

Crystal Structure of the Yeast TFIIA/TBP/DNA Complex

James H. Geiger, Steve Hahn, Sally Lee, Paul B. Sigler*

The crystal structure of the yeast TFIIA/TBP/TATA promoter complex was solved to 3 angstrom resolution by double-edge multiple wavelength anomalous diffraction from two different species of anomalous scattering elements in the same crystal. The large and small subunits of TFIIA associate intimately to form both domains of a two-domain folding pattern. TFIIA binds as a heterodimer to the side of the TBP/TATA complex opposite to the side that binds TFIIB and does not alter the TBP/DNA interaction. The six-stranded β -sandwich domain interacts with the amino-terminal end of TBP through a stereospecific parallel β -strand interface and with the backbone of the TATA box and the 5'-flanking B-DNA segment. The four-helix-bundle domain projects away from the TBP/TATA complex, thereby presenting a substantial surface for further protein-protein interactions.

Initiation of transcription by RNA polymerase II (Pol II) requires a set of basal transcription factors that together with Pol II form a stable preinitiation complex (PIC) at core promoter elements (1, 2). PIC formation is initiated by the binding of the multiprotein complex TFIID to the TATA box of the core promoter through its central component, the TATA-binding protein (TBP).

The basal transcription factor TFIIA associates with the PIC through interactions with the TBP subunit of TFIID (3). TFIIA stimulates both basal and activated transcription (4–7), and cells depleted of TFIIA show reduced levels of Pol II transcription (6, 7). One probable function of TFIIA is to counteract the effect of inhibitors of transcription that target TBP, possibly by competing with the inhibitors for binding sites near or on TBP (1, 2). Additional evidence suggests that TFIIA acts at an early step of PIC formation to stabilize an active form of an activator/TFIID/TFIIA/DNA complex that is then rapidly assembled into a functional PIC (8). In support of this role, TFIIA can greatly accelerate the formation of, and stabilize, an activator/TFIID/DNA complex (9, 10).

Yeast TFIIA (γ TFIIA) consists of two subunits encoded by the genes TOA1 (large subunit) and TOA2 (small subunit) (11), both of which are essential for growth of yeast. The molecular mechanism of γ TFIIA function is conserved, as γ TFIIA can substitute for human TFIIA (hTFIIA) in transcription in vitro (6, 12, 13). Comparison of the TFIIA large-subunit sequences from yeast, human, and *Drosophila* shows amino- and carboxyl-terminal homologous seg-

ments of 56 and 71 residues, respectively (14–16) (Fig. 1). Extensive deletion analysis demonstrated that the highly variable sequence segment separating the homologous regions is not essential (7). In addition, deletion and alanine-scanning mutagenesis have identified regions of the large subunit likely to be involved in TBP and DNA interaction (7). Recent determination of the structures of TBP/TATA complexes illuminated the stereochemical details of the interaction of TBP with the minor groove of the TATA box and the concomitant distortion of the DNA caused by this interaction (17–19). The recently determined structure of the TBP/TFIIB/TATA complex (20) establishes the interaction between TFIIB and the TBP/TATA complex and positions this factor in relation to the PIC assembly. Here, we report the crystal structure of a yeast TFIIA/TBP/DNA complex refined to 3.0 Å resolution. The structure shows the stereochemistry by which TFIIA binds to the TBP/promoter complex and furthers our understanding of the architecture of the PIC assembly.

Structure determination. For structural studies we chose the smallest version of γ TFIIA that still retained full biological activity. Deletion of nonconserved residues 55 to 215 from the large subunit (LSU) produced a slow-growth and temperature-sensitive phenotype (7). Smaller internal deletions in the TFIIA LSU were analyzed for genetic complementation of a TOA1 deletion strain. We found that deletion of residues 95 to 209 of the LSU was the largest internal deletion that completely complemented the TOA1 deletion. A recombinant variant of γ TFIIA produced in *Escherichia coli* containing a full-length small subunit (SSU) and an LSU in which residues 95 to 209 were deleted (Δ 113LSU) complemented the transcription defect in extracts depleted of γ TFIIA (7). The specific activity of Δ 113LSU assayed by

in vitro transcription was less than 50% that of full-length TFIIA (21). This Δ 95–209 variant of the LSU in combination with the full-length SSU was crystallized with the carboxyl-terminal conserved core of γ TBP (γ TBPc) and an 18-base pair (bp) fragment of the yeast CYC1 promoter (Fig. 1). Details leading to the formation of diffraction-quality crystals are given in the legend to Table 1.

The crystal structure was solved by multiple-wavelength anomalous diffraction (MAD) (22, 23) at 3.0 Å with the use of both selenium and bromine x-ray absorption edges simultaneously. Crystals grown with selenomethionine (Se-Met)-substituted TBP and TFIIA (both subunits), and an oligonucleotide substituted with five bromine atoms, were used in the experiment. Phases from a molecular replacement solution with the γ TBPc/CYC1 TATA box complex (18) as a search model aided in the location of the anomalous scattering sites. The initial MAD-phased electron density map was improved by solvent flattening and gave interpretable density for TBP, DNA, and most of the β -sandwich region of TFIIA. Combination of the MAD phases with those of the refined partial model allowed us to complete the interpretation of the map.

The model was refined against one of the MAD data sets ($\lambda = 0.9235$) with multiple cycles of positional refinement and simulated annealing to 3.0 Å resolution, interspersed with manual rebuilding and extension of the model (Table 1). The present model includes residues 61 to 240 of TBP, residues 5 to 122 of the TFIIA SSU, residues 11 to 60 and 220 to 286 of the TFIIA LSU, all 18 base pairs of the DNA, and no water molecules. The structure has an *R* value of 23.2% (data $> 2\sigma$) and free *R* value of 35.0% (data $> 2\sigma$). An example of the electron density phased by MAD and improved by solvent flattening (DPHASES) is shown in Fig. 2. The structure has excellent stereochemistry with no violations of ϕ, ψ torsional restraints in the backbone.

Overall structure of TFIIA. TFIIA interacts with the TBP/promoter complex as a heterodimer consisting of two major structural elements: (i) a six-stranded β -sandwich domain that is composed of two β sheets oriented 75° to one another and (ii) a left-handed four-helix-bundle domain (Fig. 3). The two subunits of TFIIA are intimately associated in both domains. The conserved carboxyl-termini of each subunit contribute three strands to the β sandwich, whereas the conserved amino-termini of each subunit contribute two helices to the four-helix bundle. Even though the two TFIIA subunits share no sequence homology, the phylogenetically conserved regions of the two subunits have similar overall folds beginning with two helices and ending in three strands (Fig. 3). Three antiparallel

J. H. Geiger and P. B. Sigler are in the Department of Molecular Biophysics and Biochemistry and the Howard Hughes Medical Institute, Yale University, New Haven, CT 06510, USA. S. Hahn and S. Lee are at the Fred Hutchinson Cancer Research Center, A1-162, Seattle, WA 98104, USA.

*To whom correspondence should be addressed.

strands from each of the two subunits meet in the middle of the β sandwich and form a parallel β -strand interaction between L- β 1 and S- β 1 (24). A large interface securely orients the two domains relative to one another and defines one of the three distinct hydrophobic cores of the protein. The other two hydrophobic cores are the inner core of the four-helix-bundle domain and the interface between the two β sheets of the β -sandwich domain. In addition to the two major domains, an acidic loop is packed against one side of the β -sandwich domain containing residues 220 to 240 of the LSU and is stabilized by interaction with SK119 (25) and several residues near the amino-terminus of the SSU. The 48 residues of Δ 113LSU between LF60 (25) to LS220 are

disordered in this structure, but this length is more than sufficient to span the gap between these two residues in the structure. The fold of the β -sandwich domain has not, to our knowledge, been characterized previously. In fact, the TFIIA structure is distinct from any protein structures archived in the protein databank.

Most of the buried residues make intersubunit interactions, thereby forming a solvent-excluded surface area of 6000 \AA^2 between the two subunits. This explains the necessity of providing both subunits for proper folding (11); indeed, it is not possible to imagine either of these subunits existing as a structural or functional unit without the other. Although TFIIA is an $\alpha_2\beta_2$ tetramer in solution (11), it is an $\alpha\beta$ dimer when bound to the

TBP/TATA complex. This behavior is similar to that of TBP, which is a dimer in solution but binds the TATA box as a monomer.

Structure of the complex. TFIIA/TBP interactions. TFIIA binds exclusively to the amino-terminal half of TBP, which places the TFIIA contact site on the downstream subdomain of the pseudosymmetrical TBP. However, the severely arched nature of the TBP promoter complex allows the TFIIA to reach back and face the minor groove of the DNA just upstream of the TATA box (Fig. 4). All of the TFIIA/DNA and TFIIA/TBP interactions are made with the β -sandwich domain of TFIIA; the four-helix bundle projects away from the TBP/TATA complex. The edge of the β sheet, comprising S- β 1, docks against the β 2 strand of TBP in parallel

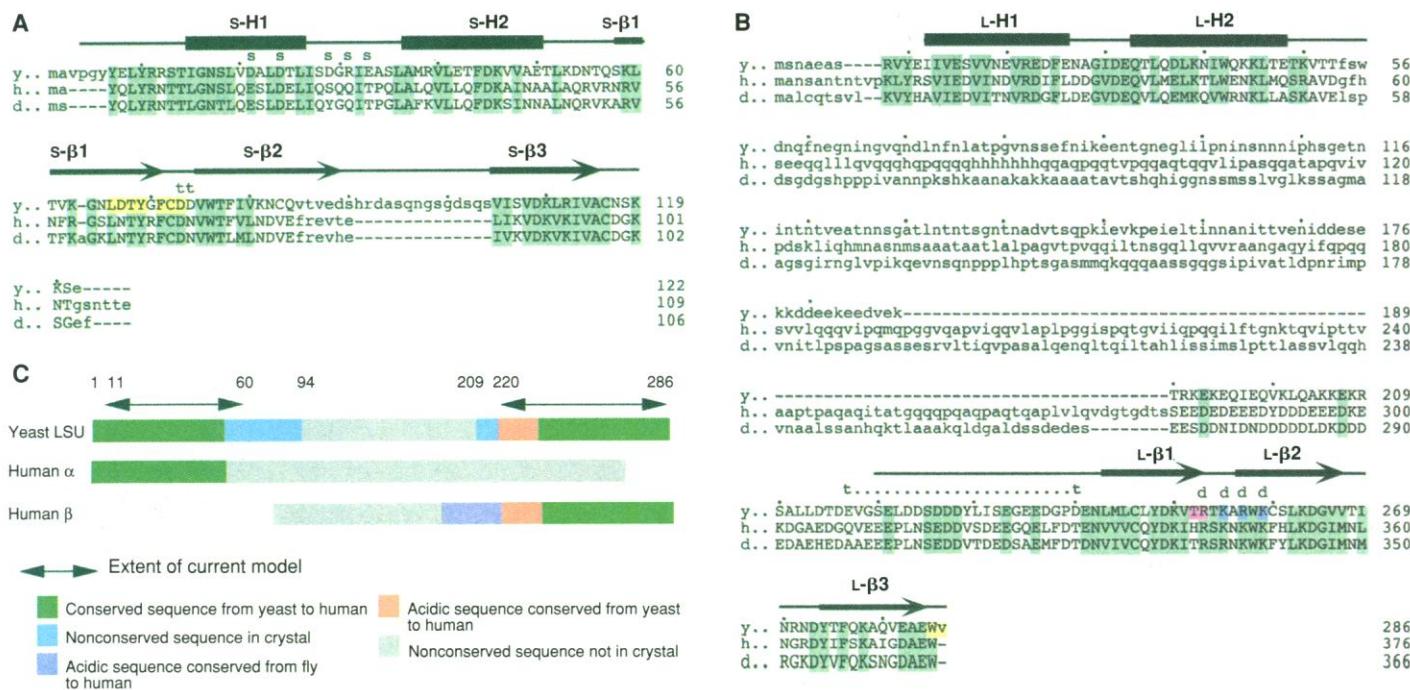


Fig. 1. Amino acid sequences for yeast (y), human (h), and *Drosophila* (d) TFIIA subunits. (A) TFIIA small subunit (SSU) sequences. (B) TFIIA large subunit (LSU) sequences. Residues that are identical or similar between yeast, human, and *Drosophila* are shaded green. Secondary structure elements for yTFIIA are indicated above the sequence; a bar indicates an α helix, and an arrow indicates a β sheet. Residues in the yeast LSU that contact DNA in the crystal structure are shaded red. Residues that appear to interact with variable geometry are highlighted in violet. Residues in both yeast subunits that contact TBP are shaded yellow. Residues that when substituted with alanine reduce or eliminate interaction of yTFIIA with TBP are indicated by "t" above the sequence (7). Sequence segments that when deleted eliminate interaction of yTFIIA with TBP are indicated by "t...t." Residues that when substituted with alanine bind TBP but are defective in formation of the yTFIIA/TBP/DNA complex are indicated by "d" above the sequence (7). Residues that when substituted with alanine give a temperature-sensitive phenotype but bind TBP and TBP/DNA complexes normally are indicated by "s" above the SSU sequence (7). (C) Schematic comparison of the yeast TFIIA LSU with the human α and β subunits. Shown are sequences conserved among species from yeast to human (green), nonconserved sequence contained in the Δ 113LSU used in our crystal (blue), an acidic sequence conserved from *Drosophila* to human (purple), an acidic sequence conserved from yeast to human (red), and nonconserved sequences deleted in the Δ 113LSU construct or found in the human protein (gray). (D) Canonical TATA sequence found in eukaryotes.

(E) Sequences of the two oligonucleotides corresponding to sequence from the yeast CYC1 promoter and used in the crystallization. The red box indicates the 8 bp contacted by TBP in the TFIIA/TBP/TATA complex. Numbering is as described in the text. (F) Comparison of the oligonucleotides used in the present study to that used in the structure of the yTBP/CYC1/TATA-box complex (18). The red box indicates the sequence bound by TBP in the present ternary complex, and the black box represents the sequence bound by TBP in the yTBP/CYC1-TATA box complex. The alignment of TFIIA sequences was performed with MACAW (48). Abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.

orientation, essentially joining the nine-stranded TBP β sheet to the six-stranded sheet of TFIIA (Fig. 4B). This strand-to-strand interface, along with the adjoining regions, constitutes the majority of the interface between the two proteins (Fig. 4B). The effect of this strand-to-strand interaction is to orient the rest of the TFIIA molecule on the upstream face of the TBP/promoter complex (26), thereby causing the β -sheet domain to form a wall abutting the first eight bases of the upstream DNA as described below.

Although modest in area (1600 Å² of solvent-excluded surface), the TFIIA/TBP interface is conserved and stereospecific

(Figs. 1 and 5). The interaction of SY69 and LW285 in this region is especially noteworthy (Fig. 4C). These two large hydrophobic residues join to insert themselves into a single cavity in TBP, the walls of which are defined by the TBP side chains of E93, R105, and R107 and the main-chain atoms of M104, R105, and I106. In addition, the phenolic hydroxyl of SY69 forms hydrogen bonds with R105 and E93. The appropriate conformation of the guanidinium group of R107 is stabilized by a salt bridge with the carboxyl-terminus of the LSU. Both E93 and R105 interact with the TATA box in the TBP/TATA binary complex (18), and

these interactions are retained in the ternary complex, with little change in conformation. In addition, the main-chain atoms of R105 and I106 make hydrophobic interactions with LW285. These residues are positioned by Y139 of TBP, which accepts a hydrogen bond from the main-chain NH of I106 and is packed securely against these residues. The γ TBP Y139A, K138A (27) double mutant is completely deficient in TFIIA binding, at least in part as a result of the loss of this critical TBP Y139 buttressing interaction. The interactions that occur between TFIIA and TBP are summarized in Fig. 5. As shown, most of the interactions

Table 1. Summary of γ TFIIA structure determination. The two subunits of γ TFIIA were overexpressed in two separate strains of *E. coli*, isolated in 6 M urea, and refolded as described (11). A modified version of the LSU was used that lacks the nonconserved residues 95 to 209. This construct fully rescues a WT TOA1 deletion in yeast and has about 50% the lower specific activity in *in vitro* transcription assays relative to WT TFIIA (44). This is the shortest construct of the LSU that behaves like the wild type in these assays and was necessary for the production of diffraction-quality crystals. The refolded TFIIA was concentrated with Centriprep 10 (Amicon) concentrators and purified by SOURCE Q anion exchange chromatography (Pharmacia). To form the ternary complex, we mixed γ TFIIA in stoichiometric amounts with the 18-bp blunt-ended oligonucleotide containing the CYC1 promoter TATA box (Fig. 1C), purified by Mono Q anion exchange chromatography, and the 185-residue carboxyl-terminal core of yeast TBP was prepared as described (18). Diffraction-quality crystals of this complex were obtained from sitting drops containing 45 mM ammonium acetate, 15 mM sodium acetate, 5% (v/v) glycerol, and 2.5 mM dithiothreitol (pH 4.5) equilibrated against a reservoir of 90 mM ammonium acetate, 30 mM sodium acetate, and 5% glycerol (pH 4.5) at 18°C. Parallelepiped-shaped rods with dimensions up to 0.8 mm by 0.25 mm by 0.25 mm grew in 1 week. The space group is $P2_12_12_1$ with unit cell dimensions of $a = 61.7$, $b = 94.0$, $c = 124.0$ Å. Crystals were stabilized by slowly adding a solution containing 60 mM ammonium acetate, 5% glycerol, and 20 mM sodium acetate (pH 4.5) to the drop in which the crystals were

grown. The concentration of glycerol in this solution was gradually increased, whereas the concentration of ammonium acetate was gradually decreased by slowly adding appropriate solutions to the drop and then removing mother liquor until the concentration of the drop was 40 mM ammonium acetate and 30% glycerol—a solution suitable for flash freezing. Crystals were then mounted in nylon loops and flash-frozen in liquid propane cooled in liquid nitrogen. Highest resolution data were collected to 3.0 Å resolution at NSLS beamline X4A at 100 K on imaging plates. Five wavelengths of MAD data were collected at NSLS X4A at 100 K on imaging plates. All data were processed with DENZO and scaled with SCALEPACK. The Se-Met-substituted proteins were produced by growing the bacterial strains in minimal media, with Se-Met replacing methionine. Brominated DNA was synthesized by standard chemical techniques replacing 5-bromouracil for thymine in the sequence. Using the γ TBP/TATA-box complex as a search model, a molecular replacement solution (AMoRe) (45) provided phases to 4 Å used to calculate anomalous difference Fouriers at all five wavelengths. The bromine and selenium positions were found with the use of these difference Fouriers. Bromine and selenium sites were refined and phases were computed with ML-PHARE in which four of the wavelengths were treated as derivatives and the fifth was treated as parent, with anomalous pairs of the parent wavelength merged. MAD phases were improved by solvent flattening with DPHASES and partial model combination with ML-PHARE. The model was built through use of the program O (46), and positional and simulated annealing refinements were carried out with XPLOR (47).

	$\Delta f'/\Delta f''$ Se	$\Delta f'/\Delta f''$ Br	Crystal	Resolution limit (Å)	R_{sym} (%) [last shell]	% Coverage [last shell]			
<i>Data collection statistics (NSLS X4A)</i>									
$\lambda_1 = 0.9235$ (parent)	-2.0/3.4	-5.18/0.50	1	3.0	4.3[24.7]	97.8[92.7]			
$\lambda_2 = 0.9204$	-1.96/3.4	-9.94/0.50	1	3.0	4.3[23.7]	97.8[91.9]			
$\lambda_3 = 0.9201$	-1.96/3.4	-7.58/3.82	1	3.0	4.3[23.4]	97.5[91.8]			
$\lambda_4 = 0.9797$	-9.79/0.50	2.45/0.56	2	3.0	3.5[36.1]	88.8[71.6]			
$\lambda_5 = 0.9794$	-7.65/3.84	2.45/0.56	2	3.0	3.5[34.0]	88.2[69.7]			
<i>Phasing statistics</i>									
Resolution shell (Å)	10.0	7.5	6.0	5.0	4.3	3.75	3.3	3.0	Overall
†Phasing power (Å)									
Se/Br (λ_2)	0.22	0.29	0.38	0.33	0.39	0.29	0.17	0.11	0.23
Anomalous	0.44	0.58	0.84	0.89	0.77	0.61	0.43	0.36	0.55
Se/Br (λ_3)	0.20	0.23	0.30	0.31	0.37	0.25	0.16	0.10	0.20
Anomalous	0.47	0.63	0.83	0.91	0.84	0.69	0.47	0.41	0.60
Se/Br (λ_4)	0.52	0.86	1.22	1.31	0.91	0.77	0.64	0.60	0.80
Anomalous	0.22	0.29	0.36	0.38	0.33	0.26	0.20	0.14	0.25
Se/Br (λ_5)	0.57	0.81	1.20	1.32	0.85	0.74	0.64	0.57	0.78
Anomalous	0.43	0.54	0.62	0.69	0.63	0.53	0.41	0.27	0.48
Mean figure of merit	0.49	0.55	0.64	0.62	0.49	0.40	0.25	0.24	0.39
<i>Refinement statistics (Se/Br λ_i)</i>									
	Resolution		R factor		Free R factor		N reflections		
Data with $F > 2\sigma$	6.0–3.0		23.2		35.0		12,429		
rms deviations	Bond lengths: 0.011		Bond angles: 1.517		Impropers: 1.380				

* $R_{\text{sym}} = \sum |I_h - \langle I_h \rangle| / \sum I_h$, where $\langle I_h \rangle$ is the average intensity over Friedel and symmetry equivalents. †Isomorphous (iso) phasing power = $\sum |F_{\text{H}}| / \sum |F_{\text{H+obs}}| - |F_{\text{H+calc}}|$; anomalous (anom) phasing power = $\sum |F''_{\text{H}}| / \sum |AD_{\text{obs}}| - |AD_{\text{calc}}|$.

occur between the SSU and TBP, and little conformational change occurs in any of the TBP residues upon TFIIA binding.

TFIIA/DNA interactions. TFIIA interactions with DNA are made by β -sandwich residues in the LSU segment L252-259.

Some are explicit and well ordered, others are less so. L252 makes a hydrogen bond to a phosphate oxygen of G-3, three base pairs upstream of the TATA box, and LR253 makes a hydrogen bond to a phosphate oxygen of T-5', which is within the TATA

box. Although LK255, LR257, and LR259 all point toward the upstream DNA from the opposing face of the β -sandwich and are close enough to make direct contacts, their side chains are disordered. LR253, LK255, and LR259 have all been shown to be important for γ TFIIA/TBP/TATA complex formation (7). It is also necessary to have DNA extended at least six bases upstream of the TATA box to achieve normal binding affinity (28), suggesting that the positive electrostatic potential in this region is important for binding, although variable direct interactions cannot be ruled out.

TBP/TATA box. Bound TFIIA imposes no significant changes in the DNA structure of the complex, either in the TATA region or in flanking duplex segments. The 9 bp of DNA upstream of the TATA box and the two bases downstream of the TATA box all conform to standard B-DNA conformation. There is, however, one major difference in the DNA in this structure compared with that of the γ TBP/CYC1 TATA structure (18)—namely, in the “choice” of TATA boxes. Figure 1, E and F, shows the DNA sequence used in this structure, with the eight base pairs of the TATA box enclosed in red and the DNA used in the binary complex with the eight base pairs of the TATA box of that structure boxed in black. The position of the TATA box—that is, the eight base pairs in contact with TBP—has shifted two base pairs downstream in the ternary complex. This TATA sequence is consistent both with the consensus TATA sequence and with TATA boxes found in nature (29, 30). It differs from the adenovirus major late promoter (AdMLP) by only one base, that of the C in the eighth position, which is a G in the AdMLP. There are three possible explanations for this shift:

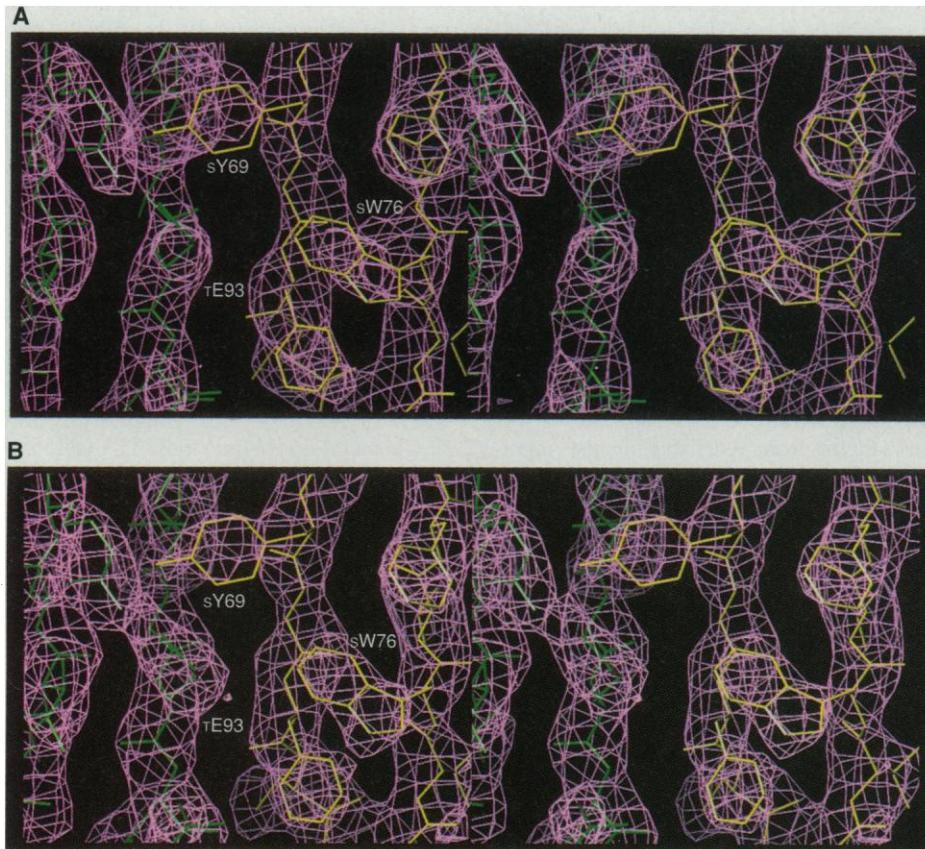


Fig. 2. Stereo view of (A) the experimental, solvent-flattened electron density map phased exclusively by MAD phasing (3.0 Å, 1.3 σ), and (B) the σ_A weighted (49) $2F_o - |F_c|$ electron density map (3.0 Å, 1.3 σ) calculated with the final refined coordinates. Shown are TBP residues (green) and TFIIA residues (yellow). Residues of TBP are labeled according to the format established for TFIIA (25), that is, TBP R-105 corresponds to TR105.

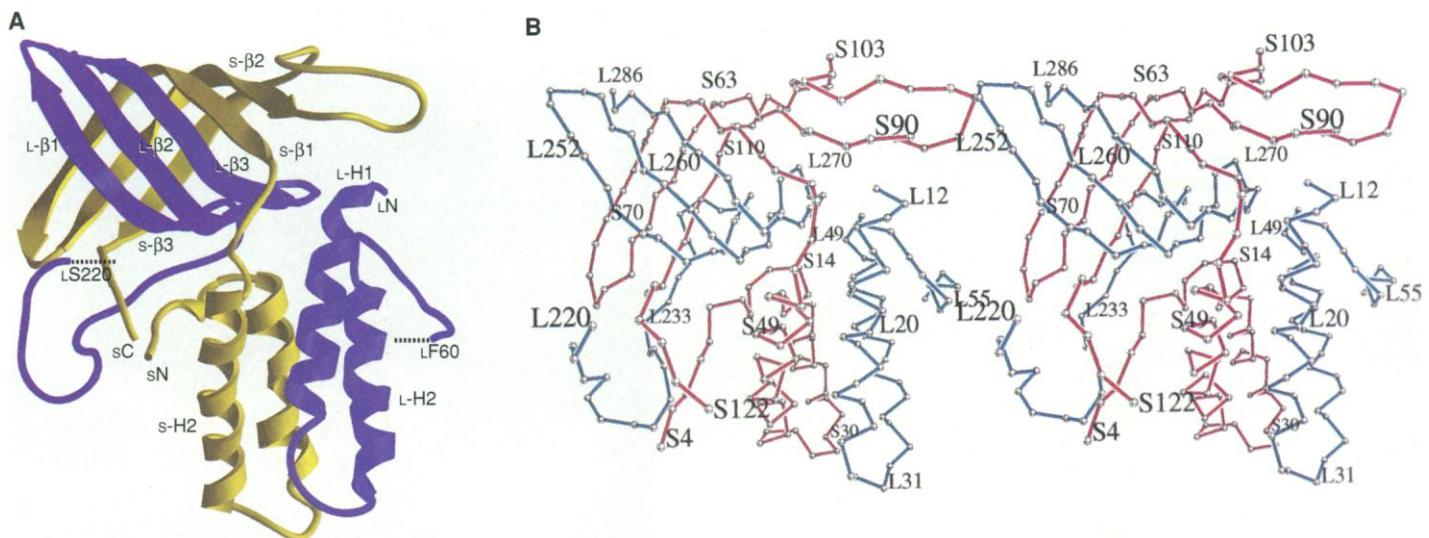


Fig. 3. Structure of γ TFIIA. (A) Ribbon (50) drawing of TFIIA with elements of secondary structure labeled. Shown are the SSU (yellow) and the LSU (blue). Dotted lines here and in the following figures represent the points at which the

electron density for the LSU is not interpretable. (B) Stereo view of the $C\alpha$ backbone of the two subunits of TFIIA with residues labeled. The LSU (blue) and the SSU (red) are indicated.

(i) The choice of TATA box has been affected by the interaction with TFIIA; (ii) the choice of TATA box is affected by crystal packing; and (iii) the choice of TATA box was affected in the γ TBPc/TATA box structure by the presence of the hairpin loop on the downstream side.

It is unlikely that the shift is due to γ TFIIA because the interactions between TBP and DNA are unchanged from those made in the binary complex, and little change is seen in the structure of the DNA or TBP. Crystal packing is an unlikely explanation as well, because both ends of the DNA are free to make crystal contacts no matter which TATA box is chosen. The presence of the hairpin loop in the original structure confers on the adjacent base pair a conformation that is distorted from B-DNA—or from TA-DNA, the DNA conformation seen in the bound TATA boxes (17–19, 31). If TBP were to bind to the hairpin construct of the binary TBP/TATA structure in the manner seen in the present ternary complex structure, the contact surface would include this distorted

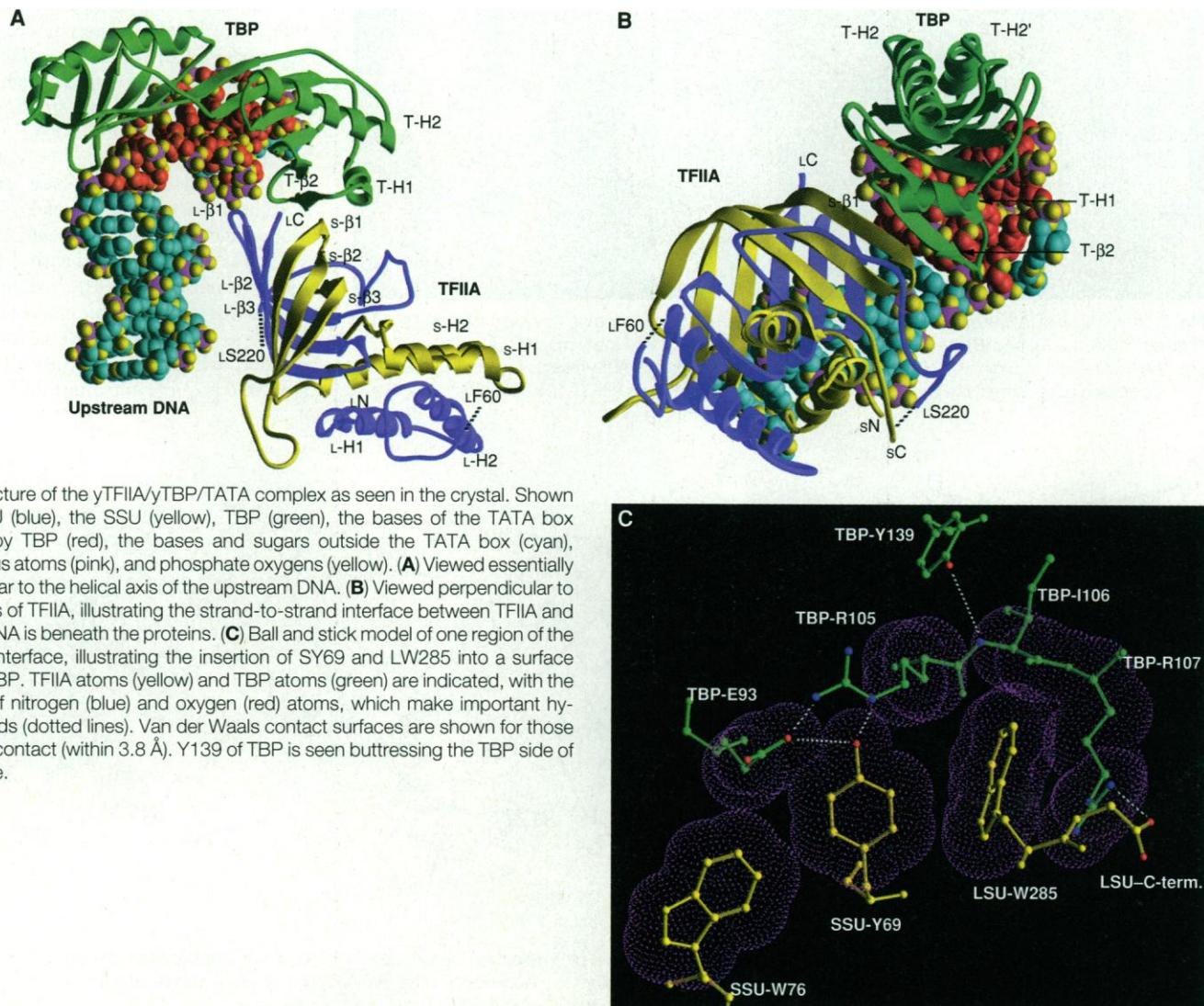
base pair as the eighth base pair in the TATA box. Thus, in the original structure of the binary complex, this hairpin-induced distortion may have forced TBP to bind to a slightly less favorable alternate choice of TATA box. Hydroxyl radical footprinting of both TBP and TBP/TFIIA/CYCl TATA complexes are ambiguous, but indicate that both TATA boxes may be occupied in solution (28).

TFIIA binds to the TBP/promoter complex by a combination of stereospecific and electrostatic interactions that exploit the unusual conformation of the TBP/promoter complex. There is no induced fit because the conformation of TBP and the TATA box is unchanged relative to the binary complex structure, which serves as a nucleoprotein scaffold upon which TFIIA and TFIIIB bind with high affinity and stereospecificity.

Mutagenesis of TFIIA/TBP interfaces. Almost all of the residues defining the interface between γ TBP and γ TFIIA in the present structure are conserved among species ranging from yeast to humans; importantly, the Y65A variant in the γ subunit of

hTFIIA (SY69A in γ TFIIA) prevents formation of the ternary complex and transcriptional activity of hTFIIA (32). Berk and co-workers have made drastic mutations to 89 of the surface residues of hTBP and assayed them for hTFIIA binding and transcriptional activation in vivo (33). Four changes—R107E, N91E, E93R, and A86E (yeast numbering is used)—specifically prevented formation of the TFIIA/TBP/TATA complex and were deficient for activated transcription in vivo. All of these residues were shown to interact directly with TFIIA in our structure (Fig. 5). Taken together, these results support the assertion that hTFIIA, like γ TFIIA, requires the interface depicted in Figs. 4 and 5 for binding.

Mutagenesis studies have also implicated direct involvement of the H2 helix of TBP in the interaction with hTFIIA (27, 34–36). Specifically, mutations of γ TBP residues K138L or K145L abrogated the formation of the γ TBP/hTFIIA/TATA complex (35, 36). Single mutations of human R235A (R135 in γ TBP) or D228A (D130 in γ TBP) decreased



binding of hTFIIA 10- and 20-fold, respectively (34). Our structure shows no direct contacts between γ TFIIA and the H2 helix of γ TBP—the H2 helix is 13.5 Å away from TFIIA at its closest approach. Consistent with our structure, γ TFIIA binds to two of

these variants, the only two so far tested, with an affinity similar to that of wild-type (WT) TBP (35), although the gel mobility of the complex is altered. This discrepancy suggests that hTFIIA complexes may have an additional and essential binding interface

that includes the basic residues in the H2 helix of TBP. What portion of hTFIIA might be making interactions with the H2 helix? Only one sequence segment is conserved in *Drosophila* and hTFIIA that is not conserved in yeast; it is a region located just amino-terminal to L220 in our structure. In both the *Drosophila* and human proteins this segment is almost entirely acidic, whereas the equivalent segment of yeast (residues L190 to L219) is mostly basic. These 30 acidic residues of *Drosophila* and hTFIIA could easily reach the H2 helix and make important interactions with the H2 helix of TBP.

TFIIA and DNA. γ TFIIA makes fewer contacts with DNA than it does with TBP, and no direct contacts are made between TFIIA and the DNA downstream of the TATA box. This finding is consistent with the hydroxyl radical footprint of the TBP/TATA complex, which is unchanged upon binding of TFIIA, and with the observation that DNA sequences that are truncated to within six bases of the TATA box do not affect TFIIA binding (28). Presumably, the sparseness of fixed direct contacts to the phosphates does not impede access of hydroxyl radicals to the minor groove. In contrast, TFIIA extends the footprint 5' to the TATA box when performed with the bulkier deoxyribonuclease I reagent (3). Results from DNA cross-linking of the yeast TFIIA/TBP/promoter complex are also in agreement with our structure, showing cross-links of the LSU to DNA up to 9 bp upstream of the TATA box (37) and 2 bp downstream of the TATA box (38). Residues of the β -sheet region of the LSU are within 10 Å (the length of the cross-linking

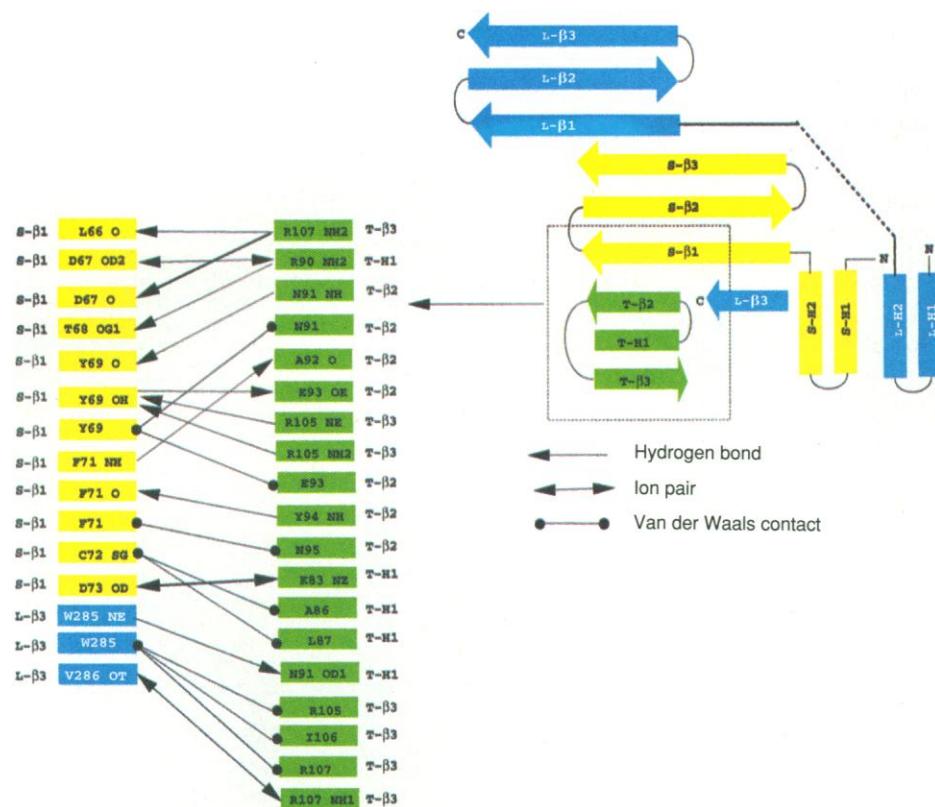


Fig. 5. Diagram of the specific interactions between TFIIA and TBP seen in the crystal. The γ TFIIA SSU (yellow), the LSU (blue), and TBP (green) are shown.

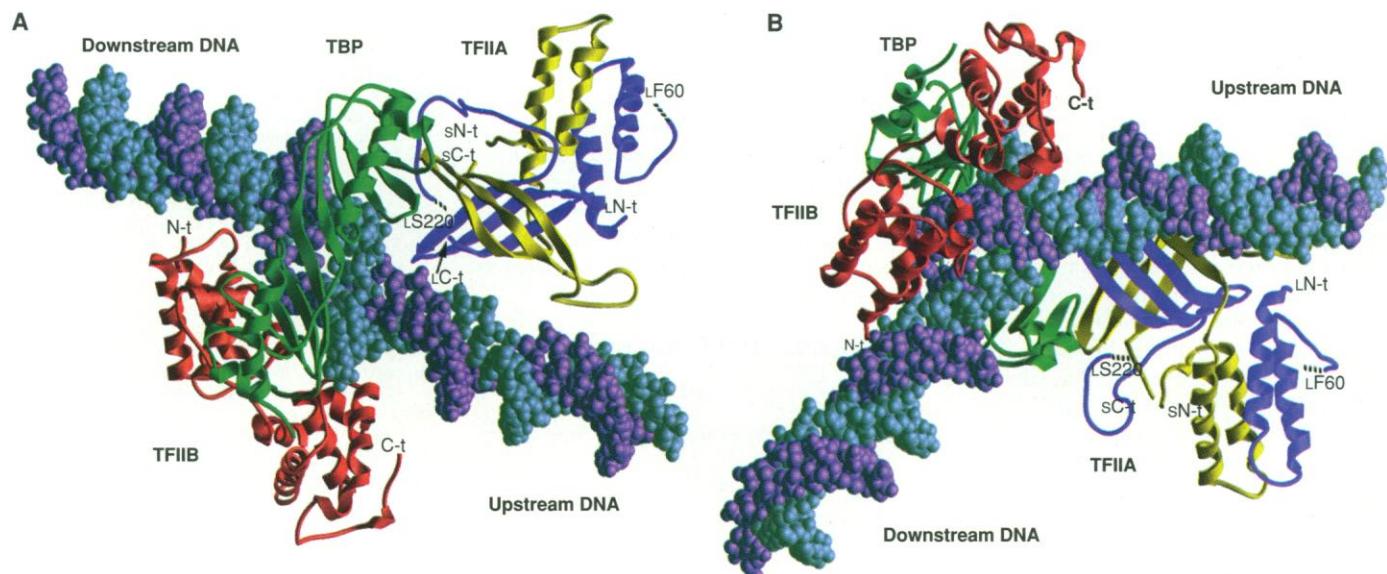


Fig. 6. Presumed model of the TBP/TFIIA/TFIIBc/promoter complex assembled computationally with the crystal structures of the TFIIBc/TBP/TATA box and the present structure of the TFIIA/TBP/TATA box (57). TFIIB (red), TBP (green), the SSU (yellow), and the LSU (blue) are indicated. Model B-form DNA is extended 20 bp in the downstream direction and 12 bp in the

upstream direction from the oligonucleotide used in the crystallization and is depicted as a space-filling model. Proteins are ribbons depictions. **(A)** View is down the pseudotwofold of TBP, approximately perpendicular to the H2 and H2' helices. **(B)** The underside of the complex rotated $\sim 180^\circ$ relative to **(A)**.

reagent) of the upstream phosphates that bear the cross-linking reagent. Cross-links with derivatized phosphate groups downstream of the TATA box could conceivably be made with the flexible segment carboxyl-terminal to L220 (13 Å from the DNA backbone). In contrast, cross-linking studies of hTFIIA indicate that the α subunit (the amino-terminus of the LSU in γ TFIIA) can interact with DNA up to 17 bp upstream of the TATA box (37, 38). These base pairs are far out of range of γ TFIIA seen in our structure. The presence of nonconserved sequence segments in this subunit makes this result difficult to interpret in the absence of a hTFIIA ternary complex structure.

Special relation to TFIIB. Unlike TFIIB (20), which interacts both upstream and downstream of the TATA box, TFIIA is located upstream of the TBP/TATA complex. Consistent with the *in vivo* and *in vitro* data, it is unlikely that TFIIA makes functionally important interactions with any of the basal machinery located downstream of the TATA box. Figure 6 depicts the complex with modeled B-DNA extending upstream and downstream from the DNA in the crystal (20). By overlaying the TBP in the hTFIIB/TBP/TATA complex with the TBP in our complex, we have modeled the TFIIB/TFIIA/TBP/TATA quaternary complex (20). As shown in Fig. 6, A and B, TFIIB is located on the opposite side of the DNA from TFIIA, and there appears to be no direct interactions between the proteins in the complex, at least in the regions of each protein seen in the structures. Even when both TFIIA and TFIIB are bound, the entire top side of TBP is free to make interactions with other factors. Indeed, only a small portion of the surface area of TFIIA is used in the interaction with the TBP/TATA complex, leaving the majority of all the proteins available for interaction with other factors.

TFIIA has also been shown to abrogate the effects of some transcriptional repressors—namely, Dr2 (topoisomerase I) (39), MOT1 (40), NC1 (41), and HMG1 (42)—suggesting that TFIIA may sterically interfere with the interaction between the repressors and the TFIID/promoter complex. This possibility seems likely for MOT1, a repressor of transcription that acts by prying TBP from the DNA by an adenosine 5'-triphosphate-dependent mechanism. If MOT1 must interact with both TBP and the upstream promoter region of DNA, it is possible that the β -sandwich domain prevents MOT1 from binding to its nucleoprotein substrate.

The four-helix-bundle domain. The four-helix-bundle domain is not involved in the interaction of γ TFIIA with TBP or DNA. However, deletion of any 10 amino acids in this domain (7) disrupts subunit association, eliminates *in vitro* activity of

γ TFIIA, and produces nonviable yeast *in vivo*, probably because these variants are unable to fold properly. Alanine-scanning mutagenesis has identified two temperature-sensitive mutants in the four-helix-bundle domain: the double-variant SD21A, SD24A; and the triple variant SD29A, SR31A, SE33A (7). The first variant is significantly defective in both *in vitro* basal and activated transcription, although as expected from the structure, it has wild-type binding affinity for the TBP/TATA complex. All of these amino acid changes are found near the bottom of the four-helix-bundle domain, implicating this domain as important for the transcriptional activity of TFIIA both *in vivo* and *in vitro*.

hTFIIA that lacks the entire α subunit is stable and active in *in vitro* activator-independent transcription assays and still retains its antirepressor activity with regard to topoisomerase I, indicating that deletion of an entire half of the four-helix-bundle domain produces at least some part of a properly folded molecule (43). α Subunit-depleted TFIIA was, however, deficient in transcriptional activation in response to GAL4-VP16 *in vitro*, further implicating the helical domain in transcriptional regulation. The four-helix-bundle domain extends in a direction perpendicular to the DNA helical axis and resembles a handle attached to the side of the complex (Fig. 6B).

Electrostatics calculations of TFIIA indicate that, with the notable exception of the surface facing the upstream DNA, virtually all of the exposed surface of TFIIA in the complex is negatively charged, especially that of the four-helix-bundle domain, indicating that TFIIA is most likely to interact with the positive surfaces of other factors.

In summary, the two subunits of TFIIA associate intimately to produce both domains of a previously uncharacterized two-domain protein that associates stereospecifically with the TBP/promoter nucleoprotein complex. The specific contacts with TBP and DNA are modest, leaving a generous surface on both TFIIA and TBP available for interaction with the multitude of factors involved in the transcription initiation process.

REFERENCES AND NOTES

1. E. Maldonado and D. Reinberg, *Curr. Opin. Cell Biol.* **7**, 352 (1995).
2. J. A. Ranish and S. Hahn, *Curr. Opin. Genet. Dev.*, in press.
3. S. Buratowski *et al.*, *Cell* **56**, 549 (1989).
4. X. Sun, D. Ma, M. Sheldon, K. Yeung, D. Reinberg, *Genes Dev.* **8**, 2336 (1994).
5. K. Yokomori *et al.*, *ibid.*, p. 2313.
6. J. Ozer *et al.*, *ibid.*, p. 2324.
7. J. J. Kang, D. T. Auble, J. A. Ranish, S. Hahn, *Mol. Cell Biol.* **15**, 1234 (1995).
8. W. Wang *et al.*, *Genes Dev.* **6**, 1716 (1992).
9. T. Chi, P. Lieberman, K. Ellwood, M. Carey, *Nature* **377**, 254 (1995).
10. N. Kobayashi *et al.*, *Mol. Cell Biol.* **15**, 6465 (1995).
11. J. A. Ranish *et al.*, *Science* **255**, 1127 (1992).

12. J. A. Ranish and S. Hahn, *J. Biol. Chem.* **266**, 19320 (1991).
13. S. Hahn *et al.*, *EMBO J.* **8**, 3379 (1989).
14. J. DeJong and R. G. Roeder, *Genes Dev.* **7**, 2220 (1993).
15. D. Ma *et al.*, *ibid.*, p. 2246.
16. K. Yokomori *et al.*, *ibid.*, p. 2235.
17. J. L. Kim *et al.*, *Nature* **365**, 520 (1993).
18. Y. Kim *et al.*, *ibid.*, p. 512.
19. J. L. Kim and S. K. Burley, *Nature Struct. Biol.* **1**, 638 (1994).
20. D. B. Nikolov *et al.*, *Nature* **377**, 119 (1995).
21. J. Ranish, personal communication.
22. W. A. Hendrickson *et al.*, *Proteins Struct. Funct. Genet.* **4**, 77 (1988).
23. V. Ramakrishnan *et al.*, *Nature* **362**, 219 (1993).
24. See Fig. 1, A and B, for the identification of the elements of secondary structure of TFIIA. Elements of TBP secondary structure have been defined previously (23).
25. LF60 refers to residue F60 of the LSU; similarly, SK119 refers to residue K119 in the SSU. This convention will be maintained throughout this paper.
26. If TBP is viewed perpendicular to its pseudotwofold in the familiar saddle view, the DNA will be projecting out at the viewer on one side, and into the page on the other. The face of TBP from which the upstream DNA emanates is the upstream face, and the other face is the downstream face of TBP.
27. L. A. Stargell and K. Struhl, *Science* **269**, 75 (1995).
28. J. H. Geiger, S. Hahn, S. Lee, P. B. Sigler, data not shown.
29. F. Nagawa and G. R. Fink, *Proc. Natl. Acad. Sci. U.S.A.* **82**, 8557 (1985).
30. S. Hahn, E. T. Hoar, L. Guarente, *ibid.*, p. 8562.
31. G. Guzikovich-Guerstein and Z. Shakked, *Nature Struct. Biol.* **3**, 32 (1996).
32. J. Ozer *et al.*, *J. Biol. Chem.* **271**, 11182 (1996).
33. G. O. Bryant, L. Martel, S. K. Burley, A. Berk, personal communication.
34. H. Tang, X. Sun, D. Reinberg, R. H. Ebright, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 1119 (1996).
35. S. Buratowski and H. Zhou, *Science* **255**, 1130 (1992).
36. D. K. Lee *et al.*, *Mol. Cell Biol.* **12**, 96 (1992).
37. B. Coulombe, J. Li, J. Greenblatt, *J. Biol. Chem.* **269**, 19962 (1994).
38. T. Lagrange, T. K. Kim, G. Orphanides, R. Ebright, D. Reinberg, personal communication.
39. A. Merino *et al.*, *Nature* **365**, 227 (1993).
40. D. T. Auble *et al.*, *Genes Dev.* **8**, 1920 (1994).
41. M. Meisterernst *et al.*, *Cell* **66**, 981 (1991).
42. H. Ge and R. G. Roeder, *J. Biol. Chem.* **269**, 17136 (1994).
43. D. Ma *et al.*, personal communication.
44. J. Ranish and S. Hahn, unpublished results.
45. J. Navaza, *Acta Crystallogr. Sect. A* **50**, 157 (1994).
46. T. A. Jones *et al.*, *ibid.* **47**, 110 (1991).
47. A. T. Brunger, *XPLOR Version 3.1 Manual* (Yale University, New Haven, 1993).
48. G. D. Schuler, S. F. Altschul, D. J. Lipman, *Proteins Struct. Funct. Genet.* **9**, 180 (1991).
49. R. Reed, *Acta Crystallogr. Sect. A* **42**, 140 (1986).
50. M. Carson, *J. Appl. Crystallogr.* **24**, 958 (1991).
51. The TFIIBc/TBP/TFIIA/TATA quaternary complex was modeled by overlaying the structure of *Arabidopsis* TBP in the TFIIBc/TBP/TATA ternary complex onto the structure of γ TBP in the γ TBP/ γ TFIIA/TATA complex and by using the resulting matrix to translate TFIIBc into the γ TBP/ γ TFIIA/TATA complex. TFIIBc refers to the amino-terminally deleted TFIIB used in the crystal structure.
52. We thank G. van Duyn (DPHASES), Z. Otwinowski (DENZO and SCALEPACK), G. Meinke, and S. Williams for assistance during data collection; C. Ogata for help at the X4A beamline; T. Lagrange, D. Reinberg, R. Ebright, P. Lieberman, and J. Ozer for communicating results before publication; A. Bohm, G. Ghosh, and B. De Decker for helpful discussions; and D. Nikolov and S. Burley for providing the coordinates of the TFIIBc/TBP/TATA box complex. Supported by grants from the National Institutes of Health to P.E.S. and S.H. J.H.G. was an American Cancer Society Fellow, and S.H. is a Leukemia Society scholar.

5 April 1996; accepted 22 April 1996