Imperfect conservation of a sigma factor-like subregion in *Xenopus* general transcription factor RAP30

Da-Wei Gong, Shigeru Hashimoto², Kenji Wada¹, Robert G. Roeder², Yoshihiro Nakatani and Masami Horikoshi^{2,3,*}

Laboratory of Molecular Biology, NINDS, ¹Laboratory of Neurochemistry, NIDCD, NIH, Bethesda, MD 20892, ²Laboratory of Biochemistry and Molecular Biology, The Rockefeller University, New York, NY 10021, USA and ³Institute of Applied Microbiology, University of Tokyo, Bunkyo-ku, Tokyo 113, Japan

Received October 21, 1992; Revised and Accepted October 27, 1992

EMBL accession no. Z15132

The initiation of transcription in eukaryotes requires multiple general factors and RNA polymerase (1). RAP30 (as well as RAP74) was initially identified as an RNA polymerase II associated protein by affinity chromatography on immobilized RNA polymerase II (2). Later, using an antibody against RAP30, it was found that RAP30 corresponds to the small subunit of TFIIF (3). TFIIF is a tetramer of two RAP30 and two RAP74 subunits. Recent data indicate that RAP30 participates in recruitment of RNA polymerase II into the pre-existing TFIID-TFIIB-promoter complex (4). Furthermore, it has been shown that TFIIF plays a role in the elongation process (5). Therefore, understanding the structure-function relationships of this unique and interesting factor will help elucidate the mechanisms involved in eukaryotic transcription initiation and elongation. The cDNA for human RAP30 contains a region encoding a segment similar to subregions 1b and 2.1 of the prokaryotic transcription initiation factor sigma (6). Previously it was shown that sigma subregion 2.1 is involved in the sigma interaction with core RNA polymerase (7). In dramatic confirmation of the proposed homology, RAP30 has been recently shown to competitively interact with the prokaryotic core RNA polymerase (8). This observation postulates the presence of a universal functional and structural domain between prokaryotic sigma and eukaryotic RAP30.

Here we report the cDNA cloning of *Xenopus* RAP30 and compare its primary structure to that of the human RAP30 (Figure 1). *Xenopus* RAP30 cDNAs were isolated by screening a *Xenopus* oocyte cDNA library under low stringency conditions using human RAP30 cDNA as a probe. The deduced amino acid sequence of *Xenopus* RAP30 shows 83% identity and 91% similarity to the human sequence. Although several insertions are present in the N-terminal half of *Xenopus* RAP30, the C-terminal half has no insertions or deletions. This suggests the presence of two large structural subdomains in RAP30: a C-terminal half, containing the sigma-similarity region, which might have a core function and an N-terminal half which might have some type of regulatory function. Surprisingly, *Xenopus* RAP30 has an insertion of 12 amino acids within the region corresponding to the sigma subregion 1b in human RAP30 (6). Thus, it is likely that only the region corresponding to sigma subregion 2.1 and its surrounding region is responsible for interaction with RNA

polymerase II (7). The data on the sequence of *Xenopus* RAP30 should be useful for additional structural and functional studies.

ACKNOWLEDGEMENTS

We thank Dr Mike Brenner for a critical reading of the manuscript and Drs Jerry Thomsen and Doug A. Melton for the *Xenopus laevis* cDNA library. M.H. was an Alexandrine and Alexander L. Sinsheimer Scholar. A part of this study was supported by grants from the National Institutes of Health to R.G.R. and M.H. and by the Pew Trusts to The Rockefeller University.

REFERENCES

- Roeder, R.G. (1991) *Trends Biochem. Sci.* **16**, 402-408.
- Sopta, M., Carthew, R.W. and Greenblatt, J. (1985) *J. Biol. Chem.* **260**, 10353-10361.
- Flores, O., Maldonado, E., Burton, Z.F., Greenblatt, J. and Reinberg, D. (1988) *J. Biol. Chem.* **263**, 10812-10816.
- Flores, O., Lu, H., Killeen, M., Greenblatt, J., Burton, Z.F. and Reinberg, D. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 9999-10003.
- Flores, O., Maldonado, E. and Reinberg, D. (1989) *J. Biol. Chem.* **264**, 8913-8921.
- Sopta, M., Burton, Z.F. and Greenblatt, J. (1989) *Nature* **341**, 410-414.
- Lesley, S.W. and Burgess, R.R. (1989) *Biochemistry* **28**, 7728-7734.
- McCracken, S. and Greenblatt, J. (1991) *Science* **253**, 900-902.
- Horikoshi, M., Fujita, H., Wang, J., Takada, R. and Roeder, R.G. (1991) *Nucleic Acids Res.* **19**, 5436.

Xenopus	1	MSAEKGELDLNGAKONTGMVLKPKYLAQOWAKATGRGEVGLKRLVKNO	50
	1	- R T V V S S A T	49
human	51	GKTEVSFTLNEELASIQDIGGKPASSLVSTPREHFFLLQSVGGQTLVLT	100
	50	R D N H -- A V F	97
	101	ESLSGQSEDKSENVRVIDKLALGIVVHRAECRPAASDNYMQRKQTEES	150
	98	S ----- S Q E RL L	135
	151	SKPKRQSQOLEKAVTSYKPVSHQYNIYEYERKKDDGKRAVDKHKVLD	200
	136	V L D V T A R E A QH	185
	201	MLFSAFEKHQYNIKLDVDTQKPVYTLKEILRDIGIYNNKGTHTKNTWEL	250
	186	L V KE VQ V I	235
	251	KPEYRHYQGDKSD	264
	236	E	249

Figure 1. Amino acid sequence of *Xenopus laevis* RAP30 and its comparison with human RAP30 (6, 9). For human RAP30, only the amino acids different from *Xenopus* RAP30 are shown. Bar (—) indicates deleted amino acids in human RAP30. Sigma-similarity regions are indicated as region 1b, 2.1 and 2.2 according to Sopta *et al.* (6). Putative nucleotide binding domain (a) and CREB-like DNA binding domain (b), presented by Sopta *et al.* (6), are shown by lines.

* To whom correspondence should be addressed at: Laboratory of Biochemistry and Molecular Biology, The Rockefeller University, New York, NY 10021, USA