Cell, Vol. 94, 573–583, September 4, 1998, Copyright ©1998 by Cell Press

Solution Structure of a TBP–$TAF_{II}230$ Complex: Protein Mimicry of the Minor Groove Surface of the TATA Box Unwound by TBP

Dingjiang Liu,* Rieko Ishima,*‖ Kit I. Tong,*
Stefan Bagby,* Tetsuro Kokubo,†#
D. R. Muhandiram,‡ Lewis E. Kay,‡
Yoshihiro Nakatani,† and Mitsuhiko Ikura*§
*Division of Molecular and Structural Biology
Ontario Cancer Institute
Department of Medical Biophysics
University of Toronto
Toronto, Ontario M5G 2M9
Canada
†Laboratory of Molecular Growth Regulation
National Institute of Child Health and Human Development
National Institutes of Health
Bethesda, Maryland 20892
‡Protein Engineering Network Centres of Excellence
Departments of Molecular and Medical Genetics, Biochemistry, and Chemistry
University of Toronto
Toronto, Ontario M5S 1A8
Canada

Summary

General transcription factor TFIID consists of TATA box–binding protein (TBP) and TBP-associated factors (TAF_{II}s), which together play a central role in both positive and negative regulation of transcription. The N-terminal region of the 230 kDa *Drosophila* TAF_{II} ($dTAF_{II}230$) binds directly to TBP and inhibits TBP binding to the TATA box. We report here the solution structure of the complex formed by $dTAF_{II}230$ N-terminal region (residues 11–77) and TBP. $dTAF_{II}230_{11-77}$ comprises three α helices and a β hairpin, forming a core that occupies the concave DNA-binding surface of TBP. The TBP-binding surface of $dTAF_{II}230$ markedly resembles the minor groove surface of the partially unwound TATA box in the TBP–TATA complex. This protein mimicry of the TATA element surface provides the structural basis of the mechanism by which $dTAF_{II}230$ negatively controls the TATA box–binding activity within the TFIID complex.

Introduction

Transcription initiation of protein-encoding genes requires assembly of a multiprotein preinitiation complex (PIC) on TATA-containing core promoters that consists of RNA polymerase II and at least six general transcription factors (GTFs) designated TFIIA, -IIB, -IID, -IIE, -IIF, and -IIH (reviewed in Orphanides et al., 1996; Roeder, 1996). The TBP subunit of TFIID recognizes and binds the TATA element of the core promoter in a sequence-specific manner to nucleate PIC assembly. This process is stimulated by TFIIA. TFIIB then enters the complex through interaction with TBP and DNA and creates a platform for binding of the RNA polymerase II–TFIIF complex. Subsequent recruitment of TFIIE and TFIIH completes PIC formation. More recent studies, however, have shown that RNA polymerase II, a subset of GTFs, and coactivators may enter the PIC as a preassembled unit (reviewed in Orphanides et al., 1996). Although the assembly pathway most relevant to the in vivo situation remains unclear, binding of the core promoter by TFIID is a crucial point for regulation of transcription: numerous transcription activators recruit TFIID via direct interactions or perhaps induce conformational changes in TFIID that affect recruitment or function of other factors (reviewed in Burley and Roeder, 1996; Orphanides et al., 1996; Roeder, 1996). Conversely, TFIID is also targeted by negative regulators to repress transcription. For instance, negative cofactor NC2, also known as the Dr1–DRAP1 complex, binds to TBP-bound promoters and inhibits entry of TFIIA and TFIIB (reviewed in Kaiser and Meisterernst, 1996).

The TAF_{II}s themselves are active participants in transcriptional regulation (reviewed in Burley and Roeder, 1996; Roeder, 1996). While TAF_{II}s may function as positive cofactors, at least in part by providing interaction sites for activators, TAF_{II}s also act as negative cofactors in some contexts by preventing and/or destabilizing TFIID interaction with the core promoter. An early study with partially purified TFIID demonstrated that TFIID binds to a weak core promoter less efficiently than TBP alone (Nakatani et al., 1990). Later experiments with highly purified TFIID indicated that TFIID binds poorly to a variety of core promoters lacking strong initiator sequences (Aso et al., 1994). Studies with a cell-free system lacking TAF_{II}s revealed that TAF_{II}s impair functional PIC formation (Oelgeschläger et al., 1998). An investigation employing partially reconstituted TFIID complexes showed that human (h) $TAF_{II}250$ (reviewed in Burley and Roeder, 1996) represses the basal transcription activity of TBP (Verrijzer et al, 1995; Guermah et al., 1998). In agreement with these results, we found that $hTAF_{II}250$ (T. K., unpublished results) and homologous $dTAF_{II}230$ (also known as $dTAF_{II}250$) and yeast (y) $TAF_{II}145$ (also known as $yTAF_{II}130$) (reviewed in Burley and Roeder, 1996) inhibit TBP binding to the TATA box when two recombinant polypeptides are mixed in vitro (Kokubo et al., 1993, 1998). The N-terminal 156 residues of $dTAF_{II}230$ (Figure 1) inhibit TATA box binding through direct interaction with TBP (Nishikawa et al., 1997), suggesting that the TBP affinity for the $dTAF_{II}230$ N-terminal region is comparable to that for the TATA box. The K_d for the TBP–TATA interaction is 4 nM (Perez-Howard et al., 1995). The $dTAF_{II}230$ N-terminal region has been dissected into subdomains I (residues 11–77) and II (residues 82–156), which are thought to bind to the concave and convex surfaces of TBP (Kokubo et al., 1994; Nishikawa et al., 1997). In addition to their inhibitory function,

§To whom correspondence should be addressed.
‖Present address: Molecular Structural Biology Unit, National Institute of Dental Research, National Institutes of Health, Bethesda, Maryland 20892.
#Present address: Division of Gene Function in Animals, Nara Institute of Science and Technology, Ikoma, Nara 630-01, Japan.

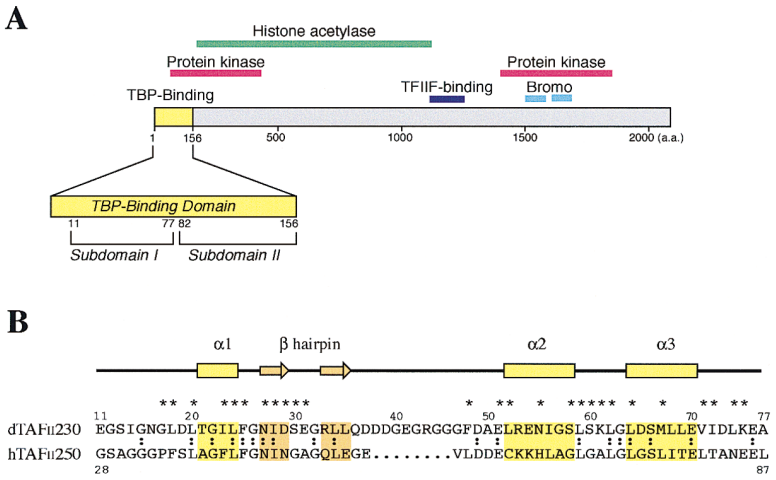

Figure 1. Features and Sequence Alignment of $TAF_{II}230$

(A) Schematic depiction of full-length *Drosophila* $TAF_{II}230$. Proposed domain locations and functions are shown.

(B) Sequence alignment of the N-terminal TBP-binding domain (residues 11–77) of *Drosophila* $TAF_{II}230$ with the corresponding region of human $TAF_{II}250$. Secondary structure elements determined in the present study are indicated with yellow boxes for α helices and orange arrows for β strands of the β hairpin. Residues conserved between the human and *Drosophila* proteins are connected by two dots, and residues involved in the TBP interaction are marked with an asterisk.

$hTAF_{II}250$ and $dTAF_{II}230$ possess enzymatic activities: a histone acetylase activity (Mizzen et al., 1996) that is also found in some coactivators (for example, Ogryzko et al., 1996; Yang et al., 1996) and could be important in transcription activation in the context of chromatin, and protein kinase activity (Dikstein et al., 1996) that may be required to direct transcription of some genes in vivo (O'Brien and Tjian, 1998).

Structural studies of GTFs have contributed greatly to our understanding of the macromolecular interactions supporting assembly of the transcription PIC. The crystal structures of TBP (Nikolov et al., 1992; Chasman et al., 1993) and the TBP–TATA box complex (Kim et al., 1993a, 1993b) revealed a large concave undersurface on TBP formed by a curved, ten-stranded, antiparallel β sheet that makes minor groove and phosphate–ribose contacts with a partially unwound form of the 8 base pair TATA element. Further details of PIC assembly have been revealed by the crystal structures of two ternary complexes, the TFIIB–TBP–TATA complex (Nikolov et al., 1995; Kosa et al., 1997) and the TFIIA–TBP–TATA complex (Geiger et al., 1996; Tan et al., 1996). We report here the three-dimensional structure in solution of the complex comprising residues 11–77 of *Drosophila* $TAF_{II}230$ (hereafter referred to as $dTAF_{II}230_{11–77}$) and the core domain (residues 49–240) of *Saccharomyces cerevisiae* (*S. cerevisiae*) TATA-binding protein (hereafter referred to as TBP), representing novel structural information on negative regulation of the transcription initiation machinery. This study establishes the structural basis for the TBP–$dTAF_{II}230$ interaction and provides insights into the mechanism for negative regulation of transcription within the TFIID multisubunit protein complex, including an intriguing observation of protein mimicry of DNA structure.

Results and Discussion

Structure Determination

We decided to pursue structural analysis of a mixed species complex (*S. cerevisiae* TBP and *Drosophila* TAF) for the following reasons: the TBP core sequence is highly conserved between species (80% sequence identity between *S. cerevisiae* and *Drosophila*), so the TBP core structures will be essentially identical; $dTAF_{II}230_{11–77}$ has similar affinities for *S. cerevisiae*, *Drosophila*, and human TBPs (T. K. and Y. N., unpublished results); *Drosophila* TBP expression was poor; and $dTAF_{II}230_{11–77}$ and *S. cerevisiae* TBP produced a stable complex. The first two of these facts indicate that the *Drosophila/S. cerevisiae* complex is physiologically relevant.

The ^{1}H-^{15}N heteronuclear single quantum coherence (HSQC) spectrum of free TBP displays many broadened peaks (data not shown), presumably due to the formation of a dimer or multimer in solution (Coleman et al., 1995; Perez-Howard et al., 1995). In contrast, the HSQC spectrum of ^{15}N-labeled TBP complexed with unlabeled $dTAF_{II}230_{11–77}$ is highly dispersed (Figure 2A). This is consistent with the gel filtration behavior of the complex that indicates an apparent molecular weight of 28–30 kDa (data not shown), corresponding to a 1:1 complex of TBP (M_r 21.3 kDa) and $dTAF_{II}230_{11–77}$ (M_r 6.8 kDa).

Agreement between the observed and expected numbers of backbone NH cross peaks and excellent peak dispersion in ^{1}H-^{15}N HSQC spectra of ^{15}N-labeled $dTAF_{II}230_{11–77}$ bound to unlabeled TBP (Figure 2B) indicate that $dTAF_{II}230_{11–77}$ adopts a single folded conformation. On the other hand, ^{1}H-^{15}N HSQC and CD spectra of free $dTAF_{II}230_{11–77}$ indicate that $dTAF_{II}230_{11–77}$ by itself is largely unfolded (Figure 2C). During a titration experiment in which ^{1}H-^{15}N HSQC spectra of $dTAF_{II}230_{11–77}$ were recorded with successive additions of TBP, most peaks corresponding to the unfolded state decrease in intensity and appear at new positions corresponding to the folded state, consistent with the high affinity of $dTAF_{II}230_{11–77}$ for TBP (K_d approximately 10^{-9} M). The excellent dispersion in the ^{1}H-^{15}N HSQC and other NMR spectra of both TBP and $dTAF_{II}230_{11–77}$ in the binary complex permitted structure determination of both polypeptides (Figure 3 and Table 1).

$dTAF_{II}230_{11–77}$ Structure

In the complex with TBP, $dTAF_{II}230_{11–77}$ adopts an elongated conformation comprising a core domain and a long loop (Figure 3), with approximate maximal dimensions of 40 Å × 12 Å × 24 Å (25 Å × 12 Å × 24 Å if the loop is omitted). The pairwise root–mean–square (rms) deviation of the NMR-derived structures is 0.70 Å for

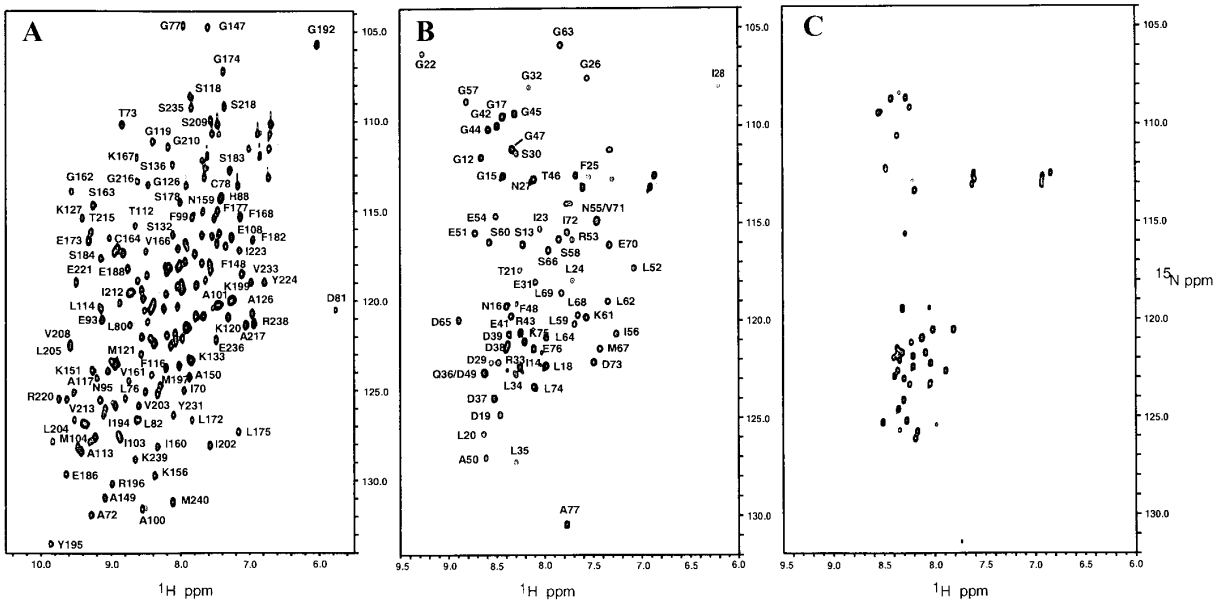

Figure 2. Backbone Fingerprint NMR Spectra of the TBP–dTAF$_{II}$230$_{11–77}$ Complex

(A) ^{1}H-^{15}N HSQC spectrum of ^{15}N-labeled TBP in complex with unlabeled dTAF$_{II}$230$_{11–77}$.

(B) ^{1}H-^{15}N HSQC spectrum of ^{15}N-labeled dTAF$_{II}$230$_{11–77}$ in complex with unlabeled TBP.

(C) ^{1}H-^{15}N HSQC spectrum of ^{15}N-labeled dTAF$_{II}$230$_{11–77}$ alone.

In (A) and (B), peaks are labeled (where peak separation allows) using the one-letter amino acid code and sequence position of the corresponding residue.

backbone atoms and 1.09 Å for heavy atoms (N-terminal residues 11–16 and C-terminal residues 75–77, which are unstructured, were not included in this calculation). The core domain contains three α helices (α1, residues 21–24; α2, 52–58; α3, 64–70) and a β hairpin (residues 27–35), with Ser-30, Glu-31, and Gly-32 forming the hairpin loop. Hydrophobic helix α1 lies at the center of the dTAF$_{II}$230$_{11–77}$ fold (Figure 4). The β hairpin packs against helix α1 via hydrophobic contacts between Ile-28 and Gly-22 and Ile-23. The hairpin loop region contacts Leu-18 to the N-terminal side of helix α1, and Leu-34 contacts Phe-25 just C-terminal of helix α1. Hydrophobic contacts between helices α1 and α3 are made by Thr-21 and Leu-24 with Leu-64 and Leu-68. These are supplemented by Leu-18 and Leu-20 contacts with Val-70 and Ile-72. Gly-22 in helix α1 makes hydrophobic contacts with Ile-56 at the center of helix α2. Conservation of amino acid sequence (34% identity and 66% similarity) between the N-terminal TBP-binding regions of dTAF$_{II}$230 (residues 11–77) and hTAF$_{II}$250 (residues 28–87) suggests that this region of hTAF$_{II}$250 adopts a similar fold to dTAF$_{II}$230 when complexed with TBP. A search of the Protein Data Bank using the DALI (Holm and Sander, 1993) and SSAP (Orengo et al., 1992) algorithms for protein structures similar to that of dTAF$_{II}$230$_{11–77}$ identified no structures with a Z score greater than 0.7 (DALI) or a score above 77 (SSAP).

dTAF$_{II}$230$_{11–77}$ contains the sequence LFG between helix α1 and the β hairpin (residues 24–26). Such a sequence is found in all members of the Kip/Cip family of cyclin-dependent kinase inhibitors (Russo et al., 1996). The LFG motif is also conserved between hTAF$_{II}$250 and dTAF$_{II}$230 (Figure 1B) (in yTAF$_{II}$145, Leu is replaced by Ile), suggesting that it is structurally significant. Indeed, by virtue of its extended conformation, the LFG motif in dTAF$_{II}$230$_{11–77}$ plays a key role in the formation of a large hydrophobic patch (Figure 4) that is responsible for the interaction with TBP. Phe-25 in the LFG motif lies at the center of the extensive hydrophobic patch and makes numerous contacts with hydrophobic residues of TBP (see below). In the crystal structure of p27^{Kip1} bound to the cyclin A–Cdk2 complex (Russo et al., 1996), the LFG sequence also assumes a partially extended conformation and binds a shallow hydrophobic groove of cyclin A. The LFG motif is probably adopted to provide an exposed hydrophobic cluster available for interaction with a partner protein.

Other residues involved in the exposed hydrophobic patch include Ile-28 (β hairpin), Leu-34 (β hairpin), Phe-48, Leu-52 (helix α2), Ile-56 (helix α2), Leu-59, and Leu-64 (helix α3). The hydrophobic patch also includes an interlocking zipper-like arrangement between Leu-18, Leu-20, Ile-23, Leu-24 and Leu-74, Ile-72, Val-71, Leu-68 (Figure 4). The β-hairpin sheet is arranged such that the polar and charged side chains of Asn-27 and Asp-29 lie on the same molecular face as the hydrophobic patch. Consequently, with respect to the central helix α1, one side of the hydrophobic patch is more hydrophilic than the other side. The significance of this asymmetry is discussed in the TBP–dTAF$_{II}$230 Interface section.

TBP Structure

The solution structure of TBP bound to dTAF$_{II}$230$_{11–77}$ is essentially identical to the saddle-like crystal structure of TBP in the TBP–DNA binary complex (Kim et al.,

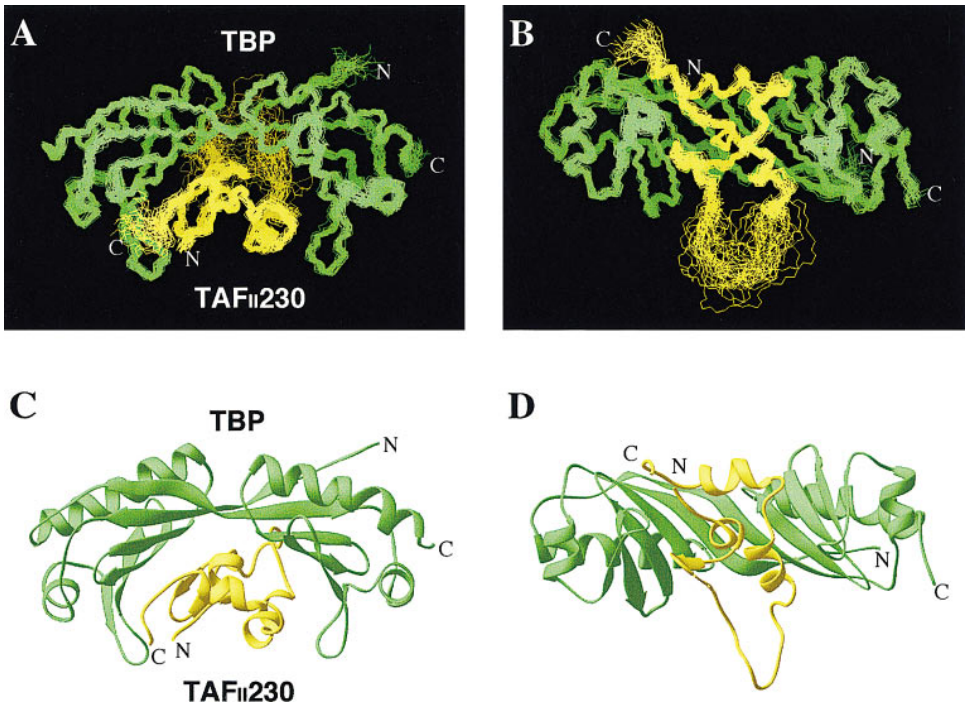

Figure 3. Best-Fit Superposition of the Family of NMR-Derived Structures and Polypeptide Fold of the TBP–dTAF$_{II}$230$_{11–77}$ Complex

(A) Best-fit superposition of 34 NMR-derived structures. Only backbone N, Cα, and C′ atoms are displayed. TBP is shown in green and dTAF$_{II}$230$_{11–77}$ in yellow. Residues 11–16 have been omitted.

(B) A different view of the family of NMR-derived structures generated from the view in (A) by rotating through 90° along the horizontal axis. These figures were produced using Insight II.

(C and D) Ribbon representation of the energy-minimized average structure of the TBP–dTAF$_{II}$230$_{11–77}$ complex. The color schemes and orientations of the complex are the same as in (A) and (B). These figures were generated using Ribbons (Carson, 1991).

1993a, 1993b). The atomic rms difference between *S. cerevisiae* TBP in complex with dTAF$_{II}$230$_{11–77}$ and with the TATA element (Kim et al., 1993a) is 1.4 ± 0.2 Å when backbone atoms of residues 64–236 are superimposed. The corresponding rms difference between the dTAF$_{II}$ 230$_{11–77}$-bound and free TBP (Chasman et al., 1993)

Table 1. Structural Statistics

Rms deviations from experimental distance restraints (Å)	
All (3956)	0.031 ± 0.003
Interresidue sequential (|i − j| = 1)(557/248)	0.039 ± 0.006
Interresidue short-range (1 < |i − j| ≤ 5)(460/193)	0.055 ± 0.006
Interresidue long-range (|i − j| ≥ 5)(1097/185)	0.012 ± 0.002
Intraresidue (443/373)	0.008 ± 0.003
Hydrogen bond (182/14)	0.011 ± 0.002
Intermolecular (204)	0.017 ± 0.004
Rms deviations from experimental dihedral restraints (°) (286)	0.17 ± 0.05
Rms deviations from idealized covalent geometry	
Bonds (Å)	0.0017 ± 0.0001
Angles (°)	0.49 ± 0.01
Impropers (°)	0.37 ± 0.01
Energies (kcal mol^{-1})	
F_{NOE}[b]	36.3 ± 10.4
F_{cdih}[b]	0.60 ± 0.40
F_{repel}[c]	32.3 ± 7.5
$E_{L\text{-}J}$[d]	−312 ± 44

[a] The number of each type of restraint used in the structure calculation is given in parentheses. The number of TBP restraints is given first, followed by the number of dTAF$_{II}$230$_{11–77}$ restraints. None of the structures exhibit distance violations greater than 0.50 Å or dihedral angle violations greater than 5°.

[b] F_{NOE} and F_{cdih} were calculated using force constants of 50 kcal mol^{-1} Å^{-2} and 200 kcal mol^{-1} rad^{-2}, respectively.

[c] F_{repel} was calculated using a final value of 4.0 kcal mol^{-1} Å^{-4} with the van der Waals hard sphere radii set to 0.75 times those in the parameter set PARALLHSA supplied with X-PLOR (Brünger, 1993).

[d] $E_{L\text{-}J}$ is the Lennard-Jones van der Waals energy calculated with the CHARMM empirical energy function and is not included in the target function for simulated annealing calculation.

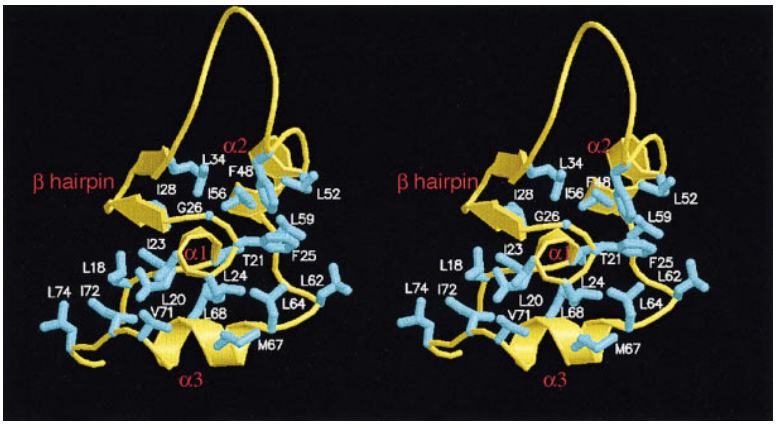

Figure 4. $dTAF_{II}230_{11–77}$ Hydrophobic Core

Stereoview of the $dTAF_{II}230_{11–77}$ fold in the TBP–$dTAF_{II}230_{11–77}$ complex, showing residues involved in hydrophobic core formation. The backbone is shown as a yellow ribbon. Secondary structure elements are labeled. This figure was generated using MOLSCRIPT (Kraulis, 1991) and rendered using Raster3D (Merritt and Bacon, 1997).

structures is 2.2 ± 0.2 Å. The TBP molecular saddle consists of two similar α/β structural motifs, related by approximate intramolecular 2-fold symmetry (Nikolov et al., 1992), that correspond to two imperfect tandem repeats in the sequence (reviewed in Hernandez, 1993). In the ensemble of NMR-derived TBP–$dTAF_{II}230_{11–77}$ structures (Figure 3), the pairwise rms deviation for TBP is 0.67 Å for backbone atoms and 1.09 Å for heavy atoms (residues 61–239 superimposed). These values improve to 0.51 Å and 1.05 Å when residues 64–144 of the N-terminal motif only are superimposed and to 0.44 Å and 0.95 Å when residues 155–236 of the C-terminal motif only are superimposed, indicating that the higher rms deviation for the entire molecule originates from the poor definition of the relative positions of the two motifs. This may reflect hinge motion between the two α/β motifs. This possibility is supported by the modest conformational changes that occur upon TBP binding to the TATA element in the TBP–TATA (Kim et al., 1993b) and TFIIB–TBP–TATA complexes (Nikolov et al., 1995), involving a twisting motion of one TBP motif with respect to the other (Kim and Burley, 1994).

TBP–$dTAF_{II}230$ Interface

The protein–protein interface consists of the TBP concave undersurface and the extensive hydrophobic patch on $dTAF_{II}230_{11–77}$. The total surface area buried at the interface is 2950 ± 140 Å^2, similar to the value (about 3150 Å^2) reported for the interface between *Arabidopsis thaliana* TBP and the adenovirus major late promoter TATA element (Kim and Burley, 1994). The surface area of $dTAF_{II}230_{11–77}$ in contact with TBP is therefore 1475 ± 70 Å^2, which corresponds to 25% of the entire accessible surface area of $dTAF_{II}230_{11–77}$. Seventy percent of the surface area buried between TBP and $dTAF_{II}230_{11–77}$ is hydrophobic, close to the corresponding value of 74% for the TBP–TATA complex (Kim and Burley, 1994).

The hydrophobic interface formed between $dTAF_{II}230_{11–77}$ and the N-terminal motif of TBP includes many methyl-containing and phenylalanine residues: the side chains of $dTAF_{II}230_{11–77}$ residues Ile-23, Ile-28, Leu-68, and Val-71 contact numerous TBP residues, including Ile-103, Thr-112, Leu-114, Val-122, and Thr-124 (Figures 5A and 5B). Leu-18, Leu-20, Ile-72, and Leu-74 interact with two aromatic TBP residues, Phe-99 and Phe-116. The interface formed between $dTAF_{II}230_{11–77}$ and the C-terminal motif of TBP includes $dTAF_{II}230_{11–77}$ residues Phe-25, Leu-52, Leu-59, and Leu-64 in contact with TBP residues Val-161, Ile-194, Leu-205, Val-213, and Thr-215 (Figures 5A and 5B). As a result of the quasi-2-fold symmetry of the TBP structure, the C-terminal half of the TBP concave surface contains two phenylalanines, Phe-190 and Phe-207, corresponding to Phe-99 and Phe-116 in the N-terminal motif. $dTAF_{II}230_{11–77}$ residues Leu-59 and Leu-62 contact Phe-190, while Glu-51 and Leu-52 contact Phe-207.

The solution structure suggests that electrostatic interactions and hydrogen bonds play a role in determining the orientation of $dTAF_{II}230_{11–77}$ relative to the TBP saddle. Although the degree of precision limits our ability to define with certainty these interactions, the structures obtained here do indicate several possibilities for intermolecular electrostatic interactions and hydrogen bonds. Potential electrostatic interactions between TBP and $dTAF_{II}230_{11–77}$ involve the following pairs of residues: Arg-98 and Glu-31, Arg-105 and Asp-29, Lys-120 and Asp-73, Lys-211 and Glu-51, and Lys-218 and Glu-70. The basic TBP residues in all of these pairs are involved in electrostatic interactions with the phosphate backbone of the TATA element (Arg-98 and T6′/T7′, Arg-105 and T5′, Lys-211 and A1′/T2′, Lys-120 and A7, and Lys-218 and A6) in the TBP–TATAAAG complex (Kim et al., 1993b; Kim and Burley, 1994). Lys-61 of $dTAF_{II}230_{11–77}$ may also interact with Glu-186 of TBP, which is located at the C-terminal stirrup and does not contact DNA in the TBP–DNA complex. Possible hydrogen bonds involve the hydroxyl groups of the symmetry-related TBP residues Thr-124 and Thr-215 with the backbone carbonyl oxygens of Ile-23 and Leu-24 of $dTAF_{II}230_{11–77}$ as acceptor atoms.

The 2-fold symmetry of the TBP concave surface is disrupted at three positions: Lys-97, Arg-98, and Thr-112 in the N-terminal motif correspond to Glu-188, Leu-189, and Val-203 in the C-terminal motif (Figures 5A and 5C). The C-terminal motif is consequently more hydrophobic than the N-terminal motif. The side chain hydroxyl group of Thr-112 in the TBP N-terminal motif probably forms a hydrogen bond with the side chain carbonyl group of Asp-29. The corresponding C-terminal motif residue, Val-203, interacts with $dTAF_{II}230_{11–77}$ hydrophobic residues Phe-25, Leu-64, and Met-67. As mentioned above, Arg-98 in the N-terminal TBP stir-

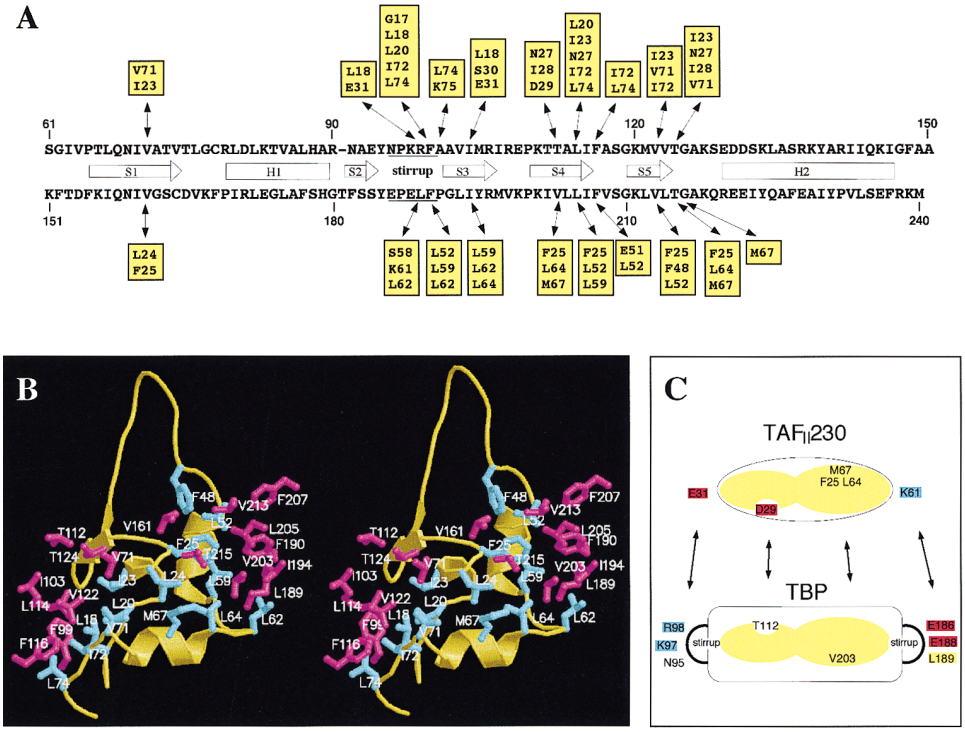

Figure 5. Molecular Interface between TBP and dTAF$_{II}$230$_{11–77}$

(A) Summary of residue pairs for which intermolecular NOEs between TBP and dTAF$_{II}$230$_{11–77}$ are observed. The dTAF$_{II}$230 residues that interact with the TBP direct repeat α/β motifs (top, N-terminal TBP motif, and bottom, C-terminal TBP motif) are listed inside yellow boxes. Secondary structure elements of the TBP motifs are represented by arrows (β strands) and boxes (α helices).

(B) Stereoview showing hydrophobic interactions between TBP and dTAF$_{II}$230$_{11–77}$. TBP and TAF$_{II}$230$_{11–77}$ side chains are shown in magenta and cyan. The dTAF$_{II}$230$_{11–77}$ backbone is shown as a yellow ribbon. The TBP backbone is omitted for clarity.

(C) Schematic diagram representing the interactions that confer asymmetry on dTAF$_{II}$230$_{11–77}$ binding to TBP. Acidic residues are highlighted with a red box, basic residues with a blue box, and hydrophobic regions are yellow. Whereas Thr-112 is involved in a putative side chain–side chain hydrogen bond with Asp-29, the corresponding C-terminal TBP motif residue, Val-203, is involved in hydrophobic interactions with Phe-25, Leu-64, and Met-67.

rup and Glu-186 in the C-terminal stirrup are involved in electrostatic interactions with Glu-31 and Lys-61 of dTAF$_{II}$230$_{11–77}$. Together, such asymmetric interactions position dTAF$_{II}$230$_{11–77}$ in a single orientation with respect to the quasi-symmetric concave surface of TBP.

Since NMR and CD spectra both indicate that free dTAF$_{II}$230$_{11–77}$ is largely unfolded but that it adopts a single folded conformation in complex with TBP, molecular recognition between TBP and dTAF$_{II}$230$_{11–77}$ can be described as a local folding process. The tight TBP–dTAF$_{II}$230$_{11–77}$ interaction and extensive degree of induced folding suggest that the heat capacity change of folding in this case is large. Such induced folding is becoming a common theme in protein–DNA (Spolar and Record, 1994), protein–RNA (Frankel and Smith, 1998), and protein–protein molecular recognition events. Other examples of induced folding involving transcription factors are the NFATC1 DNA-binding domain upon complexation with DNA (Zhou et al., 1998), the VP16 activation domain upon binding to human TAF$_{II}$31 (Uesugi et al., 1997), the phosphorylated kinase-inducible domain of CREB upon binding to the KIX domain of CBP (Radhakrishnan et al., 1997), and possibly the p53 activation domain upon binding to the MDM2 oncoprotein (Kussie et al., 1996).

Correlation with Mutagenesis Studies

The structure of the TBP–dTAF$_{II}$230$_{11–77}$ complex is consistent with previous mutagenesis studies on TAF$_{II}$230 (Kokubo et al., 1994). Block-alanine-scanning mutagenesis was used to mutate ten nonoverlapping blocks of amino acids in the N-terminal 80 residues of dTAF$_{II}$230. Of the ten block mutants, those involving residues 19–28 and 67–71 showed dramatic effects on TBP binding. These regions correspond to helices α1 and α3 and part of the β hairpin, which are all crucial to the intimate dTAF$_{II}$230–TBP interaction. Deletion mutagenesis also indicated that Leu-20, Phe-25, and Glu-70 in dTAF$_{II}$230 are important for TBP binding and inhibition of TBP-promoter interaction. These residues, located at either end of helix α1 and at the C terminus of helix α3, are involved in interactions at the center of the TBP–TAF$_{II}$230$_{11–77}$ interface. On the other hand, block mutation of residues 37–41 and 43–47 to alanines produced no significant effects on TBP binding by dTAF$_{II}$230. These residues correspond to the long loop, which makes no direct contact with TBP (Figure 3). Together with the fact that hTAF$_{II}$250 lacks eight residues in the region corresponding to this long loop (Figure 1B), the mutation and structural results suggest that the loop is unimportant for binding to TBP.

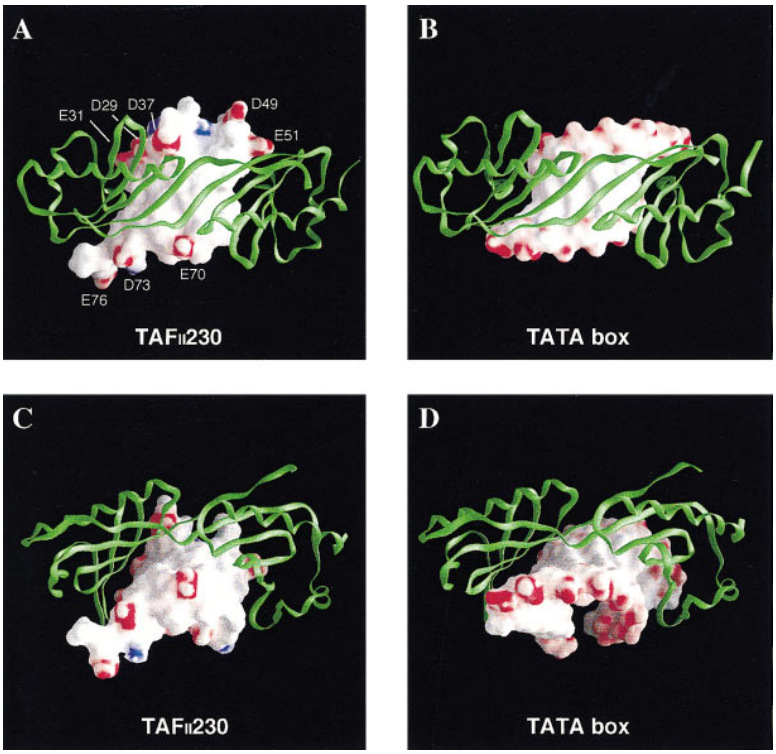

Figure 6. $dTAF_{II}230$ Mimics the Partially Unwound Minor Groove Surface of the TATA Box in the TBP–TATA Complex Structure

(A and C) The molecular surface of the TBP-binding domain of $dTAF_{II}230_{11–77}$ in the TBP–$dTAF_{II}230_{11–77}$ complex. The loop between helix α2 and the β hairpin has been omitted. In (A), acidic residues in $dTAF_{II}230_{11–77}$ that are close to TBP are labeled. The negatively charged side chains of Asp-29, Glu-31, Glu-51, Glu-70, and Asp-73 mimic the positions of negatively charged phosphate groups in the unwound minor groove of the TATA element that interact with Arg-105, Arg-98, Lys-211, Lys-218, and Lys-120 of TBP (Kim et al., 1993b; Kim and Burley, 1994).

(B and D) The molecular surface of the TBP-binding surface of the TATA element in the TBP–TATA complex (Kim et al., 1993a). Molecular surfaces are colored according to electrostatic potential. Blue corresponds to positive potential, red to negative potential. TBP is shown as a green ribbon. The TBP orientation is the same in (A) and (B) and in (C) and (D). The views in (C) and (D) are generated by an approximately 90° rotation of the views in (A) and (B). These figures were generated using GRASP (Nicholls, 1993).

In mutagenesis studies on TBP (Nishikawa et al., 1997), the most dramatic effect was observed with the Leu-114-Lys mutation that abolished the TBP–$dTAF_{II}230$ interaction. The mutations Phe-99-Lys, Ile-103-Lys, Phe-116-Lys, and Val-122-Lys also showed significant but slightly smaller effects on TBP binding to $dTAF_{II}230$. In the solution structure of the TBP–$dTAF_{II}230_{11–77}$ complex, these mutated TBP hydrophobic residues are located in a cluster around Leu-114 and pack against the hydrophobic patch of $dTAF_{II}230_{11–77}$ (Figures 4 and 5B). These TBP residues are also critical for TBP–DNA interaction (Kim et al., 1993a, 1993b). Mutation of Leu-114 to alanine, however, produced no significant effects on TBP binding by $dTAF_{II}230$ (J. Nishikawa and Y. N., unpublished results), presumably due to the small change in side chain hydrophobicity. Phe-116 and Phe-207, which occupy homologous positions in the N- and C-terminal motifs of TBP, both contribute to hydrophobic interactions with DNA bases. The Phe-116-Lys mutation produced a greater effect on TBP–$dTAF_{II}230$ interaction than the Phe-207-Lys mutation (Nishikawa et al., 1997). This may reflect subtle differences in the Phe-116 and Phe-207 interactions with $dTAF_{II}230_{11–77}$.

$dTAF_{II}230$ Mimics the TATA Minor Groove Surface

$dTAF_{II}230_{11–77}$ and the TATA element bind to the same surface of TBP. How does the protein–protein interaction compare with the protein–DNA interaction? To our surprise, the $dTAF_{II}230_{11–77}$ surface that interacts with the TBP concave surface markedly resembles the DNA surface in the TBP–DNA complex (Kim and Burley, 1994). The arch-shaped surface of $dTAF_{II}230_{11–77}$ is similar to that of the partially unwound minor groove surface of the TATA element bound to TBP (Figure 6). Both $dTAF_{II}230_{11–77}$ and the TATA element have an extensive hydrophobic surface: in $dTAF_{II}230_{11–77}$, numerous methyl-containing residues, together with Phe-25 and Phe-48, cluster to form the hydrophobic patch (Figures 4 and 5B), and in the TATA element, nucleotide bases in the unwound minor groove form the DNA hydrophobic surface. The fraction of buried surface area that is hydrophobic is very similar in the TBP–$dTAF_{II}230_{11–77}$ and TBP–TATA (Kim and Burley, 1994) complexes (70% versus 74%). Furthermore, the negatively charged side chains of $dTAF_{II}230_{11–77}$ residues Asp-29, Glu-31, Glu-51, Glu-70, and Asp-73 mimic the positions of negatively charged phosphate groups in the partially unwound minor groove of the TATA element (Figures 6A and 6B). As noted previously, potential electrostatic interactions at the TBP–$dTAF_{II}230_{11–77}$ interface involve the same TBP lysines and arginines that make direct contact with the TATA element phosphate backbone in TBP–DNA complexes (Kim et al., 1993a, 1993b; Kim and Burley, 1994).

An important difference between the TATA element and $dTAF_{II}230_{11–77}$ is that the TATA element surface is highly symmetric, whereas the $dTAF_{II}230_{11–77}$ surface is asymmetric. This asymmetry results in closer matching of surface charge and hydrophobicity between $dTAF_{II}230_{11–77}$ and TBP. The relevant interactions have been described above (TBP–$dTAF_{II}230_{11–77}$ Interface) and are depicted in Figure 5C. The complementary asymmetry between the two polypeptide surfaces confers a specific relative orientation on the TBP–$dTAF_{II}230_{11–77}$ interaction.

A macromolecular mimicry hypothesis has been put forward (Nyborg et al., 1996) based on the observation that part of the *Thermus aquaticus* elongation factor EF-G structure mimics the tRNA structure in the ternary complex of EF–Tu–GTP and tRNA (Nissen et al., 1995).

The structure of the TBP–dTAF$_{II}$230$_{11–77}$ binary complex demonstrates transcription factor (dTAF$_{II}$230) mimicry of a DNA surface (the exposed minor groove of the TATA element partially unwound by TBP) upon binding to a sequence-specific DNA-binding protein (TBP). The macromolecular mimicry observed in translation (Nyborg et al., 1996) and in transcription (this study) shows that protein structures can mimic both RNA and DNA structures. A more localized molecular mimicry of protein–nucleotide interaction in a protein–protein complex has been observed in the uracil–DNA glycosylase inhibitor protein complex (Mol et al., 1995; Savva and Pearl, 1995).

Implications for Regulation of Transcription

Direct inhibition of TBP interaction with the TATA element is an efficient way to shut down gene transcription. The extensive overlap between the TBP–dTAF$_{II}$230$_{11–77}$ protein–protein interface and the TBP–TATA protein–DNA interface implies that removal of dTAF$_{II}$230$_{11–77}$ from the TBP concave surface is required for TBP–TATA binding. Although the mechanism by which this removal is effected remains to be determined, the ability of dTAF$_{II}$ 230 subdomain I to dissociate prebound TBP from the TATA box (Kokubo et al., 1994) indicates that another portion or portions of dTAF$_{II}$230 or other regulatory factor(s) are involved. Indeed, it has been proposed that the C-terminal portion of dTAF$_{II}$230 may be involved in control of the N-terminal inhibitory domain (Kokubo et al., 1994).

TBP has been identified as a target for numerous cellular and viral transcription activators (reviewed in Hernandez, 1993). Although the exact mechanisms are not clear, VP16 activation domain binds functionally to the concave surface of TBP in vitro (Kim et al., 1994). In support of the in vivo relevance of this interaction, yeast genetic analyses revealed that various TBP mutants that cannot mediate activated transcription map to the concave surface (Arndt et al., 1995; Lee and Struhl, 1995). One of these TBP mutants (Leu-114-Lys) is unable to bind VP16 and to mediate VP16-dependent activation both in vitro (Kim et al., 1994) and in vivo (Lee and Struhl, 1995). Importantly, the same TBP mutation eliminates TBP binding to subdomain I of dTAF$_{II}$230 (Nishikawa et al., 1997). Consistent with these effects of the Leu-114-Lys mutation, subdomain I and the VP16 activation domain compete for binding to wild-type TBP (Nishikawa et al., 1997). On the other hand, experiments with functionally homologous yTAF$_{II}$145 demonstrated that subdomain II and TFIIA bind competitively to the convex TBP upper surface (Kokubo et al., 1998). These results suggest that transcriptional regulation involves, in part, competition for TBP between the dTAF$_{II}$230 N-terminal inhibitory domain and positive regulators, including transcriptional activators and TFIIA.

Several lines of evidence suggest structural similarity between activation domains and the dTAF$_{II}$230 inhibitory domain. In common with various acidic activation domains such as those of VP16, GAL4, and GCN4, dTAF$_{II}$230$_{11–77}$ is rich in acidic residues with a pI of 3.84. As reported for various acidic activators (reviewed in Drysdale et al., 1998), moreover, mutagenesis experiments (Kokubo et al., 1993; Nishikawa et al., 1997) and the predominantly hydrophobic nature of the TBP–dTAF$_{II}$230$_{11–77}$ interface (Figure 5) show that dTAF$_{II}$230$_{11–77}$ hydrophobic residues play an important role in TBP interaction. In further support of the proposed structural similarity, mutation of TBP residue Leu-114, which interacts directly with the dTAF$_{II}$230$_{11–77}$ hydrophobic patch, abolishes interaction with both dTAF$_{II}$230$_{11–77}$ and the VP16 activation domain (Nishikawa et al., 1997). In common with dTAF$_{II}$230$_{11–77}$, induced folding has been reported for several activation domains: VP16 activation domain induced by hTAF$_{II}$31 (Uesugi et al., 1997), CREB induced by coactivator CBP (Radhakrishnan et al., 1997), and p53 activation domain induced by the p53 inhibitor MDM2 (Kussie et al., 1996). These observations support the possibility that activators compete with dTAF$_{II}$230 subdomain I for TBP in part by mimicking the structure of dTAF$_{II}$230$_{11–77}$.

Conclusions

The present study provides concrete evidence that the N-terminal region of dTAF$_{II}$230 interacts directly with the concave undersurface of TBP and supports the previously suggested mechanism by which the N-terminal region of dTAF$_{II}$230 inhibits TATA box binding by TBP (Kokubo et al., 1994; Nishikawa et al., 1997). While the structural analyses of VP16 (Uesugi et al., 1997) and CREB (Radhakrishnan et al., 1997) activation domains demonstrate coil-to-helix induced folding processes upon binding to a partner protein, the present study features a more complex induced folding process whereby dTAF$_{II}$230$_{11–77}$ adopts a fold consisting of three α helices and a β hairpin upon binding to TBP. These elements combine to form an extensive hydrophobic surface that, together with a number of charged side chains, closely matches the hydrophobicity and charge distribution of the concave undersurface of TBP. As a result, dTAF$_{II}$230$_{11–77}$ mimics not only the extensive hydrophobic surface seen in the minor groove surface of the partially unwound TATA element bound to TBP, but also several electrostatic interactions between TBP and DNA phosphate groups. This molecular mimicry of a DNA surface, together with other examples including protein mimicry of globular tRNA conformation (Nyborg et al., 1996), indicates that protein structures are versatile and can mimic nonpeptide surfaces. Further investigation is required to understand how the TBP–dTAF$_{II}$230 interaction is regulated to facilitate TBP binding to the TATA box with consequent nucleation of transcription initiation complex assembly.

Experimental Procedures

Overexpression and Purification of TBP and dTAF$_{II}$230$_{11–77}$

The TBP gene from *S. cerevisiae* was subcloned into 6His-pET-11d. The vector was transformed into BL21(DE3) (Novagen) for protein overexpression. Cells were grown at 30°C and induced with 0.5 mM IPTG for 2–4 hr. Protein was purified using a standard Ni-NTA affinity column protocol (Qiagen) and then dialyzed overnight against 50 mM Tris-HCl (pH 8.4), 150 mM KCl, 2.5 mM $CaCl_2$, and 20% glycerol. The hexa-histidine tag and the N-terminal 48 amino acids of the expressed protein were removed by trypsin proteolytic cleavage at 4°C. SP Sepharose cation exchange resin (Pharmacia) was used to remove trypsin and other remaining contaminants. The final purified

protein contained amino acids 49–240 of *S. cerevisiae* TBP, as confirmed by electrospray ionization mass spectrometry.

DNA encoding *Drosophila* $TAF_{II}230_{11–77}$ was subcloned into pGEX-2T (Pharmacia) for expression as a glutathione S-transferase (GST) fusion protein (Kokubo et al., 1994). The vector was transformed into BL21(DE3) (Novagen). Cells were grown at 37°C and induced with 0.5 mM IPTG for 3 hr. The GST–$dTAF_{II}230_{11–77}$ fusion protein was purified with glutathione sepharose using a standard protocol (Pharmacia). $dTAF_{II}230_{11–77}$ was cleaved from GST by resuspending the resin in 50 mM Tris-HCl (pH 8.4), 150 mM KCl, 2.5 mM $CaCl_2$ and reacting with thrombin at room temperature for 1–4 hr with mixing. $dTAF_{II}230_{11–77}$ was further purified by FPLC using a mono-Q column (Pharmacia) to remove thrombin and any other impurities.

Preparation of Isotope-Labeled Proteins

Isotope-labeled protein samples were obtained by bacterial expression in M9 minimal medium containing ^{15}N-ammonium chloride and/or $^{13}C_6$-D-glucose and/or ^2H_2O. Several different types of isotope labeling were carried out for NMR studies. For TBP, samples included unlabeled, uniformly ^{15}N-labeled, $^{15}N/^{13}C$-labeled, $^{15}N/^2H$-labeled, and $^{15}N/^{13}C/^2H$-labeled (for both types of ^2H-labeled samples, >98% ^2H_2O was used in the M9 medium). $dTAF_{II}230_{11–77}$ samples included unlabeled, uniformly ^{15}N-labeled, $^{15}N/^{13}C$-labeled, and $^{15}N/^{13}C/^2H$-labeled (for the $dTAF_{II}230_{11–77}$ ^2H-labeled sample, 60% ^2H_2O was used in the M9 medium).

Isolation of TBP–$dTAF_{II}230_{11–77}$ Complex

The TBP–$dTAF_{II}230_{11–77}$ complex was formed by mixing purified protein solutions at low TBP concentration (<0.2 mg/ml) in the presence of 0.3–0.5 M KCl. Unlabeled protein was always kept in excess. The mixed solution was dialyzed overnight at 4°C with buffer containing 20 mM HEPES (pH 7.5), 10% glycerol, 0.5 mM PMSF. The dialyzed solution was applied to a mono-S column (Pharmacia), and the $dTAF_{II}230_{11–77}$–TBP complex was eluted using a salt gradient. The complex, now separated from free TBP and free $dTAF_{II}230_{11–77}$, was concentrated further using Centricon-10 (Amicon) for NMR studies.

NMR Spectroscopy

NMR samples of TBP–$dTAF_{II}230_{11–77}$ comprised one isotope-labeled protein and one unlabeled protein with a protein concentration in the range of 0.7–1.2 mM. Samples were prepared in either 95% H_2O/5% D_2O or 99.9% D_2O. The NMR sample buffer, optimized for protein stability and solubility using the microdialysis button test (Bagby et al., 1997), comprised 20 mM sodium phosphate (pH 7.0), 5% (v/v) perdeuterated glycerol, 250 mM sodium sulfate, 5 mM $MgCl_2$, 10 mM perdeuterated 1,4-dithiothreitol, 0.5 mM 4-(2-aminoethyl)-benzynsulfonyl fluoride hydrochloride, and 0.05 mM sodium azide. Sample buffers were purged with nitrogen before use.

NMR data were recorded at 25°C on Varian 600 MHz and 500 MHz NMR spectrometers. Resonance assignments for backbone ^1H, ^{13}C, and ^{15}N nuclei of both TBP and $dTAF_{II}230_{11–77}$ were obtained through 3-D HNCA/HNCOCA and 3-D HN(CA)CB/HN(COCA)CB experiments (Yamazaki et al., 1994). Side chain resonances were assigned using 3-D ^{15}N-edited TOCSY-HSQC and HCCH-TOCSY data sets supplemented with other experiments, including 3-D HNCACB, CBCA(CO)NNH, HNCO, (HB)CBCACOCAHA, and HNHB (Clore and Gronenborn, 1994; Kay and Gardner, 1997). Aromatic resonances of both TBP and $dTAF_{II}230_{11–77}$ were mainly assigned using simultaneous $^{13}C/^{15}N$-edited NOESY-HSQC spectra (Pascal et al., 1994) recorded in D_2O. Spectra were processed with the NMRPipe/NMRDraw programs (Delaglio et al., 1995) and analyzed with PIPP/CAPP (Garrett et al., 1991) and NMRView (Johnson and Blevins, 1994).

Structure Calculations

Approximate interproton distance restraints were obtained from multidimensional NOE spectra. For TBP, all the backbone amide NH-NH NOEs were obtained from a 3-D ^{15}N-edited NOESY spectrum recorded with a 200 ms NOE mixing time using a perdeuterated ^{15}N-labeled sample. For $dTAF_{II}230_{11–77}$, all the backbone amide NH-NH NOEs were obtained from a 3-D ^{15}N-edited NOESY with a 100 ms NOE mixing time using a $^{13}C/^{15}N$-labeled sample with 60% deuteration. NOEs involving side chain protons were obtained from 3-D ^{15}N-edited NOESY-HSQC and simultaneous $^{13}C/^{15}N$-edited NOESY-HSQC spectra (Pascal et al., 1994), with mixing times of 100 and 70 ms, using $^{13}C/^{15}N$-labeled TBP complexed with unlabeled $dTAF_{II}230_{11–77}$ or vice versa. In addition, a 3-D simultaneous $^{13}C/^{15}N$-edited NOESY-HSQC with a 100 ms mixing time using a $^{13}C/^{15}N$-labeled $dTAF_{II}230_{11–77}$ sample with 60% deuteration was useful to confirm the NOE assignments obtained using the $^{13}C/^{15}N$ double-labeled sample, especially for backbone NH-aliphatic NOEs. Intermolecular distance constraints were identified in 3-D $^{15}N/^{13}C$-separated, ^{13}C-filtered NOESY spectra (Lee et al., 1994; Zwahlen et al., 1997) using a $^{15}N/^{13}C$-labeled sample of TBP complexed with unlabeled $dTAF_{II}230_{11–77}$ and using a $^{15}N/^{13}C$-labeled $dTAF_{II}230_{11–77}$ sample with 60% deuteration complexed with unlabeled TBP. Mixing times of 100 ms and 70 ms were used for $dTAF_{II}230_{11–77}$ and TBP.

All NOEs were grouped into three distance ranges: 1.8–2.9 Å, 1.8–3.5 Å, and 1.8–5.0 Å, corresponding to strong, medium, and weak NOEs, respectively. Standard pseudo-atom distance corrections (Wüthrich et al., 1983) were incorporated to account for center averaging. An additional 0.5 Å was added to the upper limits for distances involving methyl groups to account for the higher apparent intensity of methyl resonances. Backbone hydrogen bond restraints within regular secondary structure elements that were consistent with backbone amide H/D exchange data were added at the final stage of the structure calculation.

Backbone dihedral angle constraints for TBP and $dTAF_{II}230_{11–77}$ were obtained from HNHA (Vuister and Bax, 1993) and HMQC-J (Kay and Bax, 1990) spectra and from chemical shift indices (Wishart and Sykes, 1994). For residues with $^3J_{HNH\alpha} < 5$ Hz, ϕ was restrained to $-50° \pm 40°$, and for residues with $^3J_{HNH\alpha} > 7$ Hz, ϕ was restrained to $-120° \pm 40°$. ψ dihedral angle constraints were included only for those residues within regular secondary structure elements ($-50° \pm 50°$ for residues in a helix and $130° \pm 50°$ for residues in a strand). The $^{13}C\alpha$ and $^{13}C\beta$ chemical shifts were used to confirm the secondary structure determination but were not used for direct structure refinement.

The structures were calculated using a simulated annealing protocol (Nilges et al., 1991) using the program X-PLOR (Brünger, 1993). The final structures were generated based on a total of 4242 interproton distance restraints. A total of 60 simulated annealing structures were calculated, and 46 structures were found with no NOE violations greater than 0.5 Å and no dihedral violation greater than 5°. These structures were subjected to further refinement, and 42 final refined structures were obtained. Structural statistics are presented in Table 1.

Acknowledgments

We thank R. Roeder and A. Hoffmann for providing the yeast TBP expression vector, and M. Inouye and C. Arrowsmith for helpful discussions. This work was supported by a grant to M. I. from the Medical Research Council of Canada, and in part by a grant from the Japan Society for Promotion of Science. D. L. and R. I. acknowledge the Human Frontier Science Program for the award of postdoctoral fellowships. M. I. is an MRC Scientist and Howard Hughes Medical Institute International Research Scholar.

Received June 29, 1998; revised July 29, 1998.

References

Arndt, K.M., Ricupero-Hovasse, S., and Winston, F. (1995). TBP mutants defective in activated transcription in vivo. EMBO J. *14*, 1490–1497.

Aso, T., Conaway, J.W., and Conaway, R.C. (1994). Role of core promoter structure in assembly of the RNA polymerase II preinitiation complex. A common pathway for formation of preinitiation intermediates at many TATA and TATA-less promoters. J. Biol. Chem. *269*, 26575–26583.

Bagby, S., Tong, K.I., Liu, D., Alattia, J.-R., and Ikura, M. (1997). The button test: a small scale method using microdialysis cells for assessing protein solubility at concentrations suitable for NMR. J. Biomol. NMR *10*, 279–282.

Brünger, A.T. (1993). X-PLOR Version 3.1: A System for X-Ray Crystallography and NMR (New Haven, CT: Yale University Press).

Burley, S.K., and Roeder, R.G. (1996). Biochemistry and structural biology of transcription factor IID (TFIID). Annu. Rev. Biochem. *65*, 769–799.

Carson, M. (1991). Ribbons 2.0. J. Appl. Crystallogr. *24*, 958–961.

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.

Clore, G.M., and Gronenborn, A.M. (1994). Multidimensional heteronuclear nuclear magnetic resonance of proteins. Methods Enzymol. *239*, 349–363.

Coleman, R.A., Taggart, A.K., Benjamin, L.R., and Pugh, B.F. (1995). Dimerization of the TATA binding protein. J. Biol. Chem. *270*, 13842–13849.

Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J., and Bax, A. (1995). NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J. Biomol. NMR *6*, 277–293.

Dikstein, R., Ruppert, S., and Tjian, R. (1996). $TAF_{II}250$ is a bipartite protein kinase that phosphorylates the basal transcription factor RAP74. Cell *84*, 781–790.

Drysdale, C.M., Jackson, B.M., McVeigh, R., Klebanow, E.R., Bai, Y., Kokubo, T., Swanson, M., Nakatani, Y., Weil, P.A., and Hinnebusch, A.G. (1998). The Gcn4p activation domain interacts specifically in vitro with RNA polymerase II holoenzyme, TFIID, and the Adap-Gcn5p coactivator complex. Mol. Cell. Biol. *18*, 1711–1724.

Frankel, A.D., and Smith, C.A. (1998). Induced folding in RNA–protein recognition: more than a simple molecular handshake. Cell *92*, 149–151.

Garrett, D.S., Powers, R., Gronenborn, A.M., and Clore, G.M. (1991). A common sense approach to peak picking in two-, three-, and four-dimensional spectra using automatic computer analysis of contour diagrams. J. Magn. Reson. *95*, 214–220.

Geiger, J.H., Hahn, S., Lee, S., and Sigler, P.B. (1996). Crystal structure of the yeast TFIIA-TBP-DNA complex. Science *272*, 830–836.

Guermah, M., Malik, S., and Roeder, R.G. (1998). Involvement of TFIID and USA components in transcriptional activation of the human immunodeficiency virus promoter by NF-κB and Sp1. Mol. Cell. Biol. *18*, 3234–3244.

Hernandez, N. (1993). TBP, a universal eukaryotic transcription factor? Genes Dev. *7*, 1291–1308.

Holm, L., and Sander, C. (1993). Protein structure comparison by alignment of distance matrices. J. Mol. Biol. *233*, 123–138.

Johnson, B.A., and Blevins, R.A. (1994). NMRView: a computer program for the visualization and analysis of NMR data. J. Biomol. NMR *4*, 603–614.

Kaiser, K., and Meisterernst, M. (1996). The human general co-factors. Trends Biochem. Sci. *21*, 342–345.

Kay, L.E., and Bax, A. (1990). New methods for the measurement of NH-CαH coupling constants in ^{15}N-labeled proteins. J. Magn. Reson. *86*, 110–126.

Kay, L.E., and Gardner, K.H. (1997). Solution NMR spectroscopy beyond 25 kDa. Curr. Opin. Struct. Biol. *7*, 722–731.

Kim, J.L., and Burley, S.K. (1994). 1.9 Å resolution refined structure of TBP recognizing the minor groove of TATAAAAG. Nat. Struct. Biol. *1*, 638–653.

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

Kim, J.L., Nikolov, D.B., and Burley, S.K. (1993b). Co-crystal structure of TBP recognizing the minor groove of a TATA element. Nature *365*, 520–527.

Kim, T.K., Hashimoto, S., Kelleher, R.J., III, Flanagan, P.M., Kornberg, R.D., Horikoshi, M., and Roeder, R.G. (1994). Effects of activation-defective TBP mutations on transcription initiation in yeast. Nature *369*, 252–255.

Kokubo, T., Takeda, R., Yamashita, S., Gong, D.W., Roeder, R.G., Horikoshi, M., and Nakatani, Y. (1993). Identification of TFIID components required for transcriptional activation by upstream stimulatory factor. J. Biol. Chem. *268*, 17554–17558.

Kokubo, T., Yamashita, S., Horikoshi, M., Roeder, R.G., and Nakatani, Y. (1994). Interaction between the N-terminal domain of the 230-kDa subunit and the TATA box-binding subunit of TFIID negatively regulates TATA-box binding. Proc. Natl. Acad. Sci. USA *91*, 3520–3524.

Kokubo, T., Swanson, M.J., Nishikawa, J., Hinnebusch, A., and Nakatani, Y. (1998). The yeast TAF145 inhibitory domain and TFIIA competitively bind to TATA-binding protein. Mol. Cell. Biol. *18*, 1003–1012.

Kosa, P.F., Ghosh, G., DeDecker, B.S., and Sigler, P.B. (1997). The 2.1-Å crystal structure of an archaeal preinitiation complex: TATA-box-binding protein/transcription factor (II)B core/TATA-box. Proc. Natl. Acad. Sci. USA *94*, 6042–6047.

Kraulis, P. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. J. Appl. Crystallogr. *24*, 946–950.

Kussie, P.H., Gorina, S., Marechai, V., Elenbaas, B., Moreau, J., Levine, A.J., and Pavletich, N.P. (1996). Structure of the MDM2 oncoprotein bound to the p53 tumor suppressor transactivation domain. Science *274*, 948–953.

Lee, M., and Struhl, K. (1995). Mutations on the DNA-binding surface of TATA-binding protein can specifically impair the response to acidic activators in vivo. Mol. Cell. Biol. *15*, 5461–5469.

Lee, W., Revington, M.J., Arrowsmith, C.H., and Kay, L.E. (1994). A pulsed field gradient isotope-filtered 3D ^{13}C HMQC-NOESY experiment for extracting intermolecular NOE contacts in molecular complexes. FEBS Lett. *350*, 87–90.

Merritt, E.A., and Bacon, D.J. (1997). Raster3D photorealistic molecular graphics. Methods Enzymol. *277*, 503–524.

Mizzen, C.A., Yang, X.-J., Kokubo, T., Brownell, J.E., Bannister, A.J., Owen-Hughes, T., Workman, J., Wang, L., Berger, S.L., Kouzarides, T., et al. (1996). The $TAF_{II}250$ subunit of TFIID has histone acetyltransferase activity. Cell *87*, 1261–1270.

Mol, C.D., Arvai, A.S., Sanderson, R.J., Slupphaug, G., Kavli, B., Krokan, H.E., Mosbaugh, D.W., and Tainer, J.A. (1995). Crystal structure of human uracil–DNA glycosylase in complex with a protein inhibitor: protein mimicry of DNA. Cell *82*, 701–708.

Nakatani, Y., Horikoshi, M., Brenner, M., Yamamoto, T., Besnard, F., Roeder, R.G., and Freese, E. (1990). A downstream initiation element required for efficient TATA box binding and in vitro function of TFIID. Nature *348*, 86–88.

Nicholls, A.J. (1993). GRASP Manual (New York: Columbia University).

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.

Nilges, M., Kuszewski, J., and Brünger, A.T. (1991). Sampling properties of simulated annealing and distance geometry. In Computational Aspects of the Study of Biological Macromolecules by Nuclear Magnetic Resonance Spectroscopy, J.C. Hoch, F.M. Poulsen, and C. Redfield, eds. (New York: Plenum Press), pp. 451–455.

Nishikawa, J., Kokubo, T., Horikoshi, M., Roeder, R.G., and Nakatani, Y. (1997). Drosophila $TAF_{II}230$ and the transcriptional activator VP16 bind competitively to the TATA box-binding domain of the TATA box-binding protein. Proc. Natl. Acad. Sci. USA *94*, 85–90.

Nissen, P., Kjeldgaard, M., Thirup, S., Polekhina, G., Reshetnikova, L., Clark, B.F., and Nyborg, J. (1995). Crystal structure of the ternary complex of Phe-tRNAPhe, EF-Tu, and a GTP analog. Science *270*, 1464–1472.

Nyborg, J., Nissen, P., Kjeldgaard, M., Thirup, S., Polekhina, G., Clark, B.F.C., and Reshetnikova, L. (1996). Structure of the ternary complex of EF-Tu: macromolecular mimicry in translation. Trends Biochem. Sci. *21*, 81–82.

O'Brien, T., and Tjian, R. (1998). Functional analysis of the human $TAF_{II}250$ N-terminal kinase domain. Mol. Cell *1*, 905–911.

Oelgeschläger, T., Tao, Y., Kang, Y.K., and Roeder, R.G. (1998). Transcription activation via enhanced preinitiation complex assembly in a human cell-free system lacking TAF_{II}s. Mol. Cell *1*, 925–931.

Ogryzko, V., Schiltz, R.L., Russanova, V., Howard, B.H., and Nakatani, Y. (1996). The transcriptional coactivators p300 and CBP are histone acetyltransferases. Cell *87*, 953–959.

Orengo, C.A., Brown, N.P., and Taylor, W.R. (1992). Fast structural alignment for protein databank searching. Proteins *14*, 139–167.

Orphanides, G., Lagrange, T., and Reinberg, D. (1996). The general transcription factors of RNA polymerase II. Genes Dev. *10*, 2657–2683.

Pascal, S.M., Muhandiram, D.R., Yamazaki, T., Forman-Kay, J.D., and Kay, L.E. (1994). Simultaneous acquisition of ^{15}N- and ^{13}C-edited NOE spectra of proteins dissolved in H_2O. J. Magn. Reson. B. *103*, 197–201.

Perez-Howard, G.M., Weil, P.A., and Beechem, J.M. (1995). Yeast TATA binding protein interaction with DNA: fluorescence determination of oligomeric state, equilibrium binding, on-rate, and dissociation kinetics. Biochemistry *34*, 8005–8017.

Radhakrishnan, I., Pérez-Alvarado, G.C., Parker, D., Dyson, H.J., Montminy, M.R., and Wright, P.E. (1997). Solution structure of the KIX domain of CBP bound to the transactivation domain of CREB: a model for activator:coactivator interactions. Cell *91*, 741–752.

Roeder, R.G. (1996). The role of general initiation factors in transcription by RNA polymerase II. Trends Biochem. Sci. *21*, 327–335.

Russo, A.A., Jeffrey, P.D., Patten, A.K., Massagué, J., and Pavletich, N.P. (1996). Crystal structure of the $p27^{Kip1}$ cyclin-dependent-kinase inhibitor bound to the cyclin A-Cdk2 complex. Nature *382*, 325–331.

Savva, R., and Pearl, L.H. (1995). Nucleotide mimicry in the crystal structure of the uracil-DNA glycosylase-uracil glycosylase inhibitor protein complex. Nat. Struct. Biol. *2*, 752–757.

Spolar, R.S., and Record, M.T.J. (1994). Coupling of local folding to site-specific binding of proteins to DNA. Science *263*, 777–784.

Tan, S., Hunziker, Y., Sargent, D.F., and Richmond, T.J. (1996). Crystal structure of a yeast TFIIA/TBP/DNA complex. Nature *381*, 127–134.

Uesugi, M., Nyanguile, O., Lu, H., Levine, A.J., and Verdine, G.L. (1997). Induced α helix in the VP16 activation domain upon binding to a human TAF. Science *277*, 1310–1313.

Verrijzer, C.P., Chen, J.-L., Yokomori, K., and Tjian, R. (1995). Binding of TAFs to core elements directs promoter selectivity by RNA polymerase II. Cell *81*, 1115–1125.

Vuister, G.W., and Bax, A. (1993). Quantitative J correlation: a new approach for measuring homonuclear three-bond J(HNHa) coupling constants in ^{15}N-enriched proteins. J. Am. Chem. Soc. *115*, 7772–7777.

Wishart, D.S., and Sykes, B.D. (1994). The ^{13}C chemical-shift index: a simple method for identification of protein secondary structure using ^{13}C chemical-shift data. J. Biomol. NMR *4*, 171–180.

Wüthrich, K., Billeter, M., and Braun, W. (1983). Pseudostructures for the 20 common amino acids for use in studies of protein conformations by measurements of intramolecular proton-proton distance constraints with nuclear magnetic resonance. J. Mol. Biol. *169*, 949–961.

Yamazaki, T., Lee, W., Revington, M., Mattiello, D.L., Dahlquist, F.W., Arrowsmith, C.H., and Kay, L.E. (1994). A suite of triple resonance NMR experiments for the backbone assignment of ^{15}N, ^{13}C, ^{2}H labeled proteins with high sensitivity. J. Am. Chem. Soc. *116*, 11655–11666.

Yang, X., Ogryzko, V., Nishikawa, J., Howard, B., and Nakatani, Y. (1996). A p300/CBP-associated factor that competes with the adenoviral oncoprotein E1A. Nature *382*, 319–324.

Zhou, P., Sun, L.J., Dötsch, V., Wagner, G., and Verdine, G.L. (1998). Solution structure of the core NFATC1/DNA complex. Cell *92*, 687–696.

Zwahlen, C., Legault, P., Vincent, S.J.F., Greenblatt, J., Konrat, R., and Kay, L.E. (1997). Methods for measurement of intermolecular NOEs by multinuclear NMR spectroscopy: application to a bacteriophage l N-peptide/*box*B RNA complex. J. Am. Chem. Soc. *119*, 6711–6721.

Brookhaven Protein Data Bank ID Code

The atomic coordinates for TBP-$dTAF_{II}230_{11–77}$ have been deposited with the Brookhaven Protein Data Bank under ID code 1tba.