*Proc. Natl. Acad. Sci. USA*
Vol. 90, pp. 8174–8178, September 1993
Biochemistry

# Crystal structure of yeast TATA-binding protein and model for interaction with DNA

DANIEL I. CHASMAN*†‡, KEVIN M. FLAHERTY†, PHILLIP A. SHARP*, AND ROGER D. KORNBERG†

*Center for Cancer Research, E17-529, Massachusetts Institute of Technology, Cambridge, MA 02139; and †Department of Cell Biology, Beckman Laboratories for Structural Biology, Stanford, CA 94305-5400

*Contributed by Roger D. Kornberg, June 7, 1993*

**ABSTRACT** **The C-terminal 179-aa region of yeast (*Saccharomyces cerevisiae*) TATA-binding protein (TBP), phylogenetically conserved and sufficient for many functions, formed crystals diffracting to 1.7-Å resolution. The structure of the protein, determined by molecular replacement with coordinates from *Arabidopsis* TBP and refined to 2.6 Å, differed from that in *Arabidopsis* slightly by an angle of about 12° between two structurally nearly identical subdomains, indicative of a degree of conformational flexibility. A model for TBP–DNA interaction is proposed with the following important features: the long dimension of the protein follows the trajectory of the minor groove; two rows of basic residues conserved between the subdomains lie along the edges of the protein in proximity to the DNA phosphates; a band of hydrophobic residues runs down the middle of the groove; and amino acid residues whose mutation alters specificity for the second base of the TATA sequence are juxtaposed to that base.**

TATA-binding protein (TBP) is unique among transcription factors in its involvement in initiation by all three eukaryotic RNA polymerases (1, 2). A component of initiation factors SL1 and TFIIIB, TBP is believed to make nonspecific contacts with the promoters for RNA polymerases I and III, respectively (1–4). In the case of RNA polymerase II transcription, TBP interacts with other proteins constituting the initiation factor TFIID and specifically recognizes the TATA element of the promoter (5, 6). Despite its specificity, TBP tolerates considerable variation in the TATA sequence and is required for transcription from "TATA-less" RNA polymerase II promoters as well (7–9). Remarkable also are the number and variety of interactions between TBP and activators and coactivators of transcription demonstrated *in vitro* (e.g., refs. 10 and 11). The unusual aspects and versatility of TBP reflected in its multiple roles in transcription call for structural analysis in atomic detail.

Genes for TBP have been cloned from many organisms (refs. 12–14; GenBank accession nos. M64861 and L07754), and the gene sequence shows conservation of a C-terminal domain of about 180 aa, 81% identical between yeast and man, divided in two directly repeated regions of about 40% identical sequence. By contrast, the N-terminal domain is widely divergent in both size and sequence, ranging from as few as 18 aa in *Arabidopsis* to about 160 in humans and 176 in *Drosophila*. The conserved C-terminal domain is alone sufficient for TBP function, both *in vitro* and *in vivo* (15, 16), and is therefore the most attractive candidate for structural analysis.

The crystal structure of TBP from *Arabidopsis* was recently solved at 2.6-Å resolution by the method of multiple isomorphous replacement (12). We have now used these results to solve the structure of the conserved C-terminal domain of the yeast protein (yTBP179) by molecular replacement.§ We suggest a model for binding, with the protein lying along the minor groove of the DNA, and thus with the long axis of the protein forming an angle with the long axis of DNA similar to the pitch angle of the DNA helix.

## MATERIALS AND METHODS

**Crystallization of TBP.** A vector for expression of yTBP179 was prepared by fusion of the initiator methionine codon to codon 62 in plasmid pET11a (Novagen) by the polymerase chain reaction. Recombinant yTBP179 was expressed in *Escherichia coli* BL21(DE3) (Novagen) and purified by conventional chromatography (unpublished work). Crystals for x-ray analysis grew in 3–5 days at room temperature in hanging drops containing 2–5 μl (20–50 μg) of protein solution and 1–5 μl of crystallization buffer [200 mM sodium potassium phosphate, monobasic/600 mM sodium potassium phosphate, dibasic, pH 7.25/750 mM NaCl/0.5% (vol/vol) PEG-400/1 mM dithiothreitol] (17).

**Crystallography.** High-resolution data were collected with a RAXIS II imaging plate (Molecular Structure, The Woodlands, TX) with Rigaku rotating-anode generators (CuKα) run at 50 kV and 100 mA. Imaging-plate data were processed with the software provided with the RAXIS II. The native data used for the molecular replacement were collected in 1.5° oscillations for 30 min each and a plate-to-crystal distance of 145 mm. Thirteen thousand seven hundred seventy-four unique reflections (52,281 full, 39,008 partial reflections total) representing 92% of the total possible and having $R_{sym}$ ($F > 1\sigma F$, 20.0–2.6 Å) of 6.84% were included in the final list of averaged reflections.

Molecular replacement was performed entirely with XPLOR (18). The strongest solution to the rotation function had a value 0.4 SD above the next highest solution and 4.6 SD above the average signal and remained the strongest solution after Patterson correlation (PC) refinement of 103 strong candidate solutions (19). The solution to the translation function was also clear (see text). Refinement of the model was also performed in XPLOR. FRODO was used for manual rebuilding of the model (20). INSIGHT II (Biosym Technologies, San Diego) and MOLSCRIPT were used for the graphic displays of the model (21).

## RESULTS

**Crystallization and Structure Determination.** We prepared large amounts of recombinant proteins in *E. coli* corresponding to the full-length yeast TBP and yTBP179, a truncated form consisting of the C-terminal 179 aa fused to an initiator methionine. Despite an intensive effort, no crystals were obtained with the full-length yeast protein. Sequence hetero-

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviation: TBP, TATA-binding protein.

‡To whom reprint requests should be sent at the * address.

§The atomic coordinates have been deposited in the Protein Data Bank, Chemistry Department, Brookhaven National Laboratory, Upton, NY 11973 (reference 1TBP).

geneity of the N-terminal domain across species and its disorder in the crystals of the *Arabidopsis* TBP indicate that it may not have a discrete structure in the absence of interactions with other factors and may have interfered with crystallization of the full-length protein. Trigonal crystals grown from phosphate/NaCl solution were found for the truncated protein. The crystals were in the space group $P3_121$ or $P3_221$ with unit cell parameters $a = b = 104.1$ Å, $c = 75.5$ Å and diffracted x-rays ($\lambda = 1.08$ Å, Stanford Synchrotron Radiation Laboratory) to 1.7 Å.

A comprehensive search for isomorphous heavy atom derivatives identified several compounds capable of interacting with the crystals, as judged from changes in crystal color and in the relative intensity of diffracted spots, but no clear signals in difference Patterson maps were obtained. While our search for heavy-atom derivatives was in progress, the structure of TBP from *Arabidopsis* was reported, and the atomic coordinates of this structure were kindly provided by Stephen Burley (12). The rotation function, PC refinement, and translation functions for molecular replacement were performed in XPLOR (18) with the coordinates of the two molecules found in the asymmetric unit of *Arabidopsis* TBP crystals. The translation function unambiguously discriminated between the two alternatives for space-group assignment, having a maximal value in $P3_121$ 10.3 and 8.2 SD greater than the maximum in $P3_221$ and the second highest maximum in $P3_121$, respectively, and 17.1 SD greater than the average maximum in either space group. The packing of the TBP dimer in the best solution to the translation function showed no serious steric clashes and had a calculated solvent content of 62%. The *R* factor was 45.8% (8.0–4.0 Å), and could be reduced to 37.4% (8.0–4.0 Å) by rigid-body minimization of, successively, (*i*) the whole dimer, (*ii*) each monomer, and (*iii*) each domain of each monomer (aa 19–110 and 111–198 of *Arabidopsis* TBP). Powell minimization further reduced the *R* factor to 27.2%. The residues differing between the *Arabidopsis* and yeast sequences were substituted in the model for their *Arabidopsis* counterparts and the structure was manually repositioned, using FRODO, into electron density generated by simulated annealing "omit" maps (XPLOR). All omit maps except for those representing aa 35–55 and 90–110 of one monomer gave clear electron density for the peptide backbone. Corresponding regions of the second monomer were well ordered, confirming the trace of the model for the *Arabidopsis* protein and the similarity of the structure of yeast TBP. Manual rebuilding was followed by another round of Powell minimization and then both steps were repeated, bringing the *R* factor to 23%. Simulated annealing refinement of these coordinates followed by Powell positional refinement, overall *B*-factor refinement, and tightly restrained refinement of individual *B* factors reduced the *R* factor (6.0–2.6 Å) to a best value of 21.1%. In our model, bond angles deviate (rms) by 1.576° and bond lengths by 0.008 Å from ideal geometry, and *B* factors for neighboring atoms differ by 2.02 Å$^2$. While not imposed during the refinement, a noncrystallographic relationship between the two monomers can be described by a 175° rotation.

**Description of the Structure.** The well-defined maximum in the translation function and the success of the simple refinement scheme indicated that the trace of the conserved C-terminal domain of yeast TBP would have a topology very similar to that of *Arabidopsis* TBP (12). Briefly, the structure consists of two subdomains with a high degree of structural similarity related by a pseudo-twofold axis of symmetry. Each subdomain corresponds to one of the direct repeats identified in the amino acid sequence of the protein. Together, the two subdomains form a continuous, 10-stranded, antiparallel β-sheet that is slightly curved. Two loops emanate from the ends of the molecule between strands S2 and S3 (and between S2′ and S3′) and, together with the curved β-sheet, form a cylindrical cleft. Two long helices (H2 and H2′) are cradled on the convex surface of the sheet, and two short helices (H1 and H1′) are found at either end of the molecule, preceding strand S2 (S2′) in the overall trace. An extended chain connecting the two subdomains and corresponding to part of the basic region in the sequence between them is the only structural element that is not represented in both subdomains, and is the only structurally asymmetric part of the C-terminal domain of TBP (Fig. 1).

The refined structure of the yeast protein gives an indication of the conformational variability of TBP. The two monomers of *Arabidopsis* TBP were found to have nearly identical conformations (0.5 Å rms, Cα positions). The two monomers of the yeast protein were slightly less similar to each other (1.1 Å rms), and both adopted slightly different conformations from the *Arabidopsis* TBP. The conformational difference appears to be a twist propagated through the whole β-sheet, together with a splaying apart of the two subdomains. The magnitude of this combined motion can be estimated for any pair of monomers by aligning corresponding subdomains separately and then determining the magnitude of a single rotation relating the two alignments. The value ranges from a maximum of 11.8° for the comparison of the first yeast monomer with the second *Arabidopsis* monomer, to 8.7° for the two yeast monomers, to a minimum of 2.6° for the alignment of the two *Arabidopsis* monomers (Fig. 2).

Additional, local conformation variability was also found. Amino acids 146–148 within the extended chain following the long helix in the first monomer, H2, appear to be capable of adopting alternative conformations. The short loop between H1 and S2 is poorly ordered in one of the monomers of yeast TBP. Amino acids 94–97 in the first half of the loop between S2 and S3 in one of the monomers are also poorly ordered.

**Structure–Function Relationships.** For interpreting the structure, it proved useful to relate the disposition of amino acids conserved and not conserved between the two subdomains to TBP functions that must have a symmetric or an asymmetric structural basis. For example, two types of sequence conservation may be considered: conservation between the symmetry-related subdomains, which must pertain to functions in which the subdomains play identical roles, and conservation across species, especially evolutionary conservation of residues that differ between the subdomains

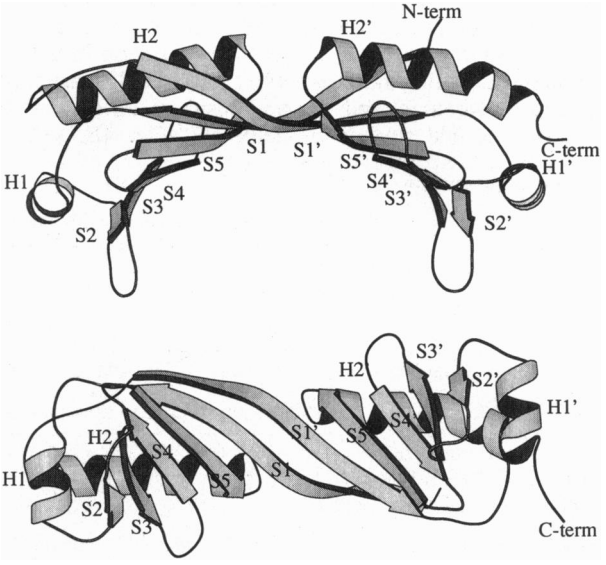


FIG. 1. C-terminal domain of TBP from yeast: Two views showing elements of secondary structure. Names of elements of secondary structure follow the convention of ref. 12 and are indicated with respect to the amino acid sequence in Fig. 3.

FIG. 2. Stereoview showing least-squares alignment of Cα positions for a monomer of the C-terminal domain of TBP from yeast (white) and a monomer of TBP from *Arabidopsis* (black).

and that are therefore likely to underly asymmetric functions in which one subdomain or the other is involved. We investigated whether residues conserved between the two subdomains formed the basis of the subdomain structure by examining the general chemical properties of these residues and their tendency to be buried in the hydrophobic core of TBP. These residues are both hydrophilic (20 residues) and hydrophobic (32 residues), and on average, are no more likely to be solvent-inaccessible than other residues from the protein (see Fig. 4*A*; blue, solvent-accessible; yellow, solvent-inaccessible). Using the method of Lee and Richards (39), we estimate that 73% of the amino acids conserved between the two subdomains are solvent-accessible, compared with 68% for the structure as a whole. The solvent-accessible residues conserved between the two subdomains are strongly localized to the concave side of the β-sheet lining the cylindrical cleft. Mutagenesis of these solvent-accessible residues leads almost invariably to disruption of DNA binding and defines this surface of the molecule as the one most likely to bind DNA (22). Interactions between TBP and the sugar-phosphate backbone of DNA would be expected to be symmetric and to underly the conservation between the two symmetry-related subdomains; these interactions appear to be a stronger constraint on the amino acid sequence than those in the hydrophobic core. Similarly, since many amino acids in the hydrophobic core differ between the two subdomains, their involvement in the conserved structure must be explained by specific covariation during evolution (Figs. 3 and 4) (23).

The second type of sequence homology, conservation across species (Figs. 3 and 4), includes most of the residues conserved between the two subdomains of the yeast sequence, as well as residues conserved in only one subdomain. Displayed on the structure, the amino acids defined by this homology form the concave surface of the β-sheet, including the loops between S2 and S3, and also a region in the first subdomain including the loop between H1 and S2, part of S1′, the loop between S3 and S4, and residues on or near the solvent-accessible surface of helix H2 (Fig. 4*B*, yellow). Sequence conservation between yeast and *Arabidopsis* TBPs is representative of that found across all species (Fig. 3).

Residues conserved across species but not between the two subdomains are of particular interest because of their likely involvement in asymmetric functions, such as interactions of TBP with other proteins in transcription complexes or sequence-specific interactions of TBP with DNA. Among the residues conserved in only one subdomain that have been studied, Lys → Leu mutations in residues 138 and 145 result in temperature-sensitive cell growth and an altered interaction with TFIIA *in vitro* (24). A suppressor of a double mutant at positions 133 and 138 was found in TDS4, a TFIIB homolog involved in polymerase III transcription (25). The double mutation of residue 205 (conserved in one subdomain) from Leu to Val and symmetrically conserved residue 194 from Ile to Phe allows yeast TBP to recognize the sequence TGTA as well as the consensus TATA. Analogous mutations in human TBP have a comparable effect on binding specificity (26). Finally, residues 93–143 are sufficient for interaction of TBP with adenoviral E1A protein (10).

The functional significance of the residues conserved in one subdomain or symmetrically conserved in both subdomains may be further revealed by the sequences of a candidate TBP from *Plasmodium* and of a TBP-like factor (TRF) from *Drosophila* (13, 14). Inclusion of these sequences in the alignment with all others reduces the fraction of residues conserved among all sequences from 59% to 26%, with the elimination of the asymmetric patterns of conservation in favor of patterns more closely resembling the conservation between the two subdomains. Residues that remain conserved in only one subdomain in this more stringent comparison may have particular functional importance (Fig. 3).

**Protein–DNA Interaction.** A plausible model for TBP–DNA interaction may be derived from existing information, though details must await the structure determination of cocrystals. Mutagenesis studies identify many of the solvent-accessible residues on the concave side of the β-sheet, conserved between the two subdomains, as important in DNA binding, so this region of TBP is, in all likelihood, its DNA-binding surface. Interactions between TBP and the sugar-phosphate

```
#in S. c. sequence      61       70        80        90       100       110       120       130       140       150
                        .        .         .         .        .   #     #         #     #   .         .         .
S. c. Protein           SGIVPTLQNIVATVTLGCRLDLKTVAL HARNAEYNPKRFAAVIMRIREPKTTALIFASGKMVVTGAKSEDDSKLASRKYARIIQKIGFAA
A. t. Protein           ----------S--N-D-K----AI-- Q---------------------------------C-------F--M-A------V--L--P-
Cons. in Two Domains          ****          *   *  *     * *  *  * *   **   *** *** * ****       *
Cons. in Closely Rel.    * * ***** **     * **  **  ********************  ***** *****  * ***  *  * *****  **  * *
Cons. in All               **           *        ** ** *      * *   *  ***** *      *             * *
                         S1/S1'          H1/H1' S2/S2'   S3/S3'    S4/S4'   S5/S5'        H2/H2'
Secondary Structure       ------------    -------- ---     -----     ----    -------  -----------------
#in S. c. sequence    151       161       171       181       191       201       211       221       231       241
                        .        .         .         .        .   #     #         #     #   .         .         .
S. c. Protein           KFTDFKIQNIVGSCDVKFPIRLEGLAFSHGTFSSYEPELFPGLIYRMVKPKIVLLIFVSGKIVLTGAKQREEIYQAFEAIYPVLSEFRKM*
A. t. Protein           --K----------------------Y--AA----------------K---------------I----M-D-T-K---N----------IQQ*
Cons. in Two Domains          ****          *   *  *     * *  *  * *   **   *** *** * ****       *
Cons. in Closely Rel.    *  *  ** * **** *** ** *   *     **** ****** *  *   ******** * ****       **  * * *    *
Cons. in All             *      *          ** ** *   *     **** * ** *          ****       ** *          *   *
```

FIG. 3. Sequence conservation in the C-terminal domain of TBP with alignment of the two domains (*Upper* and *Lower*). *S.c.*, *Saccharomyces cerevisiae*; *A.t.*, *Arabidopsis thaliana*. Identically conserved residues in each category are indicated by stars. "Cons. in Closely Rel" (conserved in closely related) refers to conservation among sequences from human, *Drosophilia, Dictyostelium, Acanthamoeba, Schizosaccharomyces pombe, Saccharomyces cerevisiae*, wheat, potato, *Caenorhabditis elegans*, and maize (see text for references). "Cons. in All" refers to amino acids conserved among the closely related sequences together with the sequence for TBP from *Plasmodium* and TBP-like protein TRF from *Drosophila*. Basic amino acids largely conserved between the two domains and among all sequences are indicated (#).

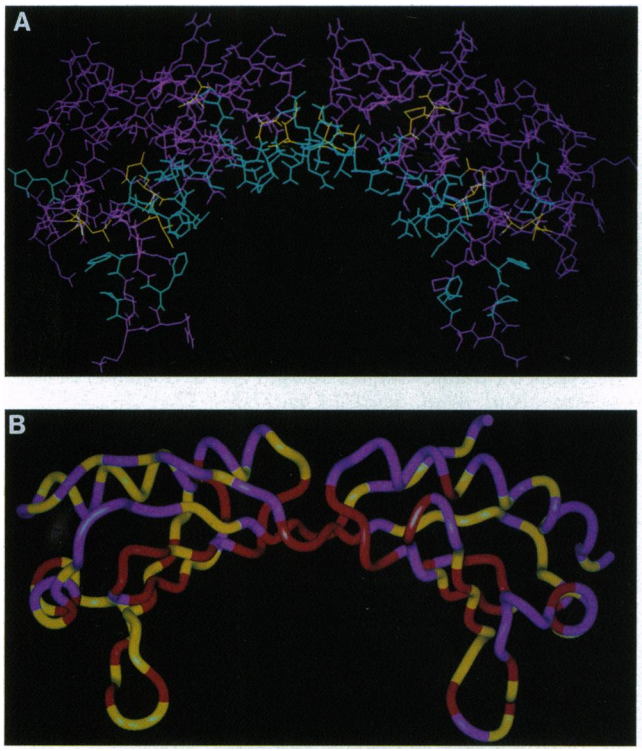


FIG. 4. (*A*) Stick diagram in stereo showing solvent-accessible (blue) and solvent-inaccessible (yellow) residues conserved between the two domains of TBP from yeast. Residues not conserved are purple. (*B*) Ribbon diagram indicating residues conserved between the two domains of TBP from yeast (red), conserved among closely related sequences but not between the two domains (yellow; see Fig. 3), and not conserved (purple).

backbone of DNA may be presumed to involve the two subdomains in a symmetric fashion, calling for alignment of the pseudo-dyad axis of the protein with a dyad of the DNA. In one model, TBP interacts with the DNA through the concave side of the β-sheet and is aligned with the long dimension of the protein perpendicular to the long axis of the DNA (12). We suggest an alternative and more specific model in which TBP is rotated about a coincident dyad with the DNA, allowing it to wrap around the DNA following the trajectory of the minor groove (Fig. 5 *A* and *B*). In this position, the 10 strands of β-sheet make nearly exclusive contact with the minor groove, consistent with biochemical characterization of the complex (27, 28) and suggestions from sequence comparisons (31, 32).

Four further observations support our model in detail. (*i*) There is an excellent fit of TBP along its entire length to the minor groove. After the relative orientation of protein and DNA in the model was determined by manual positioning, rigid body minimization of van der Waals energy in XPLOR resulted in only minor adjustment (<3–6°).

(*ii*) Four basic residues that are conserved between the two subdomains (105, 110, 120, and 127 in subdomain 1; 196, 201, 211, and 208 in subdomain 2), and also well conserved across species, lie along two rows circumscribing the presumptive DNA-binding cleft of TBP and, in our model, follow the sugar-phosphate backbones of the two DNA strands (Fig. 5 *A* and *B*). An additional basic residue (107 in subdomain 1 and 199 in subdomain 2) is also conserved, but less so. The model thus resembles the structures of many other DNA-binding proteins whose basic residues make sequence-independent contacts with DNA phosphates (29).

(*iii*) An essentially uninterrupted band of hydrophobic residues runs down the center of the DNA-binding surface, between the two rows of basic (and other hydrophilic) residues. This arrangement of hydrophobic and hydrophilic residues reinforces the argument for a uniform pattern of contacts between TBP and DNA on the concave side of the β-sheet. A uniform pattern of conservation is a natural attribute of a model in which the protein lies entirely within the minor groove rather than across both sugar-phosphate backbones of the DNA (Fig. 5*C*) (12).

(*iv*) Hydroxyl radical footprints exhibit a dyad between the fourth and fifth base pairs of the consensus TATAAA sequence, whereas phosphate ethylation protection analysis suggests a dyad 1 bp 3′ to this location (27, 28). Remarkably, for the alignment of either dyad with the pseudo-dyad of TBP, the residues whose mutation causes a change in specificity for the second base pair of the consensus sequence are juxtaposed with that very base pair (26). Mutations that result in a dominant negative allele of TBP, presumably by disrupt-

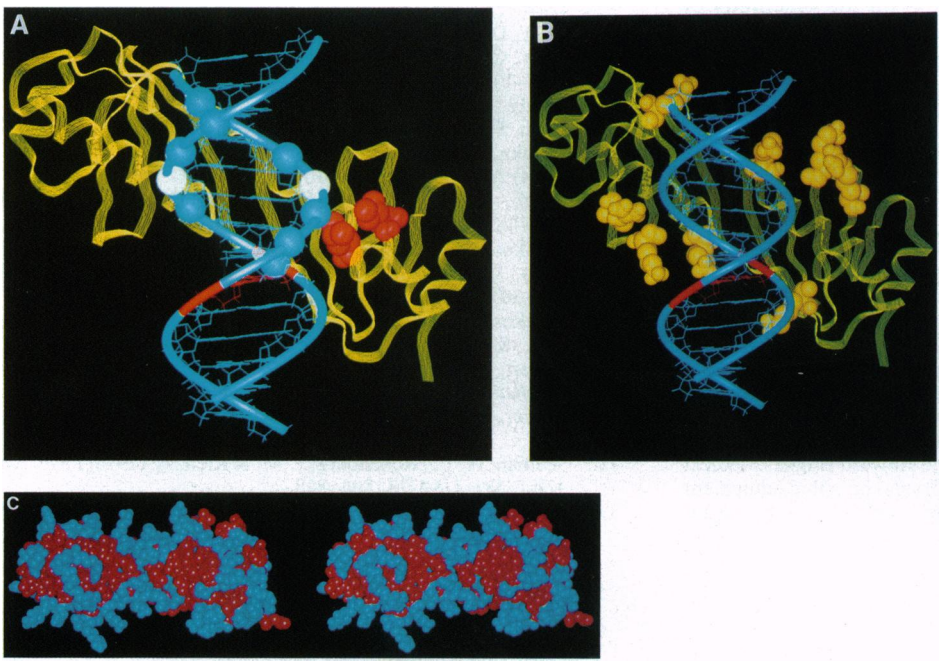


FIG. 5. Views of a model for the interaction of TBP and DNA. (*A*) TBP positioned in the minor groove of B-form DNA with the sequence 5′-GGGCTATAAAAGGG-3′. Phosphates whose ethylation interferes with TBP binding are indicated by spheres along the ribbon backbone of the DNA (27). The symmetry-related, central phosphates in the interference pattern are white and bracket the coincident dyad of TBP and the DNA. Mutations that allow TBP to bind the sequence "TGTA" in addition to the consensus "TATA" are indicated by space-filling representation (red) on the ribbon of TBP (yellow). (*B*) The model omitting the phosphates but showing the four basic residues conserved in each domain (yellow). (*C*) Stereoview of the isolated TBP monomer in space-filling representation looking into the presumptive DNA-binding cleft and showing the distribution of hydrophobic (red) and hydrophilic (blue) residues.

ing DNA binding but not interactions with other proteins, closely bracket the specificity mutation and further define the region of the protein making essential contacts with DNA (30). Finally, positioning TBP in this way can accommodate the one exception to the apparent restriction of protein–DNA contacts to the minor groove: the reported interference with TBP binding to DNA by methylation at the N7 position of guanine might be explained by contacts between amino acids in strand S1 reaching around the sugar-phosphate backbone into the major groove (27, 28).

## DISCUSSION

The structure determination of the C-terminal domain of yeast TBP reported here has confirmed the previous results on the *Arabidopsis* protein, given an indication of the degree of conformational flexibility about the central pseudo-dyad, and provided detailed information about the disposition of residues that are not conserved between the two proteins. We have presented a model for TBP–DNA interaction whose broad features are likely to be correct but which may require revision in detail. In particular, the structure of DNA in the complex must differ somewhat from the canonical B form used in our model. Gel mobility-shift analyses have shown that DNA is bent in the complex. If the bent DNA cannot lie in a single plane, then a degree of untwisting must occur as well, since the sum of writhing and untwisting in the complex is nearly zero (33, 34). There are many reasons to suspect that TBP may participate in DNA unwinding for transcription. First, $\sigma^{70}$, which resembles TBP in several respects (35), is involved in DNA melting by *E. coli* RNA polymerase holoenzyme. Second, TBP has been shown to bind single-stranded DNA with an affinity only 1 order of magnitude less and 2 orders of magnitude greater than those for the TATA element and nonconsensus double-stranded DNA, respectively. Third, the structure of TBP resembles those of two other proteins known to bind single-stranded nucleic acids, the RNA-binding-domain protein and bacteriophage T4 gene 5 protein, both of which contain extensive surfaces of antiparallel $\beta$-sheet (36, 37). Insofar as the mode of TBP–DNA interaction resembles the nucleic acid-binding interactions of these two proteins, TBP may enforce a partially single-stranded character of its binding site or facilitate unwinding in a subsequent step.

Finally, attributes of the yeast structure and our model for its interaction with DNA could account for the specificity of TBP binding *in vitro* and at polymerase II promoters, despite a paucity of features in the minor groove for discrimination of specific bases (38). Even partial unwinding of the DNA might facilitate contacts between the hydrophobic residues running down the center of the DNA-binding surface and the bases. And the conformational flexibility of the protein seen in the crystal structure might contribute not only to an unwinding interaction but also to the positioning of sequence-specific contacts between the protein and the DNA.

D.I.C. is particularly grateful to Carl Pabo for encouragement, discussions, and generous provision of x-ray facilities. We thank Mark Rould, Brigitte Raumann, Joel Pomerantz, Dan Bancroft, Stephen Harrison, and Xiaohua Zhang for helpful discussions. Coordinates for the *Arabidopsis* structure were generously provided by Stephen Burley. David McKay provided advice in the initial stages of the project. Support for the crystallographic equipment in Carl Pabo's laboratory was provided by the Howard Hughes Medical Institute and the Pew Charitable Trust. Crystallographic equipment in David McKay's laboratory was supported in part by Grant GM3992 from the National Institutes of Health. This work was supported by Grant DRG-1112 from the Damon Runyon–Walter Winchell Cancer Foundation (D.I.C.) and Grants AI21144 (R.D.K.), P01-CA42063 (P.A.S.), and P30-CA14051 (P.A.S.) from the National Institutes of Health.

1. Sharp, P. A. (1992) *Cell* **68**, 819–821.
2. Rigby, P. W. J. (1993) *Cell* **72**, 7–10.
3. Taggart, A. K., Fisher, T. S. & Pugh, B. F. (1992) *Cell* **71**, 1015–1028.
4. Huet, J. & Sentenac, A. (1992) *Nucleic Acids Res.* **20**, 6451–6454.
5. Dynlacht, B. D., Hoey, T. & Tjian, R. (1991) *Cell* **66**, 563–576.
6. Timmers, H. T., Meyers, R. E. & Sharp, P. A. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 8140–8144.
7. Hahn, S., Buratowski, S., Sharp, P. A. & Guarente, L. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 5718–5722.
8. Wobbe, C. R. & Struhl, K. (1990) *Mol. Cell. Biol.* **10**, 3859–3867.
9. Weis, L. & Reinberg, D. (1992) *FASEB J.* **6**, 3300–3309.
10. Lee, W. S., Kao, C. C., Bryant, G. O., Liu, X. & Berk, A. J. (1991) *Cell* **67**, 365–376.
11. Lee, D. K., Dejong, J., Hashimoto, S., Horikoshi, M. & Roeder, R. G. (1992) *Mol. Cell. Biol.* **12**, 5189–5196.
12. Nikolov, D. B., Hu, S.-H., Lin, J., Gasch, A., Hoffmann, A., Horikoshi, M., Chua, N.-H., Roeder, R. G. & Burley, S. K. (1992) *Nature (London)* **360**, 40–46.
13. McAndrew, M. B., Read, M., Sims, P. F. G. & Hyde, J. E. (1993) *Gene* **124**, 165–171.
14. Crowley, T. E., Hoey, T., Liu, J.-K., Jan, Y. N. & Tjian, L. Y. J. R. (1993) *Nature (London)* **361**, 557–561.
15. Kelleher, R. J., Flanagan, P. M., Chasman, D. I., Ponticelli, A. S., Struhl, K. & Kornberg, R. D. (1992) *Genes Dev.* **6**, 304–315.
16. Cormack, B. P., Strubin, M., Ponticelli, A. S. & Struhl, K. (1991) *Cell* **65**, 341–348.
17. McPherson, A. (1985) *Methods Enzymol.* **114**, 112–120.
18. Brünger, A. T. (1992) XPLOR *Manual* (Harvard Univ./Yale Univ.), v. 3.0.
19. Brünger, A. T. (1990) *Acta Crystallogr. Sect. A* **46**, 46–57.
20. Jones, T. A. (1978) *J. Appl. Crystallogr.* **11**, 268–272.
21. Kraulis, P. J. (1991) *J. Appl. Crystallogr.* **24**, 946–950.
22. Yamamoto, T., Horikoshi, M., Wang, J., Hasegawa, S., Weil, P. A. & Roeder, R. G. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 2844–2848.
23. Lim, W. A. & Sauer, R. T. (1989) *Nature (London)* **339**, 31–36.
24. Buratowski, S. & Zhou, H. (1992) *Science* **255**, 1130–1132.
25. Buratowski, S. & Zhou, H. (1992) *Cell* **71**, 221–230.
26. Strubin, M. & Struhl, K. (1992) *Cell* **68**, 721–730.
27. Lee, D. K., Horikoshi, M. & Roeder, R. G. (1991) *Cell* **67**, 1241–1250.
28. Starr, D. B. & Hawley, D. K. (1991) *Cell* **67**, 1231–1240.
29. Pabo, C. O. & Sauer, R. T. (1992) *Annu. Rev. Biochem.* **61**, 1053–1095.
30. Reddy, P. & Hahn, S. (1991) *Cell* **65**, 349–357.
31. Nash, H. A. & Granston, A. E. (1991) *Cell* **67**, 1037–1038.
32. Tanaka, I., Appelt, K., Dijk, J., White, S. W. & Wilson, K. S. (1984) *Nature (London)* **310**, 376–381.
33. Lorch, Y. & Kornberg, R. D. (1993) *Mol. Cell. Biol.* **13**, 1872–1875.
34. Horikoshi, M., Bertuccioli, C., Takada, R., Wang, J., Yamamoto, T. & Roeder, R. G. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 1060–1064.
35. Horikoshi, M., Wang, C. K., Fujii, H., Cromlish, J. A., Weil, P. A. & Roeder, R. G. (1989) *Nature (London)* **341**, 299–300.
36. Kenan, D. J., Query, C. C. & Keene, J. D. (1991) *Trends Biochem. Sci.* **16**, 214–220.
37. Brayer, G. D. & McPherson, A. (1984) *Biochemistry* **23**, 340–349.
38. Seeman, N. C., Rosenberg, J. M. & Rich, A. (1976) *Proc. Natl. Acad. Sci. USA* **73**, 804–808.
39. Lee, B. & Richards, F. M. (1971) *J. Mol. Biol.* **55**, 379–400.