Identification of two large subdomains in TFIIE-α on the basis of homology between *Xenopus* and human sequences

Yoshiaki Ohkuma[1], Shigeru Hashimoto[1], Robert G.Roeder[1] and Masami Horikoshi[1,2]
[1]Laboratory of Biochemistry and Molecular Biology, The Rockefeller University, New York, NY 10021, USA and [2]Institute of Applied Microbiology, University of Tokyo, Bunkyo-ku, Tokyo 113, Japan

Received September 29, 1992; Accepted October 6, 1992 EMBL accession no. Z14131

The general transcription factor TFIIE, a heterotetramer of $\alpha_2\beta_2$ structure (1), is one of the essential components in the transcription initiation pathway of RNA polymerase II (reviewed in 2). The deduced amino acid sequence of human TFIIE (hTFIIE) reveals several interesting structural motifs that suggest direct involvement in transcriptional regulation (3–5). Two complementary approaches have been employed to characterize structure–function relationships in TFIIE subunits. One, already completed for TFIIE-β (6), involves an analysis of the evolutionary conservation of primary structure between *Xenopus* TFIIE (xTFIIE) and hTFIIE subunits. The other involves the analysis of deletion and point mutants in putative structural motifs or conserved regions.

A hTFIIE-α cDNA fragment was used to isolate cDNAs encoding xTFIIE-α. Positive clones contained the complete open reading frame, which encodes a polypeptide of 433 amino acids with a calculated molecular mass of 49 kD. The primary structure of xTFIIE-α bears significant homology to hTFIIE- α (79.2% identity and 87.8% similarity) (see Figure 1) and most of the previously presented motifs are well conserved, as is the case for the TFIIE-β subunit (6). These motifs contain sequence similarities to σ-factors (Figure 1, boxes **a** and **b**), a leucine repeat (box **c**), a zinc finger motif (box **d**), a helix–turn–helix motif (box **e**) and a putative kinase consensus domain (box **g**). In addition, two sets of direct repeats, shown by arrows, are well conserved. A putative kinase domain (box **g**), proposed by Peterson *et al.* (5), requires two essential residues that are highly conserved among all kinases. Since one of these residues, aspartic acid 250 of the *Xenopus* protein does not fit the consensus, it is unlikely that this constitutes a functionally relevant kinase domain. Although the alanine rich regions (boxes **f** and **h**) are not conserved, the STDE region (box **i**), which contains a putative phosphorylation site, and the acidic region (box **j**) are well conserved, suggesting functional importance.

On the basis of homology this protein can be divided into two large subdomains, the amino-terminal half (residue 1–218) and the carboxy-terminal half (residue 219–433). The former is more highly conserved than the latter (90.8% vs 67.4% identity), and the collinearity of conservation is interrupted at eleven sites in the latter. This correlates with the fact that the putative functional motifs have only been identified in the former. This raises the interesting possibility that the amino-terminal half (1–218) may mediate essential core functions of the protein while the carboxy-terminal half may have a modulating function in transcription initiation. To demonstrate the functional relevance of the proposed motifs and the respective functions of the two large subdomains, we are currently constructing and analyzing relevant mutant proteins.

ACKNOWLEDGMENTS

We thank Alexander Hoffmann and Beth Moorefield for critical reading of the manuscript, and Drs Jerry Thomsen and Doug A.Melton for a *Xenopus laevis* cDNA library. Y.O. was supported by a fellowship from the Human Frontiers Scientific Program Organization and M.H. was an Alexandrine and Alexander L.Sinsheimer Scholar. This study was supported by NIH grants to R.G.R. and M.H. and by general support from the Pew Trusts to the Rockefeller University.

REFERENCES

1. Ohkuma,Y., *et al.* (1990) *Proc. Natl. Acad. Sci. USA* **87**, 9163–9167.
2. Roeder,R.G. (1991) *TIBS* **16**, 402–408.
3. Ohkuma,Y., *et al.* (1991) *Nature* **354**, 398–401.
4. Sumimoto,H., *et al.* (1991) *Nature* **354**, 401–404.
5. Peterson,M.G., *et al.* (1991) *Nature* **354**, 369–373.
6. Ohkuma,Y., *et al.* (1992) *Nucleic Acids Res.* **20**, 4363.

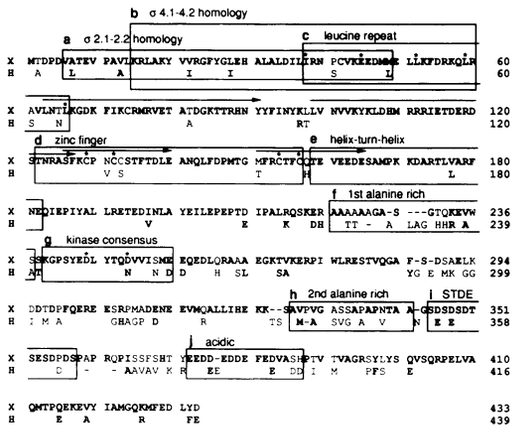

Figure 1. Deduced amino acid sequence of *Xenopus* TFIIE-α and its comparison with human TFIIE-α. For human (**H**) TFIIE-α, only the amino acids different from *Xenopus* (**X**) TFIIE-*et al.* are shown in the second lines. Identical or conserved amino acids are in bold according to the previously described grouping (6). Putative structural motifs and characteristic regions are boxed (boxes **a**–**j**). Arrows indicate the direct repeats. Relevant residues of the leucine repeat, the kinase consensus and the zinc finger motif are marked (* on top).