Downloaded from rstb.royalsocietypublishing.org

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

X-Ray Crystallographic Studies of Eukaryotic Transcription Initiation Factors

Stephen K. Burley

Phil. Trans. R. Soc. Lond. B 1996 **351**, 483-489
doi: 10.1098/rstb.1996.0046

Email alerting service Receive free email alerts when new articles cite this article - sign up in the box at the top right-hand corner of the article or click **here**

To subscribe to *Phil. Trans. R. Soc. Lond. B* go to: **http://rstb.royalsocietypublishing.org/subscriptions**

This journal is © 1996 The Royal Society

Downloaded from rstb.royalsocietypublishing.org

X-ray crystallographic studies of eukaryotic transcription initiation factors

STEPHEN K. BURLEY

Laboratories of Molecular Biophysics and Howard Hughes Medical Institute, The Rockefeller University, 1230 York Avenue, New York 10021, U.S.A.

SUMMARY

TATA box-binding protein (TBP) is required by all three eukaryotic RNA polymerases for correct initiation of transcription of ribosomal, messenger, small nuclear and transfer RNAs. Since the first gene encoding a TBP was cloned, it has been the object of considerable biochemical and genetic study. Substantial progress has also been made on structural and mechanistic studies, including our three-dimensional crystal structures of TBP, TBP bound to a consensus TATA elements, and the ternary complex of transcription factor IIB (TFIIB) recognizing TBP bound to a TATA element. The structure of apo TBP was determined at 2.1 Å resolution. This highly symmetric α/β structure represents a new DNA-binding fold, which resembles a molecular 'saddle' that sits astride the DNA. The DNA-binding surface is a novel curved, antiparallel β-sheet. The structure of TBP complexed with the TATA element of the Adenovirus major late promoter was determined at 1.9 Å resolution. Binding of the protein induces a dramatic conformational change in the DNA, by tracking the minor groove and inducing two sharp kinks at either end of the sequence TATAAAAG. Between the kinks, the right-handed double helix is smoothly curved and partly unwound, presenting a widened minor groove to TBP's concave, antiparallel β-sheet. Side chain–base interactions are completely restricted to the minor groove, and include hydrogen bonds, van der Waals contacts and phenylalanine-base stacking interactions. The structure of a TFIIB/TBP/TATA element ternary complex was determined at 2.7 Å resolution. Core TFIIB resembles cyclinA, and recognizes the preformed TBP-DNA complex via protein–protein and protein–DNA interactions. The N-terminal domain of core TFIIB forms the downstream surface of the ternary complex, where it could fix the transcription start site. The remaining surfaces of TBP and the TFIIB can interact with TBP-associated factors, other class II initiation factors, and transcriptional activators and coactivators.

1. INTRODUCTION

Eukaryotes have three distinct RNA polymerases (forms I, II, and III) that catalyse transcription of nuclear genes (Sentenac 1985). Despite their structural complexity, these multisubunit enzymes require sets of auxiliary proteins known as general transcription initiation factors to initiate transcription from corresponding class I, II, and III nuclear gene promoters (Gabrielson & Sentenac 1991; Roeder 1991; Reeder 1992; Maldonado & Reinberg 1995). TATA box-binding protein (TBP), first identified as a component of the class II initiation factor TFIID, participates in transcription by all three nuclear RNA polymerases (Nikolov & Burley 1994). Thus TBP is the first universal transcription initiation factor component (a situation formally analogous to that of essential subunits common to the three RNA polymerases).

TBP's role in transcription initiation and its regulation is best understood for genes transcribed by RNA polymerase II (Roeder 1991; Maldonado & Reinberg 1995) (figure 1).

In this setting, TBP is tightly associated with other polypeptides known as TBP-associated factors or TAFs (reviewed in Nikolov & Burley 1994). This multiprotein complex (TFIID) is a general initiation factor (Matsui *et al.* 1980) that binds to the TATA element, coordinating accretion of class II initiation factors (TFIIA, -B, -D, -E, -F, -G/J, -H, -I) and RNA polymerase II (pol II) into a functional preinitiation complex (PIC) (Roeder 1991; Zawel & Reinberg 1993). Although incapable of mimicking TFIID *in vivo* (Roeder 1991), recombinant TBP alone is competent for PIC assembly and basal or core promoter dependent transcription in the presence of the other general class II factors (Buratowski *et al.* 1989). TBP engages in physical and functional interactions with the general initiation factors TFIIA and TFIIB, the C-terminus of the large subunit of pol II, some negative cofactors (NC1, NC2, DR1) that inhibit PIC formation, some transcriptional activators, and an initiator-binding factor (TFII-I) that may be important for transcription initiation from TATA-less promoters. TFIIB is the second general transcription factor to enter the PIC, creating a TFIIB-TFIID(TBP)-DNA platform that is in turn recognized by a complex of pol II and TFIIF (pol/F). *In vitro* studies with negatively supercoiled templates demonstrated that transcription initiation can be reconstituted with TBP, TFIIB and pol II, suggesting that together TBP and TFIIB position pol

Phil. Trans. R. Soc. Lond. B (1996) **351**, 483–489
Printed in Great Britain

© 1996 The Royal Society

The Royal Society is collaborating with JSTOR to digitize, preserve, and extend access to *Philosophical Transactions: Biological Sciences.* www.jstor.org

Downloaded from rstb.royalsocietypublishing.org

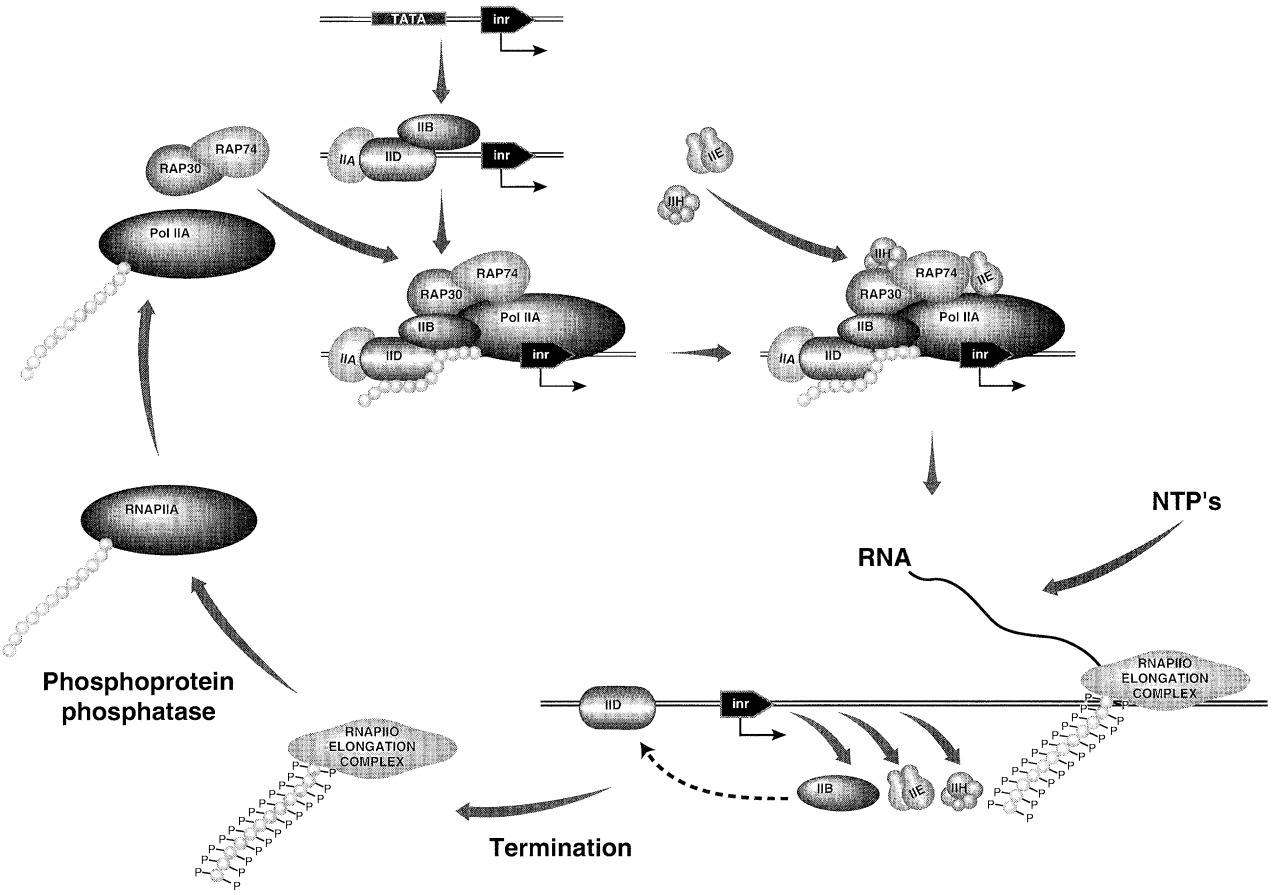


Figure 1. Preinitiation complex assembly begins with TFIID recognizing the TATA element, followed by coordinated accretion of TFIIA, TFIIB, the nonphosphorylated form of pol II and TFIIF (RAP30/RAP74), TFIIE and TFIIH. Before elongation pol II is phosphorylated by TFIIH. After termination, a phosphatase recycles pol II to its non-phosphorylated form, allowing the enzyme to reinitiate transcription. TBP (and TFIID) binding to the TATA box is an intrinsically slow step, yielding a long-lived protein-DNA complex. Efficient reinitiation of transcription can be achieved if recycled pol II re-enters the preinitiation complex before TFIID dissociates from the core promoter. Reproduced with permission (Zawel *et al.* 1993).

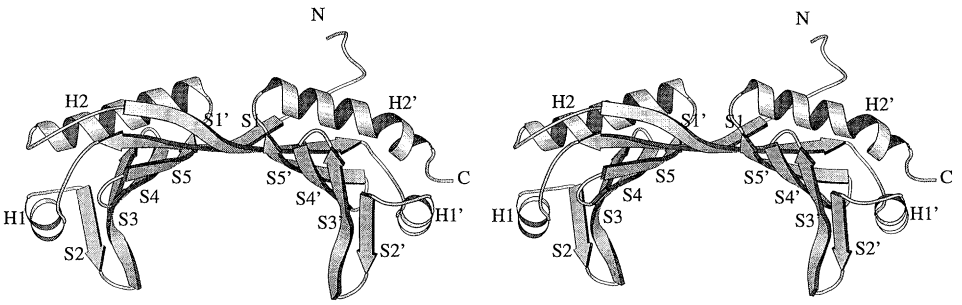


Figure 2. MOLSCRIPT (Kraulis 1991) stereodrawing of the structure of TBP2 viewed perpendicular to the internal pseudodyad axis. The N and C termini of the protein are indicated. The α-helices are shown as ribbon spirals and labelled H, the β-strands are shown as ribbon arrows (S), and loops and turns are drawn as double lines. The symbol ′ refers to the second structural domain or repeat. Reproduced with permission (Nikolov *et al.* 1992).

II (Parvin & Sharp 1993). Mutants of TFIIB alter pol II start sites in yeast, providing compelling evidence for its function as a precise spacer/bridge between TBP and pol II on the core promoter that determines the transcription start site. *In vivo* and under different conditions *in vitro*, pol II transcription initiation depends on TFIIE, TFIIF and TFIIH, and possibly TFIIA. Once PIC assembly is complete and in the presence of nucleoside triphosphates, strand separation at the transcription start site occurs to give an open complex, the C-terminal domain of the large subunit of pol II is phosphorylated, and pol II initiates transcription and is released from the promoter. During elongation, *in vitro* TFIID can remain bound to the core promoter supporting rapid reinitiation of transcription by pol II and the other general factors (figure 1) (Zawel *et al.* 1995). Core promoter binding by the TBP subunit of TFIID is an intrinsically slow step, because of the dramatic DNA deformation induced in the TATA element (Kim & Burley 1994). An abbreviated PIC assembly mechanism has also been proposed, following recent discoveries of various pol II

Downloaded from rstb.royalsocietypublishing.org

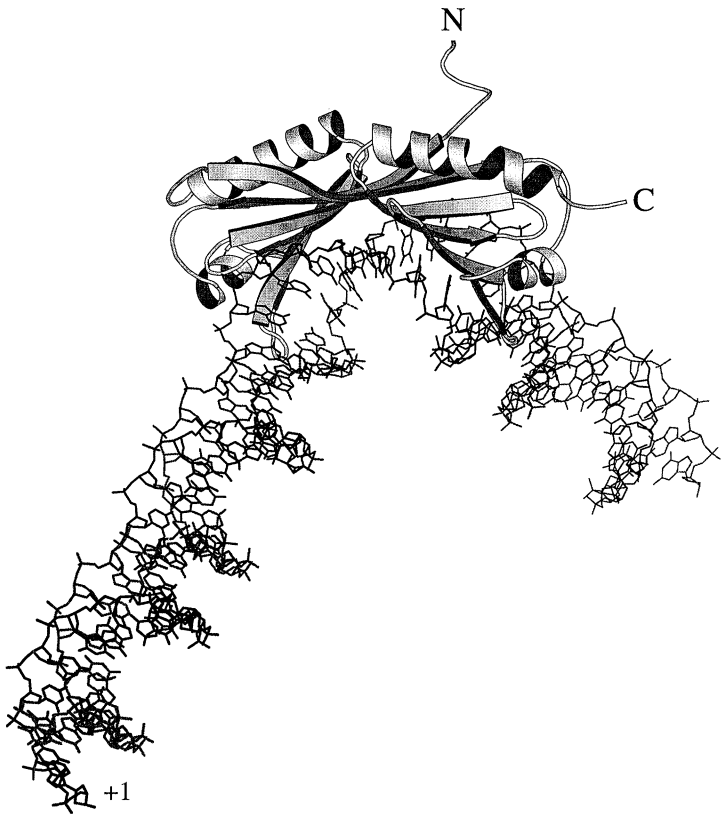


Figure 3. Three-dimensional structure of TBP2 from *Arabidopsis thaliana* complexed with the AdMLP TATA element. The molecular saddle (N- and C-termini labelled) is depicted with a ribbon drawing and the DNA is shown as a stick figure with the transcription start site labelled with +1. When TBP recognizes the minor groove of the TATA element, the DNA is kinked and unwound to present the minor groove edges of the bases to the underside of the molecular saddle. The coding strand is denoted with solid bonds. Reproduced with permission (Kim *et al.* 1993*a*).

holoenzymes containing many if not all of the general initiation factors except for TFIID (Koleske & Young 1995).

Reconstitution of the pol II preinitiation complex *in vitro* has proved remarkably successful for mechanistic studies of basal transcription initiation. However, pol II-mediated transcription is considerably more complex *in vivo*. A large number of other transcription factors, both cellular and viral in origin, regulate the precise level of messenger RNA production from class II nuclear gene promoters (reviewed in Hori & Carey 1994). These proteins are often referred to as transcriptional activators. They modulate transcription by recognizing promoter proximal and/or distal enhancer DNA targets and participating in highly specific protein–protein interactions with components of the PIC and with each other. Efficiency of RNA production from pol II promoters depends, at least in part, on the half-life of the promoter specific transcription complex and much effort is now being devoted to establishing good *in vitro* models of activator-dependent transcription initiation.

Studies of the mechanisms of action of TBP in nuclear gene transcription by RNA polymerases I (pol I) and III (pol III) are also well advanced (Reeder 1992; Hernandez 1993). A defined TBP-TAF complex, known as SL1 (selectivity factor 1), has been implicated in pol I transcription, and its three TAFs are believed to be distinct from the pol II TAFs found in TFIID (Comai *et al.* 1992). TFIIIB is the pol III-specific TBP-TAF complex consisting of at least two TAFs, one of which is similar to TFIIB (Wang & Roeder 1995).

My laboratory has been studying some of the mechanistic aspects of eukaryotic transcription initiation using X-ray crystallography and other biophysical methods. Our work has yielded structures of TBP and its complex with the TATA element of the Adenovirus major late promoter. Most recently, we described the structure of a ternary complex of TFIIB recognizing the preformed TBP-DNA complex at 2.7 Å resolution.

2. APO-TBP: A QUASI-SYMMETRIC MOLECULAR SADDLE

In 1992, we reported the structure of TBP isoform 2 (TBP2) from *Arabidopsis thaliana* at 2.6 Å resolution (Nikolov *et al.* 1992). Further progress on crystallographic studies of uncomplexed TBPs includes additional refinement of TBP2 at 2.1 Å resolution (Nikolov & Burley 1994), and a molecular replacement structure of the C-terminal 180 residues of yeast TBP (Chasman *et al.* 1993). The structure of TBP2 determined at 2.1 Å resolution is illustrated in figure 2.

Both apo-TBP structures are very similar, with two α/β-structural domains of 89–90 amino acids related by approximate intramolecular twofold symmetry. TBP2 has a relatively flexible 18 amino acid N-terminal segment. The C-terminal or core region of TBP binds to the TATA consensus sequence (TATAa/

Downloaded from rstb.royalsocietypublishing.org

tAa/t) with high affinity and slow off rate, recognizing minor groove determinants and promoting DNA bending. The N-terminal portion of TBP varies in length, shows little or no conservation among different organisms and is largely unnecessary for transcription in certain yeast strains.

TBP resembles a molecular saddle with approximate maximal dimensions $32 \times 45 \times 60$ Å. DNA binding is supported by the concave underside of the saddle, which is lined by the central eight strands of the ten-stranded anti-parallel β-sheet. The convex upper surface of TBP2 is composed of the four α-helices, the basic peptide linking the two domains, parts of strands S1 and S1′, and the non-conserved 18 N-terminal residues. This extensive upper surface binds various components of the transcription machinery (reviewed in Nikolov & Burley 1994). Each domain or structural repeat comprises approximately half of the phylogenetically conserved C-terminus of TBP, consisting of a five-stranded, curved antiparallel β-sheet and two α-helices. The two helices, lying approximately perpendicular to each other, abut the convex side of the sheet forming the hydrophobic core of each domain. The two structural domains of TBP2 are topologically identical with root mean square (r.m.s.) deviation between equivalent α-carbon atomic positions = 1.1 Å, corresponding to the two imperfect repeats in amino acid sequence (30% identical at the amino acid level and 50% identical at the nucleotide level). TBP's ancestor may, therefore, have functioned as a dimer, with gene duplication and fusion giving rise to a monomeric, quasi-symmetric TBP (Nikolov & Burley 1994).

The two crystal forms of apo-TBP each have two copies of TBP in the asymmetric unit. For TBP2 this appears to result from weak molecular self-association (buried surface area = 1700 Å^2 and measured K_d = 1 μM; D. B. Nikolov & S. K. Burley, unpublished data), which can be disrupted by dilution or addition of duplex oligonucleotides bearing a TATA element (Nikolov *et al.* 1992). There is also a report of human TBP and TFIID forming dimers at physiologic intranuclear concentration (Coleman *et al.* 1995).

3. TBP-DNA: MINOR GROOVE RECOGNITION/CORE PROMOTER BENDING

Structures of TBP2 complexed with the Adenovirus major late promoter (AdMLP) TATA element (TATAAAAG) (Kim *et al.* 1993*a*; Kim & Burley 1994), and the C-terminus of yeast TBP complexed with the yeast CYC1 -52 TATA element (TATATAAA) (Kim *et al.* 1993*b*) have been reported at 1.9 Å and 2.5 Å, respectively (figure 3).

Although the two co-crystal structures differ slightly in detail, both demonstrate an induced-fit mechanism of protein–DNA recognition. DNA-binding is mediated by the protein's curved, eight-stranded, antiparallel β-sheet, which provides a large concave surface for minor groove and phosphate-ribose contacts with the 8 b.p. TATA element. The 5′ end of standard

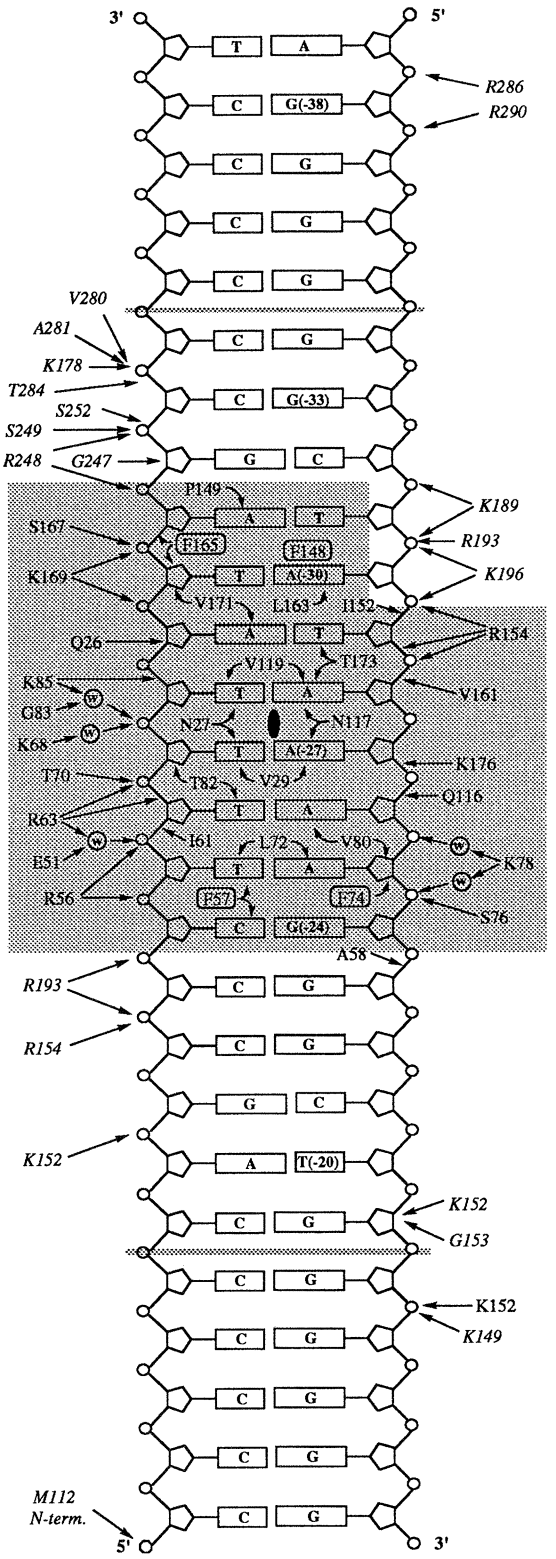


Figure 4. Schematic view of cTFIIB-TBP2-DNA interactions. TBP2-DNA interactions are restricted to the shaded area and shown with regular type (Kim & Burley 1994). cTFIIB-DNA interactions are depicted in italics. Reproduced with permission (Nikolov *et al.* 1995).

B-form DNA enters the underside of the molecular saddle, where the C-terminal portion of TBP produces an abrupt transition to an unprecedented, partly unwound form of the right-handed double helix is induced by insertion of two phenylalanine residues into the first T:A base step. Thereafter, the widened minor

Downloaded from rstb.royalsocietypublishing.org

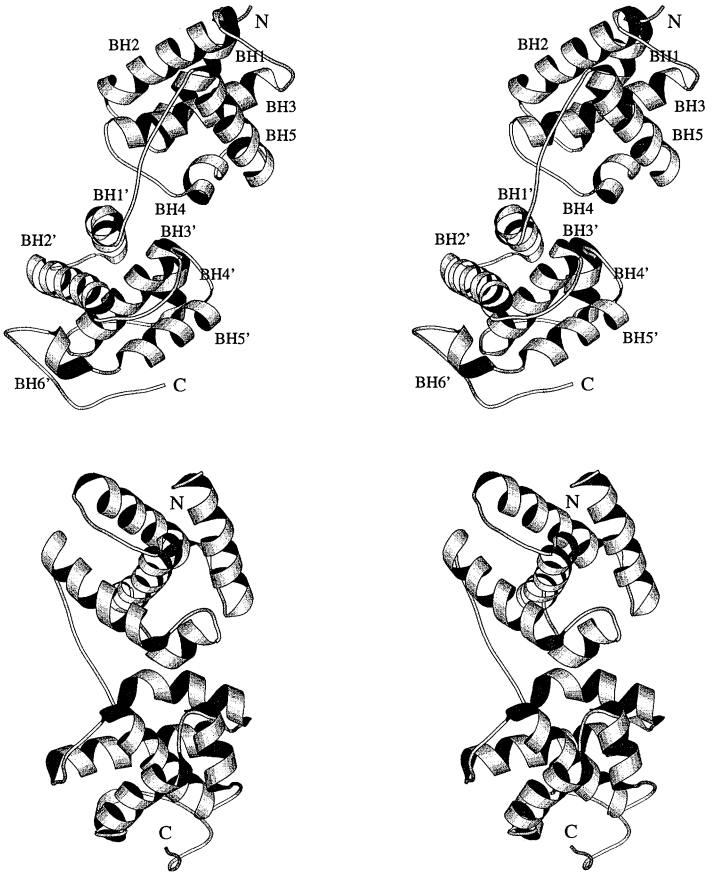


Figure 5. Structure of human cTFIIB. The N- and C-termini of the protein are labelled, and the α-helices of each domain are numbered (BH). The symbol (′) refers to the second structural domain. The BH6′ helix is absent from the first domain. Upper panel: stereodrawing of a ribbon representation of the three-dimensional structure viewed down the cleft between the two quasi-identical domains. Lower panel: Stereodrawing viewed perpendicular to the cleft. Reproduced with permission (Nikolov *et al.* 1995).

groove face of the unwound, smoothly bent DNA is approximated to the underside of the molecular saddle, burying a total surface area of about 3100 Å^2, permitting direct interactions between protein side chains and the minor groove edges of the central 6 b.p. A second large kink is induced by insertion of two phenylalanine residues in the base step between the last 2 b.p. of the TATA element, and there is a corresponding abrupt return to B-form DNA. Despite this massive distortion, Watson-Crick base pairing is preserved throughout and there appears to be no strain induced in the DNA, because partial unwinding has been compensated for by right-handed supercoiling of the double helix. Side chain-base contacts are restricted to the minor groove, including the four phenylalanines described above plus five hydrogen bonds and a large number of van der Waals contacts (figure 4). There are no water molecules mediating side chain–base interactions and the majority of the hydrogen bond donors and acceptors on the minor groove edges of the bases remain unsatisfied (13/17 in the AdMLP TATA box). Detailed analysis of the TBP2-DNA cocrystal structure at 1.9 Å resolution demonstrates that the protein also undergoes a modest conformational change on DNA binding, involving a twisting motion of one domain with respect to the other (Kim & Burley 1994).

Other biophysical methods have been used to study interactions between TBP and DNA. Site selection experiments with *Acanthamoeba* TBP showed a marked preference for a site very similar to those studied crystallographically (TATATAAG) (Wong & Bateman 1994). DNA bending by TBP in solution was confirmed using circular permutation assays (Starr *et al.* 1995). TBP binding was also shown to be enhanced by prebending of DNA towards the major groove (Parvin *et al.* 1995). TBP-DNA association kinetics have been studied by various techniques (Hoopes *et al.* 1992; Coleman & Pugh 1995; Parkhurst *et al.* 1996; Perez-Howard *et al.* 1995), and three of the four studies gave results consistent with simultaneous binding and bending with a single second-order rate constant of about 10^5 M^{-1}s^{-1}. Coleman & Pugh (1995) opted for a dramatically different model, involving dissociation of a tight human TBP dimer, tight nonspecific DNA binding by TBP and sliding of TBP on DNA. In addition, a novel chemical probe was used to demonstrate that core promoter distortion transiently extends beyond the confines of the TATA box during TBP binding (Sun & Hurley 1995).

4. TFIIB-TBP-DNA: RECOGNITION OF THE TBP-DNA COMPLEX

The crystal structure of a TFIIB-TBP-TATA element ternary complex has been determined at 2.7 Å resolution (Nikolov *et al.* 1995). Core TFIIB (cTFIIB)

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES
PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Downloaded from rstb.royalsocietypublishing.org

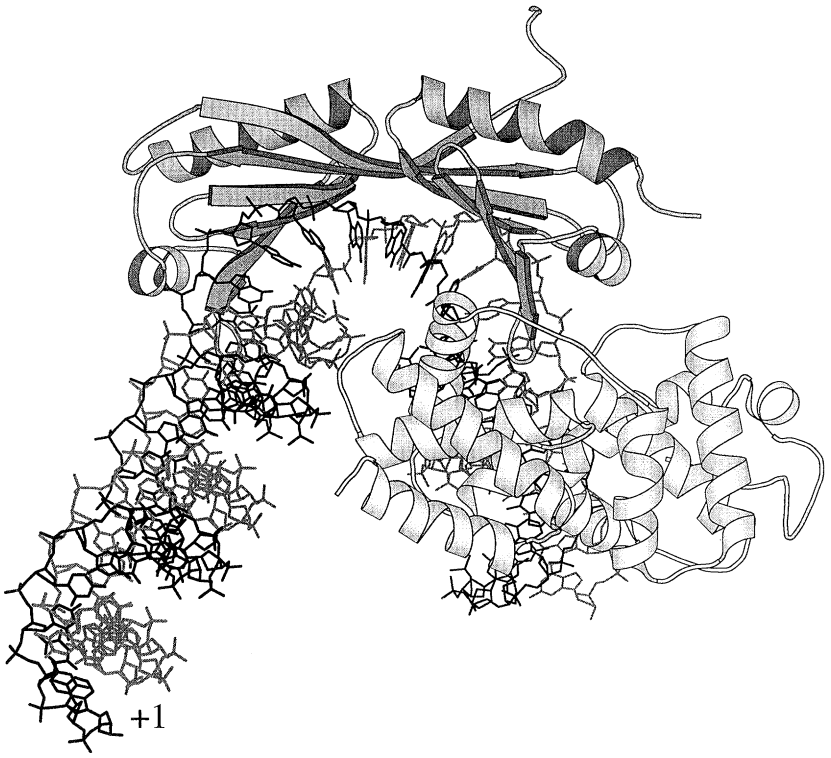


Figure 6. cTFIIB and TBP interacting with the AdMLP. Three-dimensional structure of the ternary complex of human cTFIIB recognizing TBP2 from *Arabidopsis thaliana* complexed with the AdMLP TATA element. cTFIIB (light grey) and TBP (dark grey) are depicted as shaded ribbons, and the DNA is shown as a stick figure with the transcription start site labelled with +1. The coding strand is denoted with solid bonds. The view is identical to that shown in figure 3. When cTFIIB recognizes the TBP-DNA complex there is essentially no change in trajectory of the negatively charged phospho-ribose backbone. Reproduced with permission (Nikolov *et al.* 1995).

is a two domain α-helical protein that resembles cyclin A (Jeffrey *et al.* 1995) (figure 5).

The ternary complex is formed by cTFIIB clamping the acidic C-terminal stirrup of TBP2 (S2′–S3′) in its cleft, interacting with H1′, the C-terminus, and the phospho-ribose backbone up- and downstream of the center of the TATA element (figures 4 and 6). Although the two domains of cTFIIB have the same fold, they do not have chemically identical surfaces and cannot make equivalent interactions with TBP2. Contacts between cTFIIB and the C-terminal stirrup of TBP2 are made by BH3, BH4 and BH5. The interdomain peptide interacts with H1′ and the C-terminal stirrup of TBP2. cTFIIB's BH2′-BH3′ loop interacts with the same stirrup and the C-terminus of TBP2. Despite the very extensive intermolecular contacts visualized in the ternary complex structure (total buried surface area ∼ 5600 Å^2), the structure of the TBP2-TATA element complex itself is essentially unchanged. cTFIIB recognizes the preformed TBP-DNA complex, including the path of the phosphoribose backbone of created by the unprecedented DNA deformation induced by binding of TBP. In addition to stabilizing the TBP-DNA complex, TFIIB binding contributes to the polarity of TATA element recognition. If TBP were to bind to the quasi-symmetric TATA box in the wrong orientation (i.e. the N-terminal half of the molecular saddle interacts with the 5′ end of the TATA element), the basic/hydrophobic surface of the N-terminal stirrup (S2–S3) would make unfavorable electrostatic interactions with the basic cleft of TFIIB. It is remarkable that the NMR structure of cTFIIB displays a slightly different arrangement of the two domains (Bagby *et al.* 1995), suggesting that cTFIIB undergoes a modest conformational change on recognizing the TBP-DNA complex. The first domain of cTFIIB forms the downstream surface of the cTFIIB-TBP-DNA ternary complex, where together with the putative ZN^{2+}-binding, N-terminal domain of full length TFIIB it could readily act as a bridge between TBP and pol II fixing the transcription start site. The remaining solvent-accessible surfaces of TBP (∼ 7900 Å^2) and the TFIIB (∼ 8300 Å^2) are very extensive, providing for an ample number of recognition sites for binding of TAFs, other class II initiation factors, and transcriptional activators and coactivators.

5. CONCLUSIONS AND PERSPECTIVES

X-ray crystallographic studies of apo-TBP and its complexes with two TATA elements have revealed a new quasi-symmetric protein fold, an unprecedented protein-induced DNA deformation, and a novel induced-fit mechanism of DNA recognition via contacts with the minor groove. Most recently, this work set the stage for the first structure determination of a protein recognizing a preformed protein-DNA complex, the TFIIB-TBP-TATA element ternary complex. These cocrystal structures provided direct views of two critical steps early in the assembly of the preinitiation complex, required for correct initiation of

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES
PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Downloaded from rstb.royalsocietypublishing.org

transcription by pol II. They have contributed significantly to our understanding of the precise biochemical mechanisms responsible for controlling messenger RNA production in eukaryotes, and continue to serve as a structural foundation from which to plan and interpret studies of class II nuclear gene expression.

I am grateful to J. L. Kim and D. B. Nikolov for providing illustrations for this review.

REFERENCES

Bagby, S., Kim, S., Maldonado, E., Tong, K., Reinberg, D. & Ikura, M. 1995 Solution structure of the C-terminal core domain of human TFIIB: similarity to Cyclin A and interaction with TATA-binding protein. *Cell* **82**, 857–867.

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

Chasman, D., Flaherty, K., Sharp, P. & Kornberg, R. 1993 Crystal structure of Yeast TATA-binding protein and a model for interaction with DNA. *Proc. natn. Acad. Sci. U.S.A.* **90**, 8174–8178.

Coleman, R. & Pugh, B. 1995 Evidence for functional binding and stable sliding of the TATA binding protein on nonspecific DNA. *J. biol. Chem.* **270**, 13850–13859.

Coleman, R., Taggart, A., Benjamin, L. & Pugh, B. 1995 Dimerization of TATA binding protein. *J. biol. Chem.* **270**, 13842–13849.

Comai, L., Tanese, N. & Tjian, R. 1992 The TATA-binding protein and associated factors are integral components of RNA polymerase I transcription factor, SL1. *Cell* **68**, 965–976.

Gabrielson, O. & Sentenac, A. 1991 RNA polymerase III (C) and its transcription factors. *Trends Biochem. Sci.* **16**, 412–416.

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

Hoopes, B., LeBlanc, J. & Hawley, D. 1992 Kinetic analysis of yeast TFIID-TATA box complex formation suggests a multi-step pathway. *J. biol. Chem.* **267**, 11539–11546.

Hori, R. & Carey, M. 1994 The role of activators in assembly of RNA polymerase II transcription complexes. *Curr. Opin. Gen. Dev.* **4**, 236–244.

Jeffrey, P., Russo, A., Polyak, K., Gibbs, E., Hurwitz, J., Massague, J. & Pavletich, N. 1995 Mechanism of CDK activation revealed by the structure of a cyclinA-CDK2 complex. *Nature, Lond.* **376**, 313–320.

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

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

Kim, Y., Geiger, J. H., Hahn, S. & Sigler, P. B. 1993*b* Crystal Structure of a yeast TBP/TATA-box complex. *Nature, Lond.* **356**, 512–520.

Koleske, A. & Young, R. 1995 The RNA polymerase II holoenzyme and its implications for gene regulation. *Trends Biochem. Sci.* **20**, 113–116.

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

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

Matsui, T., Segall, J., Weil, P. & Roeder, R. G. 1980 Multiple factors required for accurate initiation of transcription by purified RNA polymerase II. *J. biol. Chem.* **255**, 11992–11996.

Nikolov, D. & Burley, S. K. 1994 2.1 Å Resolution refined structure of a TATA box-binding protein TBP *Nature Struct. Biol.* **1**, 621–637.

Nikolov, D., Chen, H., Halay, E., Usheva, A., Hisatake, K., Lee, D., Roeder, R. & Burley, S. K. 1995 Crystal structure of a TFIIB-TBP-TATA element ternary complex. *Nature, Lond.* **377**, 119–128.

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

Parkhurst, K., Brenowitz, M. & Parkhurst, L. 1995 Simultaneous binding and bending of promoter DNA by TBP: Real time kinetic measurements. *Biochemistry*, **35**. (In the press.)

Parvin, J., McCormick, R., Sharp, P. & Fisher, D. 1995 Pre-bending of a promoter sequence enhances affinity for the TATA-binding factor. *Nature, Lond.* **273**, 724–727.

Parvin, J. & Sharp, P. 1993 DNA topology and a minimal set of basal factors for transcription by RNA polymerase II. *Cell* **73**, 533–540.

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

Reeder, R. 1992 Regulation of transcription by RNA polymerase I. In *Transcription regulation* (ed. S. McKnight & K. R. Yamamoto), pp. 315–348. New York: Cold Spring Harbor Laboratory Press.

Roeder, R. G. 1991 The complexities of eukaryotic transcription initiation: regulation of preinitiation complex assembly. *Trends Biochem. Sci.* **16**, 402–408.

Sentenac, A. 1985 Eukaryotic RNA polymerases. *CRC Crit. Rev. Biochem.* **18**, 31–90.

Starr, D., Hoopes, B. & Hawley, D. 1995 DNA bending is an important component of site-specific recognition by the TATA binding protein. *J. Molec. Biol.* **250**, 434–446.

Sun, D. & Hurley, L. 1995 TBP unwinding of the TATA box induces a specific downstream unwinding site that is targetted by pluramycin. *Chem. Biol.* **2**, 457–469.

Wang, Z. & Roeder, R. G. 1995 Structure and function of a human transcription factor TFIIIB subunit that is evolutionarily conserved and contains both TFIIB- and high-mobility-group protein 2 domains. *Proc. natn. Acad. Sci. U.S.A.* **92**, 7026–7030.

Wong, J. & Bateman, E. 1994 TBP-DNA interactions in the minor groove discriminate between A:T and T:A base pairs. *Nucl. Acids Res.* **22**, 1890–1896.

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

Zawel, L., Lu, H., Cisek, L., Corden, J. & Reinberg, D. 1993 The cycling of RNA polymerase II during transcription. *Cold Spring Harbor Symp. Quant. Biol.* **58**, 187–198.

Zawel, L. & Reinberg, D. 1993 Initiation of transcription by RNA polymerase II: a multi-step process. *Prog. Nucl. Acids Res. Molec. Biol.* **44**, 67–108.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Downloaded from rstb.royalsocietypublishing.org

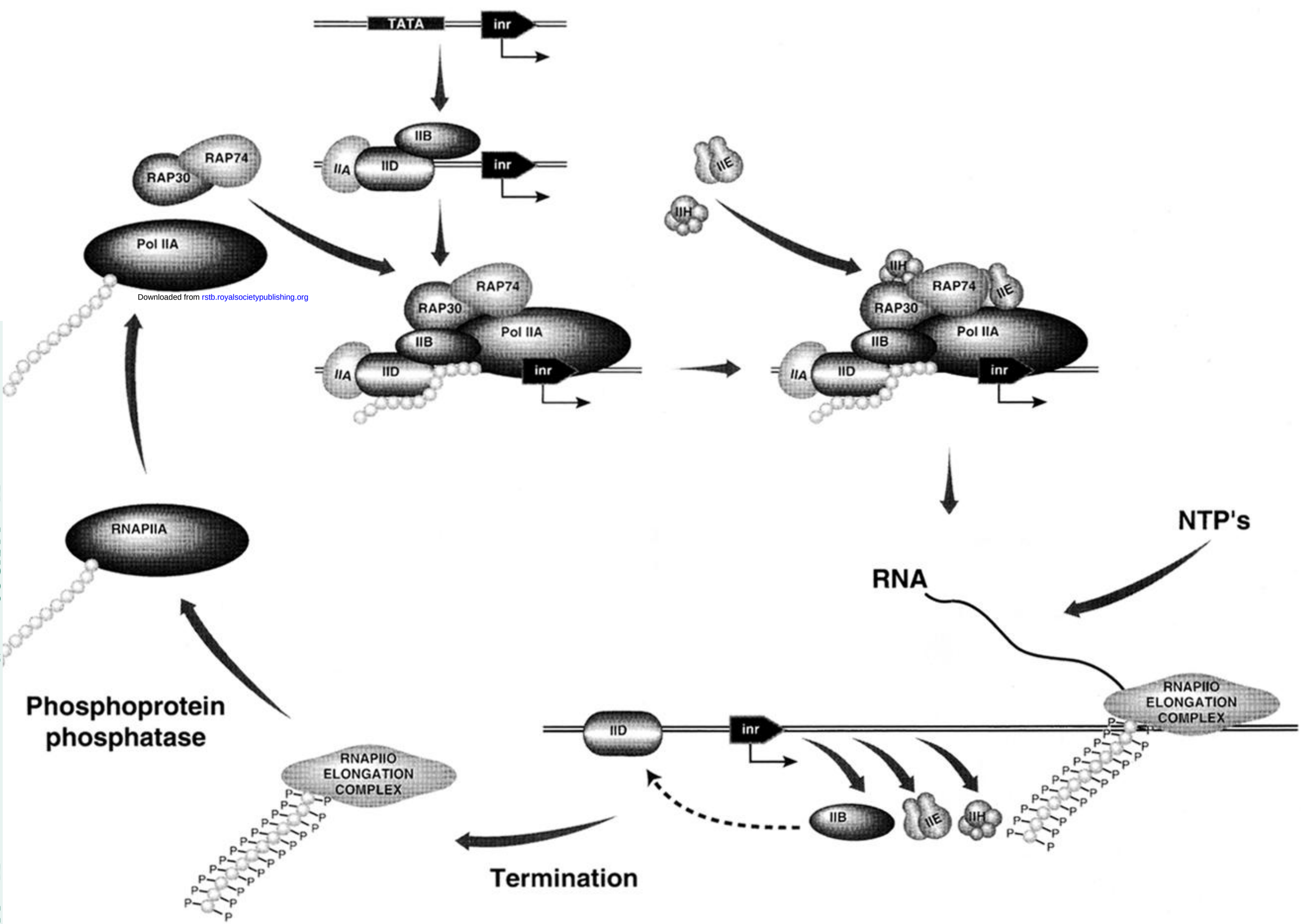
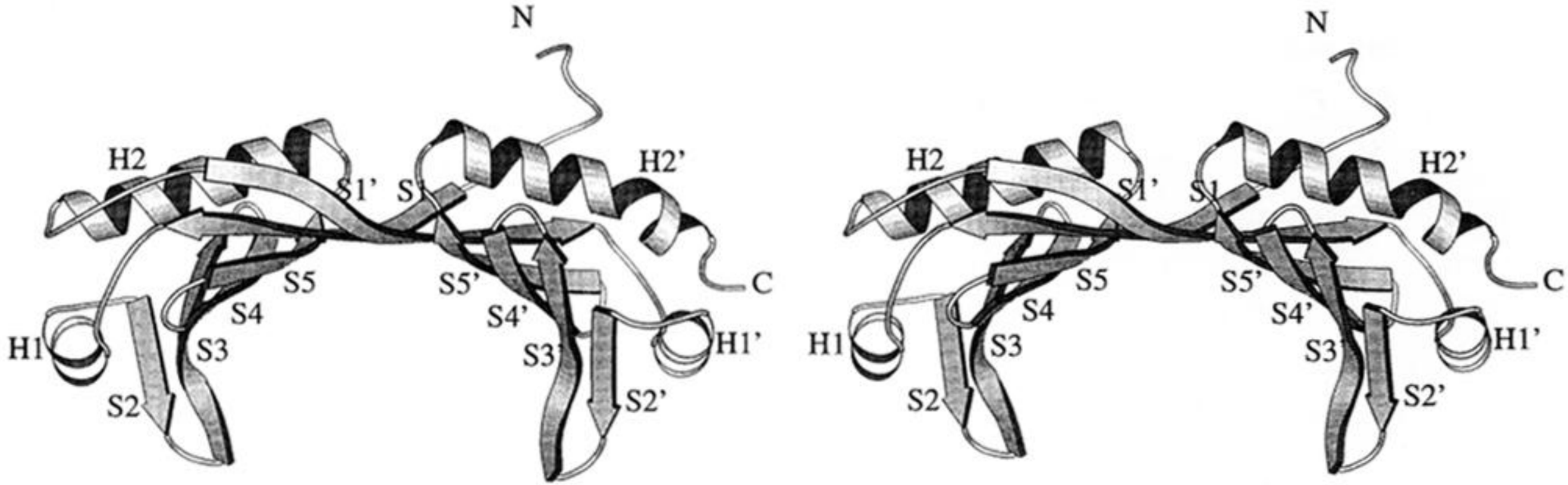


Figure 1. Preinitiation complex assembly begins with TFIID recognizing the TATA element, followed by coordinated accretion of TFIIA, TFIIB, the nonphosphorylated form of pol II and TFIIF (RAP30/RAP74), TFIIE and TFIIH. Before elongation pol II is phosphorylated by TFIIH. After termination, a phosphatase recycles pol II to its non-phosphorylated form, allowing the enzyme to reinitiate transcription. TBP (and TFIID) binding to the TATA box is an intrinsically slow step, yielding a long-lived protein-DNA complex. Efficient reinitiation of transcription can be achieved if recycled pol II re-enters the preinitiation complex before TFIID dissociates from the core promoter. Reproduced with permission (Zawel *et al.* 1993).

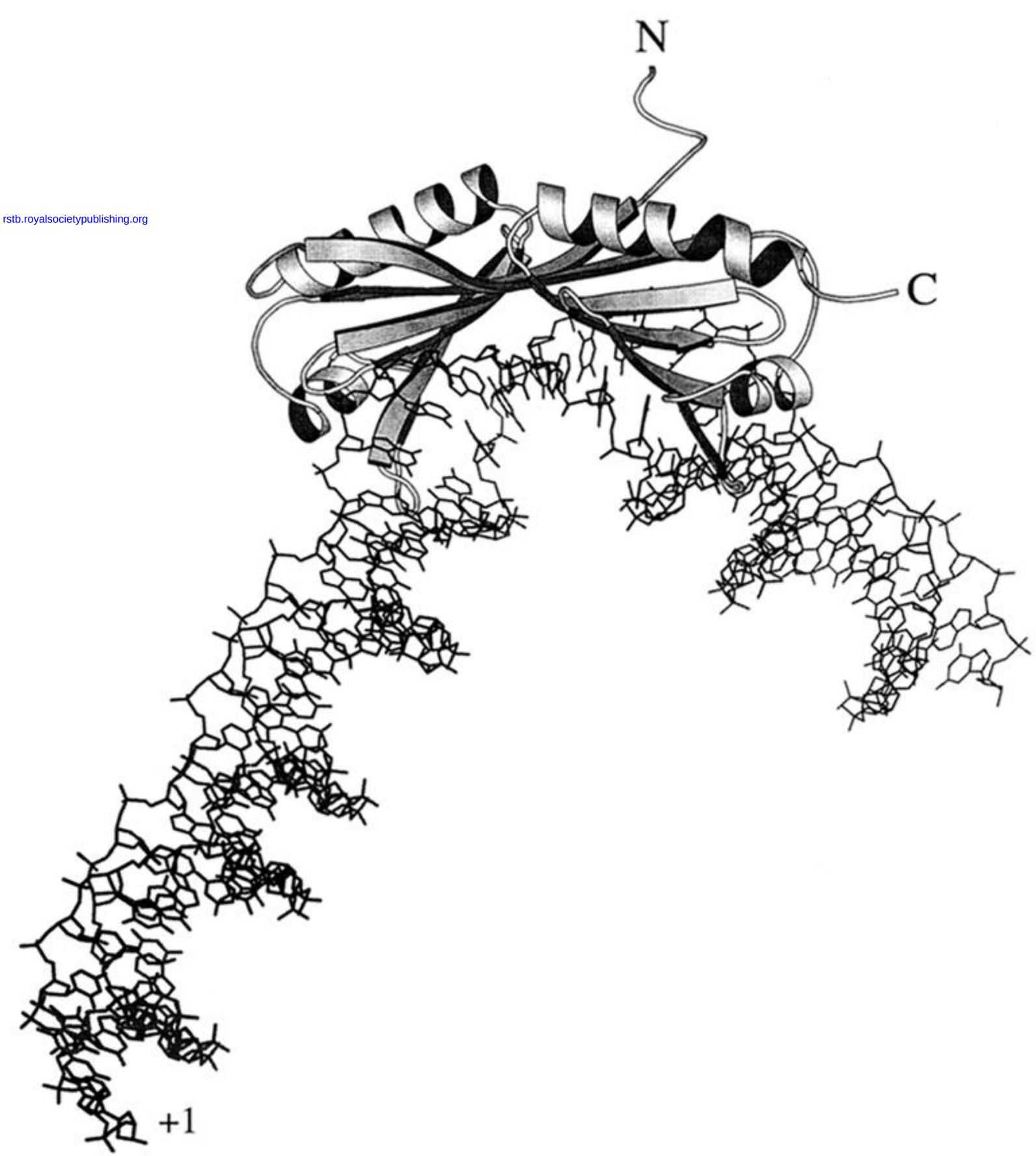
PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES



gure 2. MOLSCRIPT (Kraulis 1991) stereodrawing of the structure of TBP2 viewed perpendicular to the internal eudodyad axis. The N and C termini of the protein are indicated. The α-helices are shown as ribbon spirals and belled H, the β-strands are shown as ribbon arrows (S), and loops and turns are drawn as double lines. The symbol efers to the second structural domain or repeat. Reproduced with permission (Nikolov *et al.* 1992).

Downloaded from rstb.royalsocietypublishing.org



gure 3. Three-dimensional structure of TBP2 from *Arabidopsis thaliana* complexed with the AdMLP TATA element. ne molecular saddle (N- and C-termini labelled) is depicted with a ribbon drawing and the DNA is shown as a stick ure with the transcription start site labelled with +1. When TBP recognizes the minor groove of the TATA ement, the DNA is kinked and unwound to present the minor groove edges of the bases to the underside of the olecular saddle. The coding strand is denoted with solid bonds. Reproduced with permission (Kim *et al.* 1993*a*).

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES