Structure of the central core domain of TFIIE β with a novel double-stranded DNA-binding surface

Masahiko Okuda¹, Yoshinori Watanabe²,
Hideyasu Okamura¹, Fumio Hanaoka²,
Yoshiaki Ohkuma² and
Yoshifumi Nishimura^{1,3,4}

¹Graduate School of Integrated Science, Yokohama City University, 22-2 Seto, Kanazawa-ku, Yokohama 236-0027, ²Institute for Molecular and Cellular Biology, Osaka University, 1-3 Yamada-oka, Suita, Osaka 565-0871 and ³Genomic Sciences Center (GSC), RIKEN (The Institute of Physical and Chemical Research), 2-1 Hirosawa, Wako-shi, Saitama 351-0198, Japan

⁴Corresponding author
e-mail: nishimura@yokohama-cu.ac.jp

Human general transcription factor TFIIE consists of two subunits, TFIIE α and TFIIE β . Recently, TFIIE β has been found to bind to the region where the promoter starts to open to be single-stranded upon transcription initiation by RNA polymerase II. Here, the central core domain of human TFIIE β (TFIIE β c) has been identified by a limited proteolysis. This solution structure has been determined by NMR. It consists of three helices with a β hairpin at the C-terminus, resembling the winged helix proteins. However, TFIIE β c shows a novel double-stranded DNA-binding activity where the DNA-binding surface locates on the opposite side to the previously reported winged helix motif by forming a positively charged furrow. A model will be proposed that TFIIE stabilizes the preinitiation complex by binding not only to the general transcription factors together with RNA polymerase II but also to the promoter DNA, where double-stranded DNA starts to open to be single-stranded upon activation of the preinitiation complex.

Keywords: DNA-binding/general transcription factor/
NMR/TFIIE/winged helix

Introduction

Transcription initiation from eukaryotic protein-coding genes is a complex process requiring RNA polymerase II (Pol II) and six general transcription factors (TFIIA, TFIIB, TFIID, TFIIE, TFIIIF and TFIIH), which are necessary to assemble the preinitiation complex (PIC) at the promoter and initiate transcription (reviewed in Orphanides *et al.*, 1996; Roeder, 1996). Analysis of the PIC assembly pathway using the isolated general transcription factors has indicated that the process is initiated by the binding of the TBP subunit of TFIID to the TATA box. Then the stepwise assembly of TFIIA, TFIIB, a complex of Pol II and TFIIIF, TFIIE and TFIIH proceeds. TFIIH, which is recruited through a direct interaction with TFIIE, activates the PIC through its enzymatic activities, resulting in the open complex forma-

tion (Lu *et al.*, 1991; Ohkuma and Roeder, 1994; Dvir *et al.*, 1996). In addition to their key roles in the PIC formation and initiation, TFIIE and TFIIH play important roles in the transition from initiation to elongation (Goodrich and Tjian, 1994; Dvir *et al.*, 1997; Kumar *et al.*, 1998).

TFIIE is a heterotetramer containing two of each two subunits, TFIIE α and TFIIE β , with molecular masses of 57 and 34 kDa, respectively (Ohkuma *et al.*, 1990, 1991; Peterson *et al.*, 1991; Sumimoto *et al.*, 1991). TFIIE plays essential roles in TFIIH recruitment (Flores *et al.*, 1992) and the regulation of TFIIH activities: a kinase activity that phosphorylates the C-terminal domain (CTD) of the largest subunit of Pol II, a DNA-dependent ATPase activity and a DNA helicase activity (Lu *et al.*, 1991; Drapkin *et al.*, 1994; Ohkuma and Roeder, 1994). In addition, TFIIE plays important roles in the open complex formation (Timmers, 1994; Holstege *et al.*, 1995) and in the clearance of the initiation complex from the promoter (Goodrich and Tjian, 1994; Dvir *et al.*, 1997). In the complex structure of yeast TFIIE with Pol II revealed by two-dimensional crystallography, TFIIE interacts near the active center of Pol II (Leuther *et al.*, 1996). A photocrosslinking study showed that TFIIE β binds to the core promoter region of -14 and -2 upstream from the transcription initiation site (+1), while TFIIE α does not bind to DNA by itself (Robert *et al.*, 1996). TFIIE α binds strongly to TBP and TFIIH, and weakly to Pol II and TFIIIF, while TFIIE β binds strongly to Pol II, TFIIB, Rap30 (TFIIIF β) and TFIIH, suggesting a plausible position of TFIIE in the PIC (reviewed in Ohkuma, 1997). However, a recent study has shown that TFIIE α could directly stimulate the TBP binding to the promoter in the absence of other general transcription factors and TFIIE β could directly interact with TFIIA, suggesting new roles for TFIIE in the early stages of PIC formation (Yokomori *et al.*, 1998). To understand the detailed molecular mechanism of PIC formation, several three-dimensional structures of the general transcription factors have been determined (reviewed in Nikolov and Burley, 1997), for example, the core domain structures of TBP (Nikolov *et al.*, 1992), TFIIB (Bagby *et al.*, 1995; Zhu *et al.*, 1996), TBP associated factors (TAFs) (Xie *et al.*, 1996; Birck *et al.*, 1998), TBP and TAF (Liu *et al.*, 1998), and Rap30 (TFIIIF β) (Groft *et al.*, 1998) in their DNA-free states, in addition, the complex structures with DNA of the core domains of TBP (J.L.Kim *et al.*, 1993; Y.Kim *et al.*, 1993), TFIIB and TBP (Nikolov *et al.*, 1995), and TFIIA and TBP (Geiger *et al.*, 1996; Tan *et al.*, 1996).

Human TFIIE β , consisting of 291 amino acids, is a basic protein with a pI value of 9.5 and contains several putative structural motifs as shown in Figure 1 (Sumimoto *et al.*, 1991). By using a systematic series of deletion mutants of TFIIE β , the C-terminal two basic regions,

A

B

Fig. 1. Sequence alignment of TFIIIE β . (A) Schematic diagram of putative structural motifs and characteristic sequences of human TFIIIE β . Ser-rich, a serine-rich sequence (residues 26–71); RAP30, a region similar to the Pol II binding region of Rap30 (TFIIF β) (residues 79–111); LR, a leucine repeat motif (residues 145–163); σ 3, a region similar to the bacterial σ factor subdomain 3 (residues 163–193); BR-HLH, a basic region helix–loop–helix motif (residues 197–238); BR-HL, a basic region helix–loop sequence (residues 258–291). (B) Sequence alignment of human TFIIIE β with other TFIIIE β . Residues identified in all five species are boxed in red, residues identified in four out of the five species are boxed in pink, and conserved residues are boxed in yellow. The asterisks indicate the residues that form the hydrophobic core of human TFIIIE β c. The secondary structure of human TFIIIE β c is indicated above the sequence. Three α helical regions are indicated by H1, H2 and H3, and three β -stranded regions are indicated by S1, S2 and S3. ‘ht’ means a helical turn region.

basic helix–loop–helix and basic helix–loop regions, are found to be very important for the interactions with TFIIB, Rap30 (TFIIF β), single-stranded DNA, *Drosophila* transcriptional regulatory factor Krüppel, and TFIIE α (Okamoto *et al.*, 1998). The roles of the other regions of TFIIE β are not yet understood, although the central 227 amino acid region (residues 51–277) of TFIIE β has been demonstrated to be necessary and sufficient for both basal and activated transcription (Okamoto *et al.*, 1998). Here, we have identified a structural core domain of TFIIE β (TFIIE β c) using proteolytic digestion, and the structure of TFIIE β c has been determined using NMR. Importantly, this domain was found to have a dominant negative effect on basal transcription and, therefore, has been assumed to be functionally very important (Okamoto *et al.*, 1998). Through the three-dimensional structure analysis, it is suggested that TFIIE β c plays a role in double-stranded DNA binding of TFIIE, whose function might be important when the PIC is activated and the promoter starts opening to become single-stranded.

Results and discussion

Structure determination

The TFIIE β cDNAs have been cloned from human (Sumimoto *et al.*, 1991), *Xenopus* (Ohkuma *et al.*, 1992), *Drosophila* (Wang *et al.*, 1997), *Caenorhabditis elegans*

(S.Yamamoto, Y.Watanabe, P.J.van der Spek, T.Watanabe, H.Fujimoto, F.Hanaoka and Y.Ohkuma, in preparation) and *Saccharomyces cerevisiae* (Feaver *et al.*, 1994), showing the amino acid sequence homology among these TFIIE β proteins. To identify the structural domain of TFIIE β , the human recombinant TFIIE β was subjected to a limited proteolysis. One structural domain, consisting of the amino acid residues 66–146, and two structural domains consisting of the amino acid residues 74–152 and 211–244, have been identified by trypsin and chymotrypsin digestion, respectively. Here, the central core domain (designated as TFIIE β c) consisting of the amino acid residues 66–146, including a region putatively similar to the Pol II binding region of Rap30 (TFIIF β), has been investigated (Figure 1). This human TFIIE β c was expressed in *Escherichia coli* BL21(DE3) pLysS. Purified TFIIE β c was subjected to conventional multi-dimensional NMR measurements. Almost all signals of the main chain atoms, amide proton, amide nitrogen, α proton and α carbon for each amino acid of TFIIE β c could be assigned except for amide nitrogen signals of the three proline residues, Asn126 and the N-terminal two residues, and amide proton signals of the three proline residues, Asn126 and the N-terminal Ala66. The three-dimensional structure of TFIIE β c was determined based on a total of 1184 distance restraints, as well as 38 dihedral angle restraints

Fig. 2. Stereo view of the solution structure of human TFIIIEβc. (A) Best-fit superposition of the ensemble of the final 20 NMR structures of human TFIIIEβc. (B) The averaged structure of human TFIIIEβc. The averaged structure over the final 20 NMR structures of human TFIIIEβc is shown in a ribbon representation. These figures were made using the program MOLMOL (Koradi *et al.*, 1996).

and 20 hydrogen bond restraints, using a four-dimensional simulated annealing protocol (Nakai *et al.*, 1993).

Overview of the structure

Figure 2A shows the 20 calculated structures of TFIIIEβc and Figure 2B shows an averaged structure over 20 calculated structures. Structural statistics are summarized in Table I. With the exception of the N-terminal nine residues (66–74) and the C-terminal five residues (142–146), the backbone of TFIIIEβc is well defined. The structural quality of the ensemble of the calculated structures was checked using the program PROCHECK-NMR (Laskowski *et al.*, 1996). Almost all residues except glycine fall into allowed regions: 83.2, 15.0 and 1.0%

residues without glycine residues in the most favorable, additionally allowed and generously allowed regions, respectively. The average atomic root-mean-square (r.m.s.) deviations of the 20 simulated annealing (SA) structures about the mean coordinate positions are $0.41 (\pm 0.04)$ Å for the backbone atoms and $0.88 (\pm 0.05)$ Å for all heavy atoms.

The structure of TFIIIEβc contains three helices and a C-terminal β hairpin. The three helices, H1, H2 and H3 consist of Phe75–Gln90, Leu99–Glu105 and Leu113–Met120, respectively. A TXXE type capping box of α helix is found in the N-terminal region of H2, by forming hydrogen bonds between Thr98 and Glu101 (Harper and Rose, 1993). The region of Thr121–Val125

has a helical turn structure (ht). The C-terminal region has a β hairpin formed by a β strand with Ile130–Ile133 (S2 strand), a hairpin loop with Asp134 and Gly135, and a β strand with Lys136–Phe139 (S3 strand). Both S2 and S3 strands form a twisted anti-parallel β sheet with Leu 97 (S1 strand). H1 is longer than H2 and H3. The three helices are maintained by a hydrophobic core, formed by Val77, Leu78 and Ile81 from H1, Leu99, Ile102 and

Leu103 from H2, Trp118 and Leu119 from H3, and Ala123 and Leu124 from the helical turn as shown in Figure 3. In addition, the C-terminal region of H1, consisting of Val82, Tyr84 and Met85, interacts with Leu97 from S1, Ile130 from S2, and Tyr137 and Phe139 from S3, extending the hydrophobic core. The amino acids that participate in the hydrophobic core are well conserved in the central core domains of different species as shown in Figure 1B, indicating that other central core domains of TFIIIE β from different species also hold a three-dimensional structure similar to the human TFIIIE β c structure.

Table 1. Structural statistics

Total number of distance restraints	1224
Sequential ($i - j = 1$)	333
Short range ($1 < i - j \leq 5$)	363
Long range ($i - j > 5$)	345
Intraresidue	143
Hydrogen bond	40
Number of dihedral restraints	38
Statistic for structure calculations	$<SA>^a$
R.m.s. deviations from experimental restraints ^b	
Distance ($\text{\AA}$)	0.010 ± 0.001
Dihedral ($^\circ$)	0.0015 ± 0.0056
R.m.s. deviations from idealized covalent geometry	
Bonds ($\text{\AA}$)	0.0056 ± 0.0002
Angles ($^\circ$)	0.72 ± 0.02
Impropers ($^\circ$)	0.48 ± 0.05
Coordinate precision ($\text{\AA}$) ^c (residues 75–141)	
Backbone atoms	0.41 ± 0.04
All atoms	0.88 ± 0.05
PROCHECK Ramachandran plot statistics (%) (residues 75–141)	
Residues in most favoured regions	83.2
Residues in additional allowed regions	15.0
Residues in generously allowed regions	1.0
Residues in disallowed regions	0.8

^a $<SA>$ are the family of the final 20 simulated annealing structures including residues 66–146.

^bNone of the structures exhibited distance violations $>0.3 \text{ \AA}$, dihedral angle violations $>1^\circ$. The precision of the coordinates is defined as the average atomic r.m.s. difference between the 20 individual simulated annealing structures and the mean coordinates.

Structural similarity to winged helix proteins implies the function of TFIIIE β c

To identify the structural family of TFIIIE β c, we have compared the TFIIIE β c structure with other structures so far determined using the DALI program (Holm and Sander, 1993). Based on the Z score, which is the strength of similarity between two structures, the 13 highest similar structures to the TFIIIE β c structure are identified by Z scores >3.9 . It is interesting to find that all the 13 proteins with the highest Z scores are DNA-binding proteins: the N-terminal fragment of *E.coli* topoisomerase I (Z = 5.3; Lima *et al.*, 1994); the DNA-binding domain of the cell cycle transcription factor DP2 (Z = 5.3; Zheng *et al.*, 1999); the DNA-binding domain of HNF-3 γ (Z = 5.2; Clark *et al.*, 1993); the DNA-binding domain of catabolite gene activator protein (CAP) (Z = 5.1; Schultz *et al.*, 1991); the globular domain of histone H5 (Z = 5.1; Ramakrishnan *et al.*, 1993); the restriction endonuclease *FokI* (Z = 4.9; Wah *et al.*, 1997); the DNA-binding domain of Rap30 (TFIIIF β) (Z = 4.8; Groft *et al.*, 1998); the DNA-binding domain of OmpR (Z = 4.8; Kondo *et al.*, 1997; Martinez-Hackert and Stock, 1997); the transcription regulatory protein MotA fragment (Z = 4.6; Finnin *et al.*,

Fig. 3. The hydrophobic core of TFIIIE β c. The side chains of amino acids that form the hydrophobic core of TFIIIE β c are shown. The hydrophobic side chains from the α -helices, the β -sheet and the helical turn are shown in cyan, light-green and yellow, respectively. This figure was made using the program MOLMOL (Koradi *et al.*, 1996).

1997); the DNA-binding domain of LexA repressor ($Z = 4.4$; Fogh *et al.*, 1994); the DNA-binding domain of biotin repressor, BirA ($Z = 4.1$; Wilson *et al.*, 1992); the DNA-binding domain of arginine repressor (ArgR) ($Z = 4.0$; Ni *et al.*, 1999); and the replication terminator protein (RTP) from *Bacillus subtilis* ($Z = 3.9$; Buissiere *et al.*, 1995). It is well known that the DNA-binding domains of DP2, HNF-3 γ , histone H5, Rap30 (TFIIF β), OmpR, LexA, BirA, ArgR and RTP contain the winged helix motif, consisting of three helices and a C-terminal β hairpin.

Figure 4 shows the structural comparison of TFIIE β c with the DNA-binding domains of DP2 and HNF-3 γ . It is apparent that the architecture of TFIIE β c is very close to those of the DNA-binding domains. The DNA-bound structure of the DNA-binding domain of HNF-3 γ was established using X-ray crystallography (Clark *et al.*, 1993). The third helix of the domain, the recognition helix, locates in the major groove of DNA and two C-terminal loops (wings) contact the phosphate backbone of DNA (Figure 4). The architecture of three helices and a C-terminal β hairpin except the C-terminal two loops of HNF-3 γ resembles that of TFIIE β c. The α carbon

atoms of both domains can be superimposed with a 2.3 Å r.m.s. deviation for 58 residues. Recently the DNA-complexed structure of the heterodimer of DP2 and E2F4 has been determined (Zheng *et al.*, 1999). Both DNA-binding domains of DP2 and E2F4 contain the winged helix motif, consisting of three α helices (H1, H2 and H3) and a β sheet. The structure of TFIIE β c is closer to that of DP2 ($Z = 5.3$) than that of E2F4 ($Z = 2.8$). Although the H2 and H3 helices of DP2 are much longer than the corresponding helices of TFIIE β c (Figure 4), the arrangement of three helices and a C-terminal β hairpin of DP2 resembles that of TFIIE β c. DP2 and TFIIE β c can be superimposed with an r.m.s. deviation of 2.3 Å in the C_{α} positions of 59 residues. The DNA-binding modes of DP2 and E2F4 are similar in each; the H3 helix of each domain recognizes a specific base sequence and the N-terminal portion of the H2 (E2F4) or H1 (DP2) helix and a portion of each β sheet contacts the phosphate backbone of DNA (Figure 4). In addition, TFIIE β c, although containing a putative region similar to the Pol II binding region of Rap30, has a structure very similar to the DNA-binding domain of Rap30 (Figure 4). Actually,

Fig. 4. Comparison of TFIIE β c with winged helix proteins. Comparison of TFIIE β c (A) with the DNA-binding domains of DP2 (B), HNF-3 γ (C) and Rap30 (D). All four proteins have the same topology. In the DNA-complexed structures of the DNA-binding domains of DP2 (Zheng *et al.*, 1999) and HNF-3 γ (Clark *et al.*, 1993) amino acids that contact with DNA are shown in red. Based on the DNA titration experiment of Rap30 (Groft *et al.*, 1998) amino acids whose signals change significantly are also shown in red. This figure was made using the program GRASP (Nicholls *et al.*, 1991).

TFIIIE β c did not bind directly to Pol II (Y.Watanabe and Y.Ohkuma, unpublished data). The DNA-binding ability of Rap30 has been well investigated and the DNA-binding surface of Rap30 has been structurally mapped using NMR (Groft *et al.*, 1998), indicating a DNA-binding mode similar to other winged helix proteins (Figure 4). These results strongly suggest that TFIIIE β c is a DNA-binding domain like the winged helix proteins. However, it has been found that only the C-terminal basic regions of TFIIIE β are responsible for the single-stranded (ss)DNA binding; no double-stranded (ds)DNA-binding domain has been identified in TFIIIE β .

DNA-binding ability of TFIIIE β c

When the DNA-binding activities of TFIIIE subunits were studied, TFIIIE β could bind to both ssDNA and dsDNA by itself but TFIIIE α could not bind directly and bound to both only when added together with TFIIIE β (Y.Ohkuma, unpublished results). However, the binding efficiencies of TFIIIE β were different between ssDNA and dsDNA, and the binding to ssDNA was stronger than to dsDNA. Since the ssDNA-binding region of TFIIIE β has already been identified in the C-terminal basic helix-loop region (Okamoto *et al.*, 1998), the dsDNA-binding regions were studied using the TFIIIE β internal deletion mutants mixed with dsDNA cellulose (Figure 5A). The dsDNA-bound mutants were detected by anti-TFIIIE β rabbit polyclonal antibody after Western blotting. Three mutants, Δ 96–119, Δ 117–153 and Δ 257–291, bound more weakly to dsDNA than the other mutants (lanes 6, 7 and 12 versus lanes 1–5 and 8–11). Importantly, mutants Δ 96–119 and Δ 117–153 lack either the N- or C-terminal region of the Rap30 homology region and mutant Δ 257–291 lacks the C-terminal basic helix-loop region (Figure 5C). To examine whether these regions in TFIIIE β bind to DNA, two domains of TFIIIE β , TFIIIE β c (amino acid residues 66–146) and TFIIIE β _{245–291} (amino acid residues 245–291), were bacterially expressed and purified, and both DNA-binding activities were tested (Figure 5B). As a result, TFIIIE β c bound predominantly to dsDNA (lane 3 versus lane 2) and TFIIIE β _{245–291} bound to both ssDNA and dsDNA although the binding to ssDNA was stronger than that to dsDNA (lane 5 versus lane 6). Therefore, the dsDNA-binding regions were localized in both the Rap30 homology region and the C-terminal basic helix-loop region. Mutant Δ 257–291 bound less effectively than the two mutants Δ 96–119 and Δ 117–153 to dsDNA (Figure 5A, lane 12 versus lanes 6 and 7). There are two possible reasons for this interpretation: one is that the mutant Δ 257–291 possesses only the TFIIIE β c region but the two mutants Δ 96–119 and Δ 117–153 possess half of the TFIIIE β c region in addition to the TFIIIE β _{245–291} region; the other is that deletion of residues 257–291 causes structural changes that hinder dsDNA binding at the TFIIIE β c region. Judging from these DNA-binding activities, the main dsDNA-binding domain of TFIIIE β is likely to be the central core domain of TFIIIE β , that is, TFIIIE β c.

To detect the dsDNA binding surface of TFIIIE β c, we examined the chemical shift changes of backbone amide proton and nitrogen signals of TFIIIE β c by adding dsDNA consisting of the –14 and –2 region of the adenovirus 2 major late (Ad2ML) promoter. By adding 0.5, 1.0, 1.5 and 2.0 molar ratios of dsDNA to TFIIIE β c the chemical

Fig. 5. DNA-binding experiments using DNA affinity chromatography. (A) Double-stranded DNA binding of TFIIIE β . A series of internal deletion mutants of TFIIIE β as well as wild type [300 ng, see (C)] were incubated with double-stranded DNA cellulose as described in Materials and methods. Bound mutants were detected by Western blotting with anti-TFIIIE β antisera. TFIIIE β , either wild type or mutants, are indicated at the top. The sizes of the molecular mass markers are indicated on the left (in kilodaltons). (B) Comparison of the DNA-binding activities between the core domain and the C-terminal domain of TFIIIE β . The core domain TFIIIE β c (IIE β c) and the C-terminal domain TFIIIE β _{245–291} (IIE β _{245–291}) were expressed in *E. coli* with a His₆ tag at the N-terminus and purified through Ni-NTA cellulose. Three hundred nanograms of both domains were incubated with either single-stranded (s) or double-stranded (d) DNA cellulose resin. 'c' is 10% input. Arrows on the right indicate the position of two polypeptides TFIIIE β c (His₆-IIE β c) and TFIIIE β _{245–291} (His₆-IIE β _{245–291}). The sizes of the molecular mass markers are indicated on the left (in kilodaltons). (C) Internal deletion mutants of TFIIIE β . All mutants were designed according to the putative structural motifs and characteristic sequences as shown below. The numbers on the left indicate the actual amino acid residues deleted and are used to denote each mutant (the deleted portion of TFIIIE β is shown as a gap).

Fig. 6. Chemical shift changes. (A) Overlay of the two-dimensional ^1H - ^{15}N HSQC spectra of TFIIIE β c without and with DNA. The spectrum of 1.0 mM TFIIIE β c alone is shown in black and the spectrum after the addition of 1.0 mM dsDNA is shown in red. The residues that showed significant chemical shift changes are labeled. (B) Chemical shift changes of amide-protons and nitrogens of the TFIIIE β c backbone on the DNA titration experiment. The backbone signals of Ala66–Lys74, Gly76, Lys114 and Asn126 could not be observed in this NMR condition. The weak backbone signal of His108 was only observed in the DNA free form. Chemical shift changes after addition of 1.0 mM DNA into 1.0 mM [^{15}N]TFIIIE β c are plotted in black and red for the positive and negative changes, respectively. (C) Mapping of the residues that have chemical shift changes onto the TFIIIE β c. The residues that showed relatively large chemical shift changes (absolute chemical shift differences >0.06 p.p.m. for NH, >0.50 p.p.m. for ^{15}N) are marked in red on the backbone of TFIIIE β c.

Fig. 7. Electrostatic potential surfaces of TFIIIE β c and winged helix proteins. The electrostatic potential surface of TFIIIE β c (residues 71–143), the DNA-binding domains of DP2 (Zheng *et al.*, 1999), HNF-3 γ (Clark *et al.*, 1993) and Rap30 (Groft *et al.*, 1998). The molecular surfaces are colored according to electrostatic potential. Positive potential is shown in blue and negative potential is in red. The surface around the HTH motif of each protein is shown in the upper part with the same orientation as in Figure 4. The lower part of each protein is rotated 180° relative to the axis of each protein of the upper part. The electrostatic potential surfaces were made using the program GRASP (Nicholls *et al.*, 1991).

shifts of backbone amide proton and nitrogen signals changed gradually. In addition, we examined the chemical shift changes of the imino proton signals of dsDNA by adding TFIIIE β c. Both spectral changes are explained by a fast exchange process. Figure 6A and B shows the chemical shift changes of amide proton and nitrogen signals for a 1:1 complex of TFIIIE β c with dsDNA. It is apparent that Val77, Leu78, Asn83, Lys86, Glu105, Thr106, Gln107, Asp110, Ile111, Lys129, Lys140 and Lys142 showed significant chemical shift changes upon binding to DNA. These amino acids were mapped on the structure of TFIIIE β c as shown in Figure 6C. The binding surface of TFIIIE β c with dsDNA is identified in helix H1, the loop between helices H2 and H3, and the N- and C-termini of the β hairpin, in contrast to the DNA-binding surface in the winged helix proteins (Figure 4). This is consistent with the results from assays using the other regions of the same Ad2ML promoter as well as other promoters (data not shown).

Figure 7 shows the electrostatic potential surface of TFIIIE β c. On the surface of TFIIIE β c, there is a positive concave patch, formed by Lys86, Lys129 and Lys142, which might be a binding site for the phosphate backbone of DNA. Figure 7 shows that each of the winged helix proteins, the DNA-binding domains of DP2, HNF-3 γ and Rap30 has a positive surface around the helix–turn–helix (HTH) motif, while the corresponding surface of TFIIIE β c is rather negative (upper panel). In particular, H2, the first helix of the HTH motif of TFIIIE β c, contains four negative amino acids in the amphipathic helix of Leu–Asp–Glu–Ile–Leu–Asp–Glu, whose sequence is highly conserved in other species (Figure 1B). The positive surface of TFIIIE β c is observed on the opposite surface instead (lower panel).

Therefore, the DNA-binding modes of TFIIIE β c revealed by the NMR experiment can be explained by the electrostatic potential surface of TFIIIE β c.

Function of TFIIIE β c in the preinitiation complex

Although TFIIIE β c contains a region similar to the Pol II binding site of Rap30 (TFIIF β), its structure is very close to the winged helix motif. The physiological Pol II binding region of TFIIIE β has been investigated using various deletion mutants of TFIIIE β , and the C-terminal region, rather than the central core domain, has been identified to bind to Pol II (Y.Watanabe and Y.Ohkuma, unpublished results). Taking into account the various roles of TFIIIE β , the central core domain is likely to be the dsDNA binding region. Actually, TFIIIE β c alone interacts with dsDNA, independently of nucleotide sequence (Figure 5B; data not shown). This relatively weak dsDNA binding ability of TFIIIE β c seems to be essential for a specific function of TFIIIE. It has been puzzling how PIC is activated and consequently promoter melting occurs. Recently, evidence has accumulated showing that TFIIIE is a key factor for elucidation. For example, the previous general transcription factor–promoter DNA photo-crosslinking experiment has indicated that TFIIIE actually binds to dsDNA between positions –14 and –2 upstream from the transcription start site (+1) including the region (around –10 to –8) where dsDNA starts to be single-stranded upon promoter melting (Robert *et al.*, 1996). Also, the two-dimensional crystallography of yeast TFIIIE (yTFIIIE) with Pol II has demonstrated that yTFIIIE binds to the active center of Pol II, which is located near the transcription initiation site (Leuther *et al.*, 1996). We suggest here that the weak but significant binding ability of TFIIIE β c to dsDNA is essential for the

regulation of the position of the C-terminal region of TFIIIE β onto DNA as well as Pol II, i.e. for the activation of PIC and promoter melting. In PIC formation, the correct position of TFIIIE on the promoter DNA should be arranged by already preassembled general transcription factors, such as TFIID, TFIIA, TFIIB and TFIIF as well as Pol II. During this process, TFIIIE will bind to dsDNA using the positively charged concave surface of TFIIIE β c (Figure 7), and to Pol II near the active center (Leuther *et al.*, 1996) at the C-terminal Pol II binding region (Y.Watanabe and Y.Ohkuma, unpublished results). After PIC formation, the single-stranded DNA-binding ability of the C-terminal basic region of TFIIIE β might be necessary for promoter melting. It is also possible that a special tertiary structure of DNA formed during promoter melting, rather than normal B form DNA, could be recognized tightly by TFIIIE β c. Actually and intriguingly, a novel DNA bend was observed upon PIC formation at the position between -5 and +5 containing the +1 transcription start site (Robert *et al.*, 1998). However, further examination is needed to unveil the exact role of TFIIIE β .

Materials and methods

Expression and purification of recombinant human TFIIIE β

Recombinant human TFIIIE β was expressed in *E.coli* BL21(DE3) pLysS by induction with isopropyl- β -D-thiogalactopyranoside (IPTG) (Studier *et al.*, 1990). For purification, soluble bacterial lysates were used. His₆-tagged human TFIIIE β (His₆-TFIIIE β) was purified through an Ni-nitrilotriacetic acid (NTA) column (Qiagen) by eluting with 100 mM imidazole-HCl pH 7.9. The large scale preparation of TFIIIE β was as described before (Sumimoto *et al.*, 1991), and resulted in >95% purity as judged by Coomassie Blue staining of an SDS-polyacrylamide gel.

Identification of the central core domain of TFIIIE β

Five hundred nanograms of His₆-TFIIIE β were partially digested by 0.25 ng of trypsin or α -chymotrypsin for 4 h at 25°C. Fragments containing the His₆-tagged N-terminus of TFIIIE β were separated through an Ni-NTA column and the core domain was identified in the unbound fraction; 10% of total core domain was separated by SDS-PAGE, transferred to the PVDF membrane, and the peptide sequence was determined by the peptide sequencer (ABI 492), the rest of the sample was separated by ODS 120T reverse-phase chromatography (TOSOH) and the separated core domain was either sequenced by the peptide sequencer (ABI 492) or weighed by mass spectrometry (MALDI/MASS, PerSeptive Biosystems). Following trypsin digestion, the peptide sequence 'ALSGSSGYKF' was determined by SDS-PAGE analysis and a little shorter sequence 'ALSGSSG' was determined by reverse-phase chromatography. The measured molecular mass was 9274.3 Da. Thus, this domain was mapped from amino acid residue 66 to 146 and the calculated molecular mass of this region was 9271.90 Da. Following α -chymotrypsin digestion, the peptide sequences 'KFGVLAK' and 'SVDEEFQ' were determined by SDS-PAGE analysis. The molecular masses were 9234.6 and 4061.2 Da, respectively. These correspond to amino acid residues 74–152 and 211–244. Taking these results into consideration, the region consisting of amino acid residues 66–146 was judged to be the TFIIIE β core domain.

Sample preparation

The nucleotide sequence of the central core domain, corresponding to amino acid residues 66–146 of human TFIIIE β , was amplified by the polymerase chain reaction (PCR) using two oligonucleotides, BS1-1T (5'-GACTGCATAGGCTTTGTCTCAGGAAGCTC-3') and BS1-1B (5'-CATCAGGATCCTTCTATCTACGTTGTACTTGG-3'), digested with *Nde*I and *Bam*HI, and subcloned into the *Nde*I and *Bam*HI restriction sites of the pET3a expression vector (Novagen). Therefore, the coding region contains methionine at the N-terminus. The plasmid was transformed into *E.coli* BL21(DE3) pLysS (Novagen). The *E.coli* cells were grown at 30°C in LB broth medium or in M9 minimal media containing [¹⁵N]ammonium chloride with or without [¹³C]glucose. IPTG (1 mM) was added at OD₆₀₀ = 0.4–0.5. After 5 h growth the cells were harvested.

The cell pellet was resuspended in buffer A (20 mM Tris-HCl pH 7.0, 10% glycerol, 1 mM EDTA, 1 mM PMSF, 1 mM benzamidine and 0.1 M KCl). The cells were lysed by sonication on ice and centrifuged, and the supernatant was loaded onto the phosphocellulose (P11, Whatman) column, equilibrated with buffer A. The protein sample was eluted by a KCl linear gradient from 0.1 to 1 M. The peak fractions were pooled and the buffer was changed to buffer B (20 mM Tris-HCl pH 7.0, 10% glycerol, 1 mM EDTA and 0.1 M KCl). Then, the sample was loaded onto the S-Sepharose (Amersham Pharmacia Biotech) column, equilibrated with buffer B, and was eluted by a KCl linear gradient from 0.1 to 0.5 M. The peak fractions were pooled. The sample was concentrated to get the volume below 3 ml using centriprep membrane with mol. wt 3000 cutoff, applied onto Superdex 30 (Amersham Pharmacia Biotech) equilibrated by 20 mM Tris-HCl pH 7.0, 10% glycerol, 1 mM EDTA and 1 M KCl, and the final sample fractions were collected.

NMR spectroscopy

Sample solution for NMR experiments in H₂O is ~200 μ l of 1.0–2.0 mM protein concentration in 20 mM potassium phosphate buffer pH 6.0, 500 mM NaCl and 10% (v/v) of D₂O was added. All NMR experiments were carried out at 27°C on either a Bruker AMX2-500 or DMX-600 spectrometer equipped with a triple-resonance gradient probe. The sequential resonances of backbone were assigned using 3D CBCA (CO)NH, CBCANH experiments (Grzesiek and Bax, 1992). The side chain resonances were assigned using 2D DQF-COSY (Rance *et al.*, 1983), TOCSY (Bax and Davis, 1985), 3D HBHA(CO)NH (Grzesiek and Bax, 1992), ¹⁵N-edited TOCSY-HSQC (Marion *et al.*, 1989) and HCCH-TOCSY (Kay *et al.*, 1993) experiments. Stereospecific assignments were obtained from a combination of 3D HNHB (Archer *et al.*, 1991), 2D DQF-COSY and 3D ¹⁵N-edited NOESY-HMQC (50 ms mixing time) (Kay *et al.*, 1989) experiments. Distance information was obtained from 2D NOESY (Macura and Ernst, 1980), 3D ¹⁵N-edited NOESY-HMQC (50 and 150 ms mixing time), 3D ¹³C-edited NOESY-HSQC (Muhandiram *et al.*, 1993) and 4D ¹³C, ¹³C-edited HMQC-NOESY-HSQC (100 ms mixing time). The backbone torsion angles ϕ were obtained from 3D HNHA (Vuister and Bax, 1993) and 2D HMQC-J (Kay and Bax, 1990) experiments. Spectra were processed using NMRPipe (Delaglio *et al.*, 1995), and analyzed using the programs PIPP, CAPP and STAPP (Garrett *et al.*, 1991).

Structure calculation

Interproton distance restraints derived from NOE intensities were grouped into three distance ranges, 1.9–3.0, 1.9–4.0 and 1.9–5.0 Å, corresponding to strong, medium and weak NOEs, respectively. 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, and 2.0 Å pseudoatom correction was added to degenerate phenylalanine and tyrosine ring protons. Dihedral angles estimated from ³J_{H α} coupling constants were restrained to -90° < ϕ < -40° for ³J_{H α} $\leq$ 5.0 Hz and -160° < ϕ < -80° for ³J_{H α} $\geq$ 8.0 Hz. Hydrogen bond restraints in areas of regular secondary structure were introduced at the final stages of refinement. Structure calculations were performed using the four-dimensional simulated-annealing program EMBOSS (Nakai *et al.*, 1993) on an Indigo 2 Impact Silicon Graphics workstation. A total of 100 simulated annealing structures were calculated, and 20 structures had no NOE violations larger than 0.3 Å and no dihedral angle violations greater than 1°. Structural statistics are summarized in Table I. The atomic coordinates of the 20 structures and the averaged structure have been deposited in the Protein Data Bank as 1D8J and 1D8K, respectively. Structures were analyzed and displayed using PROCHECK-NMR (Laskowski *et al.*, 1996), GRASP (Nicholls *et al.*, 1991), MOLMOL (Koradi *et al.*, 1996) and SYBYL (Tripos Inc., St Louis, MO).

Generation of polyclonal antibody against TFIIIE β

Two hundred micrograms (100 μ l) of purified His₆-TFIIIE β (>99% pure) were mixed with the same volume (100 μ l) of complete Freund's adjuvant (Difco) and injected into each rabbit. Two weeks after the first injection, a second injection was carried out with 100 μ g (100 μ l) of purified TFIIIE β in 100 μ l of incomplete Freund's adjuvant (Difco). A third injection was carried out 2 weeks later using the same procedures as described for the second injection. Blood was collected 8 days after the third injection. We have checked by Western blotting that the generated antibody recognized all of the TFIIIE β deletion mutants used in this study (Y.Ohkuma, data not shown).

DNA-binding assays by single- and double-stranded DNA cellulose

Three hundred nanograms of each TFIIE β mutant were incubated with 7 μ l (packed volume) of either ss or ds DNA cellulose (Sigma) in a 500 μ l reaction volume of buffer C [20 mM Tris-HCl pH 7.9 at 4°C, 0.5% EDTA, 20% (vol/vol) glycerol, 0.5 mM PMSF, 10 mM 2-mercaptoethanol, 0.002% (vol/vol) Nonidet P-40] containing 100 mM KCl (BC100) with 200 μ g/ml bovine serum albumin for 4 h at 4°C with rotation. The resin was then washed twice with 500 μ l of buffer C containing 200 mM KCl (BC200), once with 500 μ l of BC100, boiled in SDS sample buffer, and analyzed by SDS-PAGE (15% acrylamide). Bound mutants were detected by Western blotting with anti-TFIIE β antisera (1:3000 dilution) as described above.

DNA titration

Titration buffer comprised 4.5 mM potassium phosphate pH 7.2 in H₂O/D₂O (10/1). A 13 bp oligonucleotide duplex sequence was 5'-GCGCGTTCGTCCT-3', derived from Ad2ML promoter from -14 to -2. DNA titration was performed by adding 0.5 mM Ad2ML promoter into 1.0 mM [¹⁵N]TFIIE β c up to 2.0 mM. Chemical shift changes of backbone amide and nitrogen signals of TFIIE β c were monitored by 2D ¹H-¹⁵N HSQC spectrum (Kay *et al.*, 1992).

Acknowledgements

We thank Kosuke Morikawa (BERI) for helpful discussion. This work was supported by grants from the Mitsubishi Foundation (to Y.N.), in part by Grants in Aid for Scientific Research on Priority Areas (06276103, 06276104, 17109102, 111154221 and 11358012 to Y.N., and 09249211 to Y.O.) from the Ministry of Education, Science and Culture of Japan and by the Core Research for Evolutional Science and Technology (CREST) from the Japan Science and Technology Corporation (to Y.O.). M.O. is a JSPS Research Fellow.

References

- Archer, S.J., Ikura, M., Torchia, D.A. and Bax, A. (1991) An alternative 3D NMR technique for correlating backbone ¹⁵N with side chain H β resonances in larger proteins. *J. Magn. Reson.*, **95**, 636–641.
- Bagby, S., Kim, S., Maldonado, E., Tong, K.I., Reinberg, D. and Ikura, M. (1995) Solution structure of the C-terminal core domain of human TFIIB: Similarity to cyclin A and interaction with TATA-binding protein. *Cell*, **82**, 857–867.
- Bax, A. and Davis, D.G. (1985) MLEV-17 based two-dimensional homonuclear magnetization transfer spectroscopy. *J. Magn. Reson.*, **65**, 355–360.
- Birck, C., Poch, O., Romier, C., Ruff, M., Mengus, G., Lavigne, A.C., Davidson, I. and Moras, D. (1998) Human TAF_{II}28 and TAF_{II}18 interact through a histone fold encoded by atypical evolutionary conserved motifs also found in the SPT3 family. *Cell*, **94**, 239–249.
- Buissiere, D.E., Bastia, D. and White, S.W. (1995) Crystal structure of the replication terminator protein from *B. subtilis* at 2.6 Å. *Cell*, **80**, 651–660.
- Clark, K.L., Halay, E.D., Lai, E. and Burley, S.K. (1993) Co-crystal structure of the HNF-3/fork head DNA-recognition motif resembles histone H5. *Nature*, **364**, 412–420.
- 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.
- Drapkin, R., Reardon, J.T., Ansari, A., Huang, J.C., Zawel, L., Ahn, K., Sancar, A. and Reinberg, D. (1994) Dual role of TFIIF in DNA excision repair and in transcription by RNA polymerase II. *Nature*, **368**, 769–772.
- Dvir, A.K., Garrett, P., Chalut, C., Egly, J.-M., Conaway, J.W. and Conaway, R.C. (1996) A role for ATP and TFIIF in activation of the RNA polymerase II preinitiation complex prior to transcription initiation. *J. Biol. Chem.*, **271**, 7245–7248.
- Dvir, A., Conaway, R.C. and Conaway, J.W. (1997) A role for TFIIF in controlling the activity of early RNA polymerase II elongation complexes. *Proc. Natl Acad. Sci. USA*, **94**, 9006–9010.
- Feaver, W.J., Henry, N.L., Bushnell, D.A., Sayre, M.H., Brickner, J.H., Gileadi, O. and Kornberg, R.D. (1994) Yeast TFIIE. Cloning, expression and homology to vertebrate proteins. *J. Biol. Chem.*, **269**, 27549–27553.
- Finnin, M.S., Cicero, M.P., Davies, C., Porter, S.J., White, S.W. and Kreuzer, K.N. (1997) The activation domain of the MotA transcription factor from bacteriophage T4. *EMBO J.*, **16**, 1992–2003.
- Flores, O., Lu, H. and Reinberg, D. (1992) Factors involved in specific transcription by mammalian RNA polymerase II. *J. Biol. Chem.*, **267**, 2786–2793.
- Fogh, R.H., Otteleben, G., Rüterjans, H., Schnarr, M., Boelens, R. and Kaptein, R. (1994) Solution structure of the LexA repressor DNA-binding domain determined by ¹H NMR spectroscopy. *EMBO J.*, **13**, 3936–3944.
- 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.
- Goodrich, J.A. and Tjian, R. (1994) Transcription factors IIE and IIH and ATP hydrolysis direct promoter clearance by RNA polymerase II. *Cell*, **77**, 145–156.
- Groft, C.M., Uljon, S.N., Wang, R. and Werner, M.H. (1998) Structural homology between the Rap30 DNA-binding domain and linker histone H5: Implications for preinitiation complex assembly. *Proc. Natl Acad. Sci. USA*, **95**, 9117–9122.
- Grzesiek, S. and Bax, A. (1992) Correlating backbone amide and side chain resonances in larger proteins by multiple relayed triple resonance NMR. *J. Am. Chem. Soc.*, **114**, 6291–6293.
- Harper, E.T. and Rose, G.D. (1993) Helix stop signals in proteins and peptides: The capping box. *Biochemistry*, **32**, 7605–7609.
- Holm, L. and Sander, C. (1993) Protein structure comparison by alignment of distance matrices. *J. Mol. Biol.*, **233**, 123–138.
- Holstege, F.C.P., Tantin, D., Carey, M., van der Vliet, P.C. and Timmers, H.T.M. (1995) The requirement for the basal transcription factor IIE is determined by the helical stability of promoter DNA. *EMBO J.*, **14**, 810–819.
- Kay, L.E. and Bax, A. (1990) New methods for the measurement of NH-C α H coupling constants in ¹⁵N-labeled proteins. *J. Magn. Reson.*, **86**, 110–126.
- Kay, L.E., Marion, D. and Bax, A. (1989) Practical aspects of 3D heteronuclear NMR of proteins. *J. Magn. Reson.*, **84**, 72–84.
- Kay, L.E., Keifer, P. and Saarinen, T. (1992) Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity. *J. Am. Chem. Soc.*, **114**, 10663–10665.
- Kay, L.E., Xu, G.-Y., Singer, A.U., Muhandiram, D.R. and Forman-Kay, J.D. (1993) A gradient-enhanced HCH-TOCSY experiment for recording side-chain ¹H and ¹³C correlations in H₂O samples of proteins. *J. Magn. Reson. B*, **101**, 333–337.
- Kim, J.L., Nikolov, D.B. and Burley, S.K. (1993) Co-crystal structure of TBP recognizing the minor groove of a TATA element. *Nature*, **365**, 520–527.
- Kim, Y., Geiger, J.H., Hahn, S. and Sigler, P.B. (1993) Crystal structure of a yeast TBP/TATA-box complex. *Nature*, **365**, 512–520.
- Kondo, H., Nakagawa, A., Nishihira, J., Nishimura, Y., Mizuno, T. and Tanaka, I. (1997) *Escherichia coli* positive regulator OmpR has a large loop structure at the putative RNA polymerase interaction site. *Nature Struct. Biol.*, **4**, 28–31.
- Koradi, R., Billeter, M. and Wüthrich, K. (1996) MOLMOL: A program for display and analysis of macromolecular structures. *J. Mol. Graph.*, **14**, 51–55.
- Kumar, K.P., Akoulitchev, S. and Reinberg, D. (1998) Promoter-proximal stalling results from the inability to recruit transcription factor IIH to the transcription complex and is a regulated event. *Proc. Natl Acad. Sci. USA*, **95**, 9767–9772.
- Laskowski, R.A., Rullmann, J.A.C., MacArthur, M.W., Kaptein, R. and Thornton, J.M. (1996) AQUA and PROCHECK-NMR: Programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR*, **8**, 477–486.
- Leuther, K.K., Bushnell, D.A. and Kornberg, R.D. (1996) Two-dimensional crystallography of TFIIB- and IIE-RNA polymerase II complexes: Implications for start site selection and initiation complex formation. *Cell*, **85**, 773–779.
- Lima, C.D., Wang, J.C. and Mondragón, A. (1994) Three-dimensional structure of the 67K N-terminal fragment of *E. coli* DNA topoisomerase I. *Nature*, **367**, 138–145.
- Liu, D., Ishima, R., Tong, K.I., Bagby, S., Kokubo, T., Muhandiram, D.R., Kay, L.E., Nakatani, Y. and Ikura, M. (1998) Solution structure of a TBP-TAF_{II}230 complex: Protein mimicry of the minor groove surface of the TATA box unwound by TBP. *Cell*, **94**, 573–583.

- Lu, H., Flores, O., Weinmann, R. and Reinberg, D. (1991) The nonphosphorylated form of RNA polymerase II preferentially associates with the preinitiation complex. *Proc. Natl Acad. Sci. USA*, **88**, 10004–10008.
- Macura, S. and Ernst, R.R. (1980) Elucidation of cross relaxation in liquids by two-dimensional NMR spectroscopy. *Mol. Phys.*, **41**, 95–117.
- Marion, D., Driscoll, P.C., Kay, L.E., Wingfield, P.T., Bax, A., Gronenborn, A.M. and Clore, G.M. (1989) Overcoming the overlap problem in the assignment of ^1H NMR spectra of larger proteins by use of three-dimensional heteronuclear ^1H - ^{15}N Hartmann-Hahn multiple quantum coherence and nuclear Overhauser-multiple quantum coherence spectroscopy: Application to interleukin 1 β . *Biochemistry*, **28**, 6150–6156.
- Martinez-Hackert, E. and Stock, A.M. (1997) The DNA-binding domain of OmpR: crystal structure of a winged helix transcription factor. *Structure*, **5**, 109–124.
- Muhandiram, D.R., Farrow, N.A., Xu, G.Y., Smallcombe, S.H. and Kay, L.E. (1993) A gradient ^{13}C NOESY-HSQC experiment for recording NOESY spectra of ^{13}C -labeled proteins dissolved in H_2O . *J. Magn. Reson. B*, **102**, 317–321.
- Nakai, T., Kidera, A. and Nakamura, H. (1993) Intrinsic nature of the three-dimensional structure of proteins as determined by distance geometry with good sampling properties. *J. Biomol. NMR*, **3**, 19–40.
- Ni, J., Sakanyan, V., Charlier, D., Glansdorff, N. and Van Duyne, G.D. (1999) Structure of arginine repressor from *Bacillus stearothermophilus*. *Nature Struct. Biol.*, **6**, 427–432.
- Nicholls, A., Sharp, K.A. and Honig, B. (1991) Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins Struct. Funct. Genet.*, **11**, 281–296.
- 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.
- Nikolov, D.B. and Burley, S.K. (1997) RNA polymerase II transcription initiation: A structural view. *Proc. Natl Acad. Sci. USA*, **94**, 15–22.
- Ohkuma, Y. (1997) Multiple functions of general transcription factors TFIIE and TFIIH in transcription: Possible points of regulation by trans-acting factors. *J. Biochem.*, **122**, 481–489.
- Ohkuma, Y. and Roeder, R.G. (1994) Regulation of TFIIH ATPase and kinase activities by TFIIE during active initiation complex formation. *Nature*, **368**, 160–163.
- Ohkuma, Y., Sumimoto, H., Horikoshi, M. and Roeder, R.G. (1990) Factors involved in specific transcription by mammalian RNA polymerase II: Purification and characterization of general transcription factor TFIIE. *Proc. Natl Acad. Sci. USA*, **87**, 9163–9167.
- Ohkuma, Y., Sumimoto, H., Hoffmann, A., Shimasaki, S., Horikoshi, M. and Roeder, R.G. (1991) Structural motifs and potential σ homologues in the large subunit of human general transcription factor TFIIE. *Nature*, **354**, 398–402.
- Ohkuma, Y., Hashimoto, S., Roeder, R.G. and Horikoshi, M. (1992) Structural conservation of putative functional motifs between *Xenopus* and human TFIIE- β . *Nucleic Acids Res.*, **20**, 4363.
- Okamoto, T., Yamamoto, S., Watanabe, Y., Ohta, T., Hanaoka, F., Roeder, R.G. and Ohkuma, Y. (1998) Analysis of the role of TFIIE in transcriptional regulation through structure–function studies of the TFIIE β subunit. *J. Biol. Chem.*, **273**, 19866–19876.
- Orphanides, G., Lagrange, T. and Reinberg, D. (1996) The general transcription factors of RNA polymerase II. *Genes Dev.*, **10**, 2657–2683.
- Peterson, M.G., Inostroza, J., Maxon, M.E., Flores, O., Admon, A., Reinberg, D. and Tjian, R. (1991) Structure and functional properties of human general transcription factor IIE. *Nature*, **354**, 369–373.
- Ramakrishnan, V., Finch, J.T., Graziano, V., Lee, P.L. and Sweet, R.M. (1993) Crystal structure of globular domain of histone H5 and its implications for nucleosome binding. *Nature*, **362**, 219–223.
- Rance, M., Sørensen, O.W., Bodenhausen, G., Wagner, G., Ernst, R.R. and Wüthrich, K. (1983) Improved spectral resolution in COSY ^1H NMR spectra of proteins via double quantum filtering. *Biochem. Biophys. Res. Commun.*, **117**, 479–485.
- Robert, F., Forget, D., Li, J., Greenblatt, J. and Coulombe, B. (1996) Localization of subunits of transcription factors IIE and IIF immediately upstream of the transcriptional initiation site of the adenovirus major late promoter. *J. Biol. Chem.*, **271**, 8517–8520.
- Robert, F., Douziech, M., Forget, D., Egly, J.-M., Greenblatt, J., Burton, Z.F. and Coulombe, B. (1998) Wrapping of promoter DNA around the RNA polymerase II initiation complex induced by TFIIF. *Mol. Cell*, **2**, 341–351.
- Roeder, R.G. (1996) The role of general initiation factors in transcription by RNA polymerase II. *Trends Biochem. Sci.*, **21**, 327–335.
- Schultz, S.C., Shields, G.C. and Steitz, T.A. (1991) Crystal structure of a CAP–DNA complex: The DNA is bent by 90°. *Science*, **253**, 1001–1007.
- Studier, W., Rosenberg, A.H., Dunn, J.J. and Dubendorff, J.W. (1990) Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol.*, **185**, 60–89.
- Sumimoto, H., Ohkuma, Y., Sinn, E., Kato, H., Shimasaki, S., Horikoshi, M. and Roeder, R.G. (1991) Conserved sequence motifs in the small subunit of human general transcription factor TFIIE. *Nature*, **354**, 401–404.
- 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.
- Timmers, H.T.M. (1994) Transcription initiation by RNA polymerase II does not require hydrolysis of the β - γ phosphoanhydride bond of ATP. *EMBO J.*, **13**, 391–399.
- Vuister, G.W. and Bax, A. (1993) Quantitative J correlation: a new approach for measuring homonuclear three bond J -(H^{N} - H^{α}) coupling constants in ^{15}N -enriched protein. *J. Am. Chem. Soc.*, **115**, 7772–7777.
- Wah, D.A., Hirsch, J.A., Dorner, L.F., Schildkraut, I. and Aggarwal, A.K. (1997) Structure of the multimodular endonuclease FokI bound to DNA. *Nature*, **388**, 97–100.
- Wang, X., Hansen, S.K., Ratts, R., Zhou, S., Snook, A.J. and Zehring, W. (1997) *Drosophila* TFIIE: purification, cloning and functional reconstitution. *Proc. Natl Acad. Sci. USA*, **94**, 433–438.
- Wilson, K.P., Shewchuk, L.M., Brennan, R.G., Otsuka, A.J. and Matthews, B.W. (1992) *Escherichia coli* biotin holoenzyme synthetase/bio repressor crystal structure delineates the biotin- and DNA-binding domains. *Proc. Natl Acad. Sci. USA*, **89**, 9257–9261.
- Xie, X., Kokubo, T., Cohen, S.L., Mirza, U.A., Hoffmann, A., Chait, B.T., Roeder, R.G., Nakatani, Y. and Burley, S.K. (1996) Structural similarity between TAFs and the heterotetrameric core of the histone octamer. *Nature*, **380**, 316–322.
- Yokomori, K., Verrijzer, C.P. and Tjian, R. (1998) An interplay between TATA box-binding protein and transcription factors IIE and IIA modulates DNA binding and transcription. *Proc. Natl Acad. Sci. USA*, **95**, 6722–6727.
- Zheng, N., Fraenkel, E., Pabo, C.O. and Pavletich, N.P. (1999) Structural basis of DNA recognition by the heterodimeric cell cycle transcription factor E2F-DP. *Genes Dev.*, **13**, 666–674.
- Zhu, W., Zeng, Q., Colangelo, C.M., Lewis, L.M., Summers, M.F. and Scott, R.A. (1996) The N-terminal domain of TFIIB from *Pyrococcus furiosus* forms a zinc ribbon. *Nature Struct. Biol.*, **3**, 122–124.

Received August 16, 1999; revised November 8, 1999;
accepted January 27, 2000