Conserved structural motifs between *Xenopus* and human TFIIB

Koji Hisatake, Sohail Malik, Robert G. Roeder and Masami Horikoshi*

Laboratory of Biochemistry and Molecular Biology, The Rockefeller University, New York, NY 10021, USA

Submitted November 4, 1991

EMBL accession no. X62868

TFIIB is a general transcription initiation factor that binds to the TFIID-promoter complex and recruits RNA polymerase II/TFIIF to the template, thus facilitating the formation of a functional preinitiation complex as reviewed in (1, 2). Like TFIID (3–6), TFIIB is proposed to be involved in transcriptional activation and may be one of the direct targets of acidic activators such as VP16 (7, 8). Therefore, understanding the structural basis of TFIIB function and its interactions with other general and/or regulatory transcription factors is particularly important for unraveling the mechanisms of basal and activator-dependent transcription. Recently, the human TFIIB cDNA has been cloned (9, 10). Sequence analysis revealed that TFIIB contains several interesting structural motifs that include a homology to sigma region 2.1/2.2 (10) or 2.4 (9), direct repeats (9, 10), basic regions (9, 10), an RNA polymerase homology (9), as well as an overall structural organization similar to that of TFIID (10). To define the structural domains potentially important for the various TFIIB functions, we have sought to identify the evolutionarily conserved motifs by characterizing TFIIB from other species.

Using our recently cloned human TFIIB cDNA as a probe, we isolated *Xenopus laevis* TFIIB cDNAs from an oocyte cDNA library. One of the cDNAs, pX19, contained a complete open reading frame and was subjected to sequencing on both strands. The deduced amino acid sequence predicts a 316 amino acid protein with calculated molecular weight of 34.7 kD. Amino acid sequence comparison with human TFIIB (Figure 1) indicates 94% identity with complete continuity. Most of the amino acid changes are found in the amino terminal region prior to the domain containing the structural motifs, suggesting that this region may be more divergent among different species. On the other hand, the direct repeats, sigma homology and basic regions are strongly conserved between these two species, underscoring the functional importance of these motifs. The well-conserved direct repeats in the carboxy terminus with a relatively more divergent amino terminal region is also reminiscent of the situation with TFIID.

We are currently trying to obtain TFIIB cDNAs from several different species and to deduce potentially important domains in TFIIB by comparing amino acid sequences. TFIIB mutants will also be tested in various assay systems to define functionally important domains. These studies will elucidate structural basis of TFIIB interactions with other transcriptional components and its function in transcriptional initiation.

ACKNOWLEDGMENTS

We thank Drs Jerry Thomsen and Doug A. Melton for a generous gift of *Xenopus laevis* cDNA library, Dr Hisakazu Fujita for

preparing the filters used for screening and Josephine Wang for her technical assistance in DNA sequencing. S.M. is supported by Fellowship GM13244 from the National Institutes of Health. M.H. is an Alexandrine and Alexander L. Sinsheimer Scholar. This work was supported by NIH grants CA42567, AI27397 to R.G.R. and GM45258 to M.H., by funds from Sankyo Co. Ltd. to M.H. and by general support from the Pew Trust to The Rockefeller University.

REFERENCES

1. Saltzman, A.G. and Weinmann, R. (1989) *FASEB J.* **3**, 1723–1733.
2. Sawadogo, M. and Sentenac, A. (1990) *Annu. Rev. Biochem.* **59**, 711–754.
3. Horikoshi, M., Carey, M.F., Kakidani, H. and Roeder, R.G. (1988) *Cell* **54**, 665–669.
4. Horikoshi, M., Hai, T., Lin, Y.-S., Green, M.R. and Roeder, R.G. (1988) *Cell* **54**, 1033–1042.
5. Stringer, K.F., Ingles, C.J. and Greenblatt, J. (1990) *Nature* **345**, 783–786.
6. Ingles, C.J., Shales, M., Cress, W.D., Triezenberg, S.J. and Greenblatt, J. (1991) *Nature* **351**, 588–590.
7. Lin, Y.-S. and Green, M.R. (1991) *Cell* **64**, 971–981.
8. Lin, Y.-S., Ha, I., Maldonado, E., Reinberg, D. and Green, M.R. (1991) *Nature* **353**, 569–571.
9. Ha, I., Lane, W.S. and Reinberg, D. (1991) *Nature* **352**, 689–695.
10. Malik, S., Hisatake, K., Sumimoto, H., Horikoshi, M. and Roeder, R.G. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 9553–9557.

Xenopus	MASTRIDALPKVTCNPHNDALLVEDYRAGDMICSEGLV	40
Human	-----L-----R-----I-----P-----	
Xenopus	VGDRIIDVGSEWRFTSNDKAAADPSRVGDAQNPILSGDDL	80
Human	-----S-----TK-----D-----	
Xenopus	TTMIGKGTGASAFDEFGSKYQNRRTMSSSDRAMQAFKE	120
Human	S-----A-----	
Xenopus	YTNSDRMLNPIITORTNMLEKQVYKSKLGRNDAL	160
Human	--T-A-----V-----A-----	
Xenopus	SACLYIACRQEGVPRFTKCAVSRSSKGLRGKFLIK	200
Human	-----	
Xenopus	ALETNVDLITTTGDFSRFCSMLEKINQMAATHLANAKV	240
Human	-----S-----C-----	
Xenopus	ELDLVPGSPISVAAATYASQASAEKRRKQKIGDLAGV	280
Human	-----	
Xenopus	ADVTIRQSYLTLYPRAPDLFPADFEDTPVDKFLQL	316
Human	-----	

Figure 1. Amino acid sequence of *Xenopus laevis* TFIIB and its comparison with human TFIIB. Amino acid sequences of *Xenopus* and human TFIIB are shown in the first and second lines, respectively. For human TFIIB, only the amino acids different from *Xenopus* TFIIB are shown. Arrows indicate the direct repeats. Amino acids conserved between the direct repeats are shaded. In *Xenopus* TFIIB, a non-conservative change from proline to threonine was found in one of those amino acids and is indicated by outlined letters. Bold-faced letters in the sigma homology regions indicate the residues conserved among TFIIB and the regions 2.1 and 2.2 of sigma factors. Basic amino acids in the basic regions are shown by stars.