduced in this way lie 110 Å apart in projection (center–to–center distance between TFIIB density and 25 Å cleft in Figure 5), which would correspond to 32 bp of B-form DNA. The coincidence with the distance of about 30 bp from TATA box to transcription start site characteristic of RNA polymerase II promoters suggests a simple geometrical basis for start site determination: TFIIB brings DNA in contact with RNA polymerase II through binding both a TBP–TATA box complex and the polymerase; the DNA then follows a straight path across the surface of the polymerase to the active center, where the initiation of transcription occurs. In S. cerevisiae, the straight path defines the minimum distance from TATA box to start site, and longer distances may reflect a capacity of the polymerase to scan downstream, with looping out of the intervening DNA. This hypothesis for start site determination rests on the firm identification of TFIIB and approximate active center locations and on the measurement of the shortest (straight line) path between them. The length of the path is not much affected by uncertainty in the location of TFIIB in the (z) direction perpendicular to the plane of the projection (plus or minus 30 Å in z, corresponding to plus or minus 4 Å in path length) or uncertainty in the location of the active center on the inner surface of the 25 Å cleft (plus or minus 10 Å). An important corollary of the hypothesis is that the DNA between the TATA box and transcription start site follows a straight path rather than bending around a curved surface of the enzyme. A possible channel for accommodating a linear DNA duplex between the TFIIB-binding and active sites can be identified in the polymerase structure (Figure 5, broken line) and remains to be verified by further crystallographic analysis.

A sterochemical basis for start site determination based on TFIIB–RNA polymerase II interaction is nicely consistent with previous biochemical and genetic evidence. As mentioned above, the distance from TATA box to transcription start site differs between S. cerevisiae and mammalian systems. When the TATA binding protein became available, it was swapped between systems, but without effect: the location of transcription start sites remained the same (Cavallini et al., 1988; Hahn et al., 1989; Kelleher et al., 1992). Neither the other general transcription factors nor RNA polymerase II could be exchanged in this manner, prompting the pursuit of a Schizosaccharomyces pombe system, more closely related to S. cerevisiae but with a TATA–to–start site distance similar to that in mammalian systems. Again, efforts to swap individual factors were unsuccessful, but a striking result was obtained by pairwise exchange: S. pombe TFIIB and RNA polymerase could substitute together for their counterparts in the S. cerevisiae system, and transcription start sites were shifted to those characteristic of S. pombe (Li et al., 1994). This finding demonstrated an interaction between TFIIB and RNA polymerase II that determines the location of the transcription start site and that is essential for the initiation of transcription. Genetic studies have given results along similar lines (Berroteran et al., 1994; Pinto et al., 1994), identifying mutations in TFIIB that affect the location of transcription start sites, although mutations in the large subunit of TFIIF and in at least two polymerase

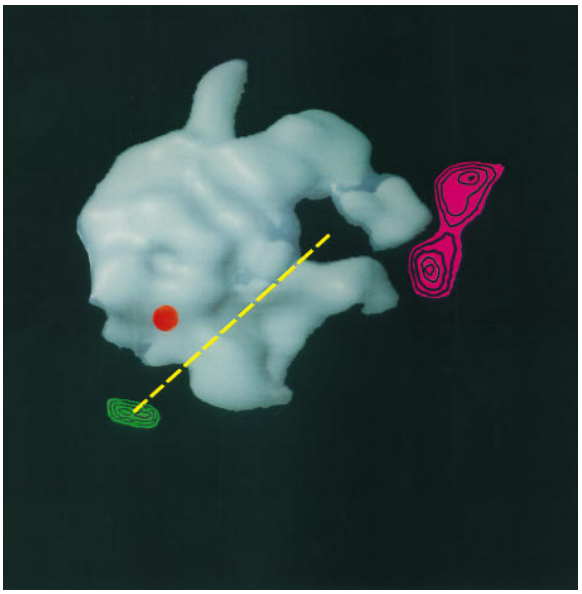

Figure 5. Composite Projection Map Showing Features of RNA Polymerase II Revealed or Surmised from Existing Evidence

TFIIB (green) and TFIIE (magenta) difference densities are superimposed on the 3-D structure of RNA polymerase II (blue; Darst et al., 1991b). The filled circle (red) indicates the approximate point at which the C-terminal repeat domain emanates from the surface of the polymerase into solution (Meredith et al., 1996). The broken yellow line extending from the TFIIB difference density to the center of the 25 Å channel thought to include the active site of the enzyme represents a possible path of DNA in the initiation complex and may lie on either the top or bottom surface of the enzyme in the direction of view (see Discussion).

subunits (Rpb1 and Rpb9) alter transcription start sites as well (Berroteran et al., 1994; Furter–Graves et al., 1994; Hull et al., 1995; Sun and Hampsey, 1995).

The location of the TFIIB binding site on RNA polymerase II reported here is consistent with other genetic and structural data. A region of about 140 amino acids near the C-terminus of the largest subunit of RNA polymerase II (Rpb1), which includes one of several blocks of sequence similarity among eukaryotic RNA polymerases, has been expressed as a fusion to an E. coli protein and shown to bind TFIIB (Xiao et al., 1994). Additionally, a mutation at a position immediately preceding this 140 residue region can alter the location of transcription start sites (Berroteran et al., 1994). The 140 residue region is followed in the sequence of Rpb1 by the so-called C-terminal repeat domain, and an antibody fragment bound to an epitope at the beginning of the C-terminal repeat domain has been resolved by 2-D crystallography (Meredith et al., 1996). The epitope is apparently located within a few Ångstroms of the point of contact of TFIIB with the surface of the polymerase (Figure 5). The proximity in amino acid sequence of the putative TFIIB binding region and the C-terminal repeat domain is thus reflected by a proximity in structure.

A question arises regarding the conformation of RNA polymerase II and its influence on the structure of the TFIIB–polymerase complex studied here. Electron crystallography of E. coli holoenzyme and core RNA polymerases has revealed a difference in the vicinity of the 25 Å cleft (Darst et al., 1989; Polyakov et al., 1995). Protein density extends about halfway around the channel in the holoenzyme but appears to surround the channel almost completely in the core polymerase. In all likelihood, the open conformation of the holoenzyme allows entry of DNA into the cleft, while the closed conformation in the core enzyme helps retain DNA and enhance processivity. The structure of RNA polymerase II in the 2-D crystals studied here resembles that of the E. coli core polymerase, especially in regard to the arm of protein density surrounding the 25 Å cleft, so RNA polymerase II in the 2-D crystals is probably in the elongation conformation. Since TFIIB is believed to be absent from the elongating enzyme, its presence in 2-D cocrystals might not be expected. The elongation conformation may, however, retain an affinity for TFIIB that is released by some other mechanism. Alternatively, the affinity for TFIIB may be much diminished, accounting for the low difference density and thus apparently low occupancy at the TFIIB binding site.

The location of TFIIE in cocrystals with RNA polymerase II reported here suggests a direct role in events involving the arm of density around the 25 Å cleft. TFIIE is clearly associated with the arm and may bridge the gap remaining after closure of the arm around the cleft (Figure 5, legend). TFIIE may therefore promote closure or stabilize or signal acquisition of the closed state. Such a role is consistent with biochemical evidence for involvement of TFIIE in late events of initiation complex assembly, including DNA transactions in the vicinity of the initiation site (Holstege et al., 1995; Zawel et al., 1995).

The results on TFIIE binding to RNA polymerase are supportive of those on TFIIB. The location of the TFIIE site remote from that for TFIIB argues against an artifact, such as nonspecific binding of many proteins to a common, functionally irrelevant site exposed in the 2-D crystals. The spacing between the TFIIB and TFIIE sites is, further, consistent with a geometrical basis for start site selection by TFIIB, based on linear distance, proposed above.

Experimental Procedures

Protein Purification

RNA polymerase II was purified from S. cerevisiae strain CB010ΔRPB4 (MATa *pep4*::HIS3 *prb1*::LEU2 *prc1*::HIS3 *rpb4*::URA3 *can1 ade2 trp1 ura3 his3 leu2–3,112 cir0*), derived from strain W303–1a. Details of the purification will be published elsewhere (D. A. B. et al., unpublished data). Recombinant wild-type and mutant yeast TFIIB were purified from E. coli (BL21/DE3) as described (Feaver et al., 1994), with several modifications. Briefly, after induction, harvest, and lysis of the cells, recombinant TFIIB was purified to homogeneity by DE52 (Whatman), Biorex70 (BioRad), hydroxyapatite (BioRad), and Sephacryl S-100HR (Pharmacia) column chromatography. Peak fractions containing pure TFIIB were precipitated with ammonium sulfate and stored at −80°C. Purity (greater than 98%) of the resulting material was assessed by SDS–polyacrylamide gel electrophoresis with Coomassie and silver staining and by dynamic light scattering. Function of the purified proteins was tested by electrophoretic mobility shift, polymerase-binding, and in vitro transcription assays (data not shown). Purification of recombinant TFIIE will be described elsewhere (D. A. B. et al., unpublished data).

Electron Crystallography

For lipid layer crystallization, a solution of RNA polymerase II (80–110 μg/ml) in crystallization buffer (50 mM Tris [pH 7.5], 60 mM ammonium sulfate, 5 mM spermidine, 10 mM DTT) was pipetted into a

cylindrical nylon well. Where transcription factors were required, they were mixed with RNA polymerase II on ice before the solution was pipetted into a well. The planar surface of the protein solution in the well was overlaid with a lipid mixture containing 0.45 mg/ml L-α-phosphatidylcholine and 0.05 mg/ml stearylamine (both from Avanti Polar Lipids) in 1:1 (v/v) chloroform/hexane. After incubation for 9–12 h at 4°C in a water-saturated argon atmosphere, a clean Pt/Pd loop was used to transfer the crystals to glow-discharged (in water vapor) carbon-coated electron microscope grids (Asturias and Kornberg, 1995). After washing with 0.05% Tween 20 and water, the grids were stained with 1% uranyl acetate.

Data Collection and Processing

All data were collected on a Phillips CM12 microscope in low dose mode and operating at 100 kV with a LaB_6 filament. Images were recorded on Kodak SO136 film at a magnification of 35,000 $\times$. Quality of the micrographs was assessed by optical diffractometry, and suitable micrographs were scanned on a Perkin Elmer flatbed scanner at a 20 μm stepsize. Crystals ranged in size from less than 1 μm^2 to greater than 6 μm^2. Average crystals in the data sets were about 2 μm^2.

Images were processed with the MRC image processing software (Amos et al., 1982; Henderson et al., 1986). Individual programs were modified locally to accommodate large crystal arrays and the large unit cell of the RNA polymerase II crystals. All data were processed in the crystallographic spacegroup p1. A unit cell contained two molecules of RNA polymerase II and had dimensions a = 225.0 $\pm$ 2.0Å, b = 225.2 $\pm$ 2.6Å, γ = 121.3 $\pm$ 0.4°. Multiple images were combined with the program ORIGTILT (Henderson et al., 1986) into one set for each projected structure. Fourier difference maps were derived by subtracting the combined RNA polymerase II data set from each of the combined cocrystal data sets. The difference maps were statistically analyzed using a Student's t test on a pixel–by–pixel basis (Meredith et al., 1996; Milligan and Flicker, 1987).

Acknowledgments

We thank Francisco Asturias, Weihau Chang, Gavin Meredith, and Claudia Poglitsch for assistance with lipid-layer crystallography and image processing; John Feaver for partially purified mutant TFIIB; Lynne Mercer for advice on electron microscopy; Ray Stevens for use of his light-scattering instrument; and Peter David for help with image processing, program modifications, and diffraction physics. Michael Hampsey and colleagues provided the plasmid containing the mutant TFIIB gene. Funding for this work was from the National Institutes of Health (grant AI21144 to R. D. K.) and from the Human Frontiers Science Program. K. K. L. is supported in part by fellowship DRG 1271 from the Cancer Research Fund of the Damon Runyon–Walter Winchell Foundation.

Received January 22, 1996; revised April 12, 1996.

References

Allison, L.A., Moyle, M., Shales, M., and Ingles, C.J. (1985). Extensive homology among the largest subunits of eukaryotic and prokaryotic RNA polymerases. Cell *42*, 599–610.

Amos, L.A., Henderson, R., and Unwin, P.N. (1982). Three-dimensional structure determination by electron microscopy of two-dimensional crystals. Prog. Biophys. Mol. Biol. *39*, 183–231.

Asturias, F.J., and Kornberg, R.D. (1995). A novel method for transfer of two-dimensional crystals from the air–water interface to specimen grids. J. Struct. Biol. *114*, 60–66.

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 TFIIB: similarity to cyclin A and interaction with TATA-binding protein. Cell *82*, 857–867.

Barberis, A., Müller, C.W., Harrison, S.C., and Ptashne, M. (1993). Delineation of two functional regions of transcription factor TFIIB. Proc. Natl. Acad. Sci. USA *90*, 5628–5632.

Beese, L.S., Derbyshire, V., and Steitz, T.A. (1993). Structure of DNA polymerase I Klenow fragment bound to duplex DNA. Science *260*, 352–355.

Benoist, C., and Chambon, P. (1981). In vivo sequence requirements of the SV40 early promoter region. Nature *290*, 304–310.

Berroteran, R.W., Ware, D.E., and Hampsey, M. (1994) The *sua8* suppressors of *Saccharomyces cerevisiae* encode replacements of conserved residues within the largest subunit of RNA polymerase II and affect transcription start site selection similarly to *sua7* (TFIIB) mutations. Mol. Cell. Biol. *14*, 226–237.

Brisson, A., Olofsson, A., Ringler, P., Schmutz, M., and Stoylova, S. (1994). Two-dimensional crystallization of proteins on planar lipid films and structure determination by electron crystallography. Biol. Cell *80*, 221–228.

Buratowski, S., and Zhou, H. (1993). Functional domains of transcription factor TFIIB. Proc. Natl. Acad. Sci. USA *90*, 5633–5637.

Buratowski, S., Hahn, S., Guarente, L., and Sharp, P.A. (1989). Five intermediate complexes in transcription initiation by RNA polymerase II. Cell *56*, 549–561.

Buratowski, S., Sopta, M., Greenblatt, J., and Sharp, P.A. (1991). RNA polymerase II–associated proteins are required for a DNA conformation change in the transcription initiation complex. Proc. Natl. Acad. Sci. USA *88*, 7509–7513.

Burton, Z.F., Killeen, M., Sopta, M., Ortolan, L.G., and Greenblatt, J.F. (1988). RAP30/74: a general transcription factor that binds to RNA polymerase II. Mol. Cell. Biol. *8*, 1602–1613.

Cavallini, B., and Huet, J. (1988). A yeast activity can substitute for the hela cell TATA box factor. Nature *334*, 77–80.

Chasman, D.I., Flaherty, K.M., Sharp, P.A., and Kornberg, R.D. (1993). Crystal structure of yeast TATA–binding protein and model for interaction with DNA. Proc. Natl. Acad. Sci. USA *90*, 8174–8178.

Conaway, R.C., and Conaway, J.W. (1993). General initiation factors for RNA polymerase II. Annu. Rev. Biochem. *62*, 161–190.

Corden, J., Wasylyk, B., Buchwalder, A., Corsi, P.S., Kedinger, C., and Chambon, P. (1980). Promoter sequences of eukaryotic protein-coding genes. Science *209*, 1406–1414.

Darst, S.A., Kubalek, E.W., and Kornberg, R.D. (1989). Three-dimensional structure of Escherichia coli RNA polymerase holoenzyme determined by electron crystallography. Nature *340*, 730–732.

Darst, S.A., Kubalek, E.W., Edwards, A.M., and Kornberg, R.D. (1991a). Two-dimensional and epitaxial crystallization of a mutant form of yeast RNA polymerase II. J. Mol. Biol. *221*, 347–357.

Darst, S.A., Edwards, A.M., Kubalek, E.W., and Kornberg, R.D. (1991b). Three-dimensional structure of yeast RNA polymerase II at 16Å resolution. Cell *66*, 121–128.

Davies, J.F., Almassy, R.J., Hostomska, Z., Ferre, R.A., and Hostomsky, Z. (1994). 2.3 Å crystal structure of the catalytic domain of DNA polymerase β. Cell *76*, 1123–1133.

Edwards, A.M., Darst, S.A., Feaver, W.J., Thompson, N.E., Burgess, R.R., and Kornberg, R.D. (1990). Purification and lipid layer crystallization of yeast RNA polymerase II. Proc. Natl. Acad. Sci. USA *87*, 2122–2126.

Feaver, W.J., Henry, N.L., Bushnell, D.A., Sayre, M.H., Brickner, J.H., Gileadi, O., and Kornberg, R.D. (1994). Yeast TFIIE. J. Biol. Chem. *269*, 27549–27553.

Furter–Graves, E.M., Hall, B.D., and Furter, R. (1994) Role of a small RNA pol II subunit in TATA to transcription start site spacing. Nucl. Acids Res. *22*, 4932–4936.

Ha, I., Roberts, S., Maldonado, E., Sun, X., Kim, L.-U., Green, M., and Reinberg, D. (1993). Multiple functional domains of human transcription factor IIB: distinct interactions with two general transcription factors and RNA polymerase II. Genes Dev. *7*, 1021–1032.

Hahn, S., Buratowski, P.A., Sharp, P.A., and Guarente, L. (1989). Isolation of the gene encoding the yeast TATA binding protein TFIID: a gene identical to the SPT15 suppressor of Ty element insertions. Cell *58*, 1173–1182.

Henderson, R., Baldwin, J.M., Downing, K.H., Lepault, J., and Zemlin, F. (1986). Structure of purple membrane from Halobacterium

halobium: recording, measurement and evaluation of electron micrographs at 3.5 Å resolution. Ultramicroscopy *19*, 147–178.

Henry, N.L., Campbell, A.M., Feaver, W.J., Poon, D., Weil, A., and Kornberg, R.D. (1994). TFIIF–TAF–RNA polymerase II connection. Genes Dev. *8*, 2868–2878.

Hisatake, K., Roeder, R.G., and Horikoshi, M. (1993). Functional dissection of TFIIB domains required for TFIIB–TFIID–promoter complex formation and basal transcription activity. Nature *363*, 744–747.

Holstege, F.C.P., Tantin, D., Carey, M., van der Vliet, P.C., and Timmers, H.T.M. (1995). The requirement for the basal transcription factor IIE is determined by the helical stability of promoter DNA. EMBO J. *14*, 810–819.

Hull, M.W., McKune, K., and Woychik, N.A. (1995). RNA polymerase II subunit RPB9 is required for accurate start site selection. Genes Dev. *9*, 481–490.

Jap, B.K., Zulauf, M., Scheybani, T., Hefti, A., Baumeister, W., Aebi, U., and Engel, A. (1992). 2-D crystallization: from art to science. Ultramicroscopy *46*, 45–84.

Kelleher, R.J., III, Flanagan, P.M., Chasman, D.I., Ponticelli, A.S., Struhl, K., and Kornberg, R.D. (1992). Yeast and human TFIIDs are interchangeable for the response to acidic transcriptional activators in vitro. Genes Dev. *6*, 296–303.

Killeen, M.T., and Greenblatt, J.F. (1992). The general transcription factor RAP30 binds to RNA polymerase II and prevents it from binding nonspecifically to DNA. Mol. Cell. Biol. *12*, 30–37.

Kim, J.L., Nikolov, D.B., and Burley, S.K. (1993a). Cocrystal structure of TBP recognizing the minor groove of a TATA element. Nature *365*, 520–528.

Kim, Y., Geiger, J.H., Hahn, S., and Sigler, P.B. (1993b). Crystal structure of a yeast TBP/TATA complex. Nature *365*, 512–520.

Kim, Y., Eom, S.H., Wang, J., Lee, D.-S. Suh, S.W., and Steitz, T.A. (1995). Crystal structure of Thermus aquaticus DNA polymerase. Nature *376*, 612–616.

Kohlstaedt, L.A., Wang, J., Friedman, J.M., Rice, P.A., and Steitz, T.A. (1992). Crystal structure at 3.5 Å resolution of HIV-1 reverse transcriptase complexed with an inhibitor. Science *256*, 1783–1790.

Kornberg, R.D., and Darst, S.A. (1991). Two-dimensional crystals of proteins on lipid layers. Curr. Opin. Struct. Biol. *1*, 642–646.

Korolev, S., Nayal, M., Barnes, W.M., Di Cera, E., and Waksman, G. (1995). Crystal structure of the large fragment of Thermus aquaticus DNA polymerase I at 2.5 Å resolution: structural basis for thermostability. Proc. Natl. Acad. Sci. *92*, 9264–9268.

Lee, S., and Hahn, S. (1995). Model for binding of transcription factor TFIIB to the TBP–DNA complex. Nature *376*, 609–612.

Li, Y., Flanagan, P.M., Tschochner, H., and Kornberg, R.D. (1994). RNA polymerase II initiation factor interactions and transcription start site selection. Science *263*, 805–807.

Maldonado, E., and Reinberg, D. (1995). News on initiation and elongation of transcription by RNA polymerase II. Curr. Opin. Cell Biol. *7*, 352–361.

Malik, S., Lee, D.K., and Roeder, R.G. (1993). Potential RNA polymerase II–induced interactions of transcription factor TFIIB. Mol. Cell. Biol. *13*, 6253–6259.

Maxon, M.E., Goodrich, J.A., and Tjian, R. (1994). Transcription factor IIE binds preferentially to RNA polymerase IIa and recruits TFIIH: a model for promoter clearance. Genes Dev. *8*, 515–524.

Meredith, G.D., Chang, W.H., Li, Y., Bushnell, D.A., Darst, S.A., and Kornberg, R.D. (1996). The C-terminal domain in the structure of RNA polymerase II. J Mol. Biol. *258*, 413–419.

Milligan, R.A., and Flicker, P.F. (1987). Structural relationships of actin, myosin, and tropomyosin revealed by cryo-electron microscopy. J. Cell Biol. *105*, 29–39.

Nikolov, D.B., Hu, S.-H., Lin, J., Gasch, A., Hoffmann, A., Horikoshi, M., Chua, N.-H., Roeder, R.G., and Burley, S.K. (1992). Crystal structure of TFIID–TATA box binding protein. Nature *360*, 40–46.

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 TFIIB-TBP-TATA–element ternary complex. Nature *377*, 119–128.

Ollis, D.L., Brick, P., Hamlin, R., Young, N.G., and Steitz, T.A. (1985). Structure of large fragment of Escherichia coli DNA polymerase I complexed with dTMP. Nature *313*, 762–766.

Pelletier, H., Sawaya, M.R., Kumar, A., Wilson, S.H., and Kraut, J. (1994). Structures of ternary complexes of rat DNA polymerase β, a DNA template-primer, and ddCTP. Science *264*, 1891–1903.

Pinto, I., Ware, D.E., and Hampsey, M. (1992). The yeast *sua7* gene encodes a homolog of human transcription factor TFIIB and is required for normal start site selection in vivo. Cell *68*, 977–988.

Pinto, I., Wu, W.-H., Na, J.G., and Hampsey, M. (1994). Characterization of *sua7* mutations defines a domain of TFIIB involved in transcription start site selection in yeast. J. Biol. Chem. *269*, 30569–30573.

Polyakov, A., Severinova, E., and Darst, S.A. (1995). Three-dimensional structure of E. coli core RNA polymerase: promoter binding and elongation conformations of the enzyme. Cell *83*, 365–373.

Sawaya, M.R., Pelletier, H., Kumar, A., Wilson, S.H., and Kraut, J. (1994). Crystal structure of rat DNA polymerase β: evidence for a common polymerase mechanism. Science *264*, 1930–1935.

Sayre, M.H., Tschochner, H., and Kornberg, R.D. (1992). Purification and properties of Saccharomyces cerevisiae RNA polymerase II general transcription factor a. J. Biol. Chem. *267*, 23383–23387.

Schultz, P., Celia, H., Riva, M., Darst, S.A., Colin, P., Kornberg, R.D., Sentenac, A., and Oudet, P. (1990). Structural study of the yeast RNA polymerase A. J. Mol. Biol. *216*, 353–362.

Schultz, P., Celia, H., Riva, M., Sentenac, A., and Oudet, P. (1993). Three-dimensional model of yeast RNA polymerase I determined by electron microscopy of two-dimensional crystals. EMBO J. *12*, 2601–2607.

Sopta, M., Carthew, R.W., and Greenblatt, J.F. (1985). Isolation of three proteins that bind to mammalian RNA polymerase II. J. Biol. Chem. *260*, 10353–10360.

Sousa, R., Chung, Y.J., Rose, J.P., and Wang, B.-C. (1993). Crystal structure of bacteriophage T7 RNA polymerase at 3.3 Å resolution. Nature *364*, 593–599.

Sousa, R., Rose, J.P., and Wang, B.-C. (1994). The thumb's knuckle: flexibility in the thumb subdomain of T7 RNA polymerase is revealed by the structure of a chimeric T7/T3 RNA polymerase. J. Mol. Biol. *244*, 6–12.

Struhl, K. (1987). Promoters, activator proteins, and the mechanism of transcriptional initiation in yeast. Cell *49*, 295–297.

Sun, Z.-W., and Hampsey, M. (1995). Identification of the gene (SSU71/TFG1) encoding the largest subunit of transcription factor TFIIF as a suppressor of a TFIIB mutation in Saccharomyces cerevisiae. Proc. Natl. Acad. Sci. USA *92*, 3127–3131.

Svejstrup, J.Q., Feaver, W.J., and Kornberg, R.D. (1996). Subunits of yeast RNA polymerase II transcription factor TFIIH encoded by the CCL1 gene. J. Biol. Chem. *2*, 643–645.

Tan, S., Conaway, R.C., and Conaway, J.W. (1995). Dissection of transcription factor TFIIF functional domains required for initiation and elongation. Proc. Natl. Acad. Sci. USA *92*, 6042–6046.

Tschochner, H., Sayre, M.H., Flanagan, P.M., Feaver, W.J., and Kornberg, R.D. (1992). Yeast RNA polymerase II initiation factor e: isolation and identification as the functional counterpart of human transcription factor IIB. Proc. Natl. Acad. Sci. USA *89*, 11292–11296.

Xiao, H., Friesen, J.D., and Lis, J.T. (1994). A highly conserved domain of RNA polymerase II shares a functional element with acidic activation domains of upstream transcription factors. Mol. Cell. Biol. *14*, 7507–7516.

Yamashita, S., Hisatake, K., Kokubo, T., Doi, K., Roeder, R.G., Horikoshi, M., and Nakatani, Y. (1993). Transcription factor TFIIB sites important for interaction with promoter-bound TFIID. Science *261*, 463–466.

Zawel, L., Kumar, K.P., and Reinberg, D. (1995). Recycling of the general transcription factors during RNA polymerase II transcription. Genes Dev. *9*, 1479–1490.