Structural conservation of putative functional motifs between *Xenopus* and human TFIIE-β

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

Submitted July 7, 1992 EMBL accession no. Z12595

Transcription initiation by RNA polymerase II requires a number of general factors that include TFIIE (reviewed in 1). Previous studies have demonstrated that TFIIE is comprised of 57 kD (TFIIE-α) and 34 kD (TFIIE-β) subunits which form a heterotetramer ($\alpha_2\beta_2$) (2), requires both α and β subunits for transcription initiation (3–5), binds to RNA polymerase II (6), and joins the preinitiation complex after RNA polymerase II and TFIIF (5). Although little is known about the precise role of TFIIE in transcriptional regulation, the characterization of structure–function relationships in TFIIE could provide relevant information. The deduced amino acid sequence of human TFIIE-β (hTFIIE-β) revealed several motifs previously implicated in DNA–protein and/or protein–protein interactions (4). To identify regions of high evolutionary conservation, and thus of presumptive functional importance, we isolated cDNAs encoding *Xenopus laevis* TFIIE-β (xTFIIE-β).

A hTFIIE-β cDNA was used as a probe to isolate a homologue from a *Xenopus* oocyte cDNA library. Clones containing a complete open reading frame (ORF) were sequenced on both strands. This putative ORF encodes a 288-amino acid polypeptide with a calculated molecular mass of 32.6 kD. Comparison of the amino acid sequence of xTFIIE-β with that of hTFIIE-β indicates 84.0% identity and 93.1% similarity with deletions at two sites (five residues) and insertions at two sites (two residues) (see Figure 1). Generally, all of the putative motifs noted previously (4, 5) are well conserved, suggesting that they are functionally relevant. These motifs include a leucine repeat region (Figure 1, box **d**), a basic region–helix–loop motif (box **g**), another recently-recognized basic region–helix–loop motif (box **f**), a potential nucleotide binding region (box **c**), and sequence similarities to σ-factors (boxes **b** and **e**), including the subregion (2.1–2.2) implicated in binding to RNA polymerase (4). Although relatively frequent amino acid changes are observed in the amino- and carboxy-terminal regions, especially in the serine-rich region (Figure 1, box **a**) and the basic and loop regions within the second basic region–helix–loop motif (box **g**), putative functional features (such as charge and potential helicity) are highly conserved. Except for the amino- and carboxy-terminal regions, the whole amino acid sequence of *Xenopus* and human can be aligned colinearly without any deletions and insertions.

Knowledge of the evolutionarily conserved sequences in hTFIIE-β and xTFIIE-β will help in the design of structure–function studies.

ACKNOWLEDGMENTS

We thank Dr Robert Kovelman and Alexander Hoffmann 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. Roeder,R.G. (1991) TIBS **16**, 402–408.
2. Ohkuma,Y., *et al.* (1990) *Proc. Natl. Acad. Sci. USA* **87**, 9163–9167.
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. Flores,O., *et al.* (1989) *J. Biol. Chem.* **264**, 8913–8921.

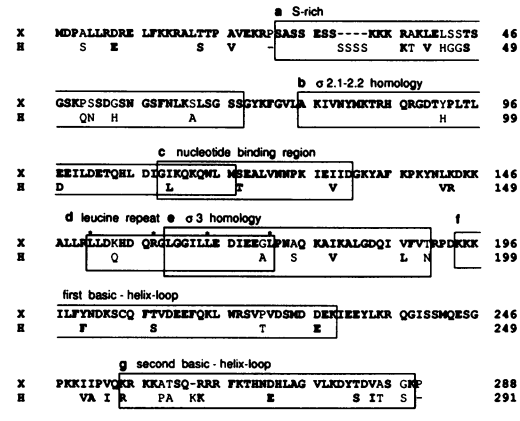

Figure 1. Deduced amino acid sequence of *Xenopus* TFIIE-β and its comparison with human TFIIE-β. For human (**H**) TFIIE-β, only the amino acids different from *Xenopus* (**X**) TFIIE-β are shown in the second lines. Identical or conserved amino acids are in bold according to the grouping (K,R,H), (D,E,N,Q), (A,I,L,V,M), (F,W,Y), (C), (G), (P) and (S,T). Putative structural motifs and characteristic regions are boxed: **a**; serine rich region, **b**; region with similarity to σ subregion 2.1–2.2, **c**; potential nucleotide binding domain, **d**; leucine repeat, **e**; region with similarity to σ region 3, **f**; 1st basic region–helix–loop motif, **g**; 2nd basic region–helix–loop motif. Relevant residues of the seven-amino acid leucine repeat are marked (* on top).

* To whom correspondence should be addressed