Cell, Vol. 71, 189–190, October 16, 1992, Copyright © 1992 by Cell Press

Letter to the Editor

TFIIB, an Evolutionary Link between the Transcription Machineries of Archaebacteria and Eukaryotes

We report here the identification by sequence similarity of a homolog of the eukaryotic transcription factor TFIIB in the archaebacterium Pyrococcus woesei. This finding has interesting implications for the evolution of the molecular machinery that controls and executes transcription.

In eukaryotes, protein-encoding genes are transcribed by RNA polymerase II, modulated by general and specific transcription factors (Sawadogo and Sentenac, 1990; Parvin et al., 1992). General transcription factors associate with promoter elements and interact with each other in a controlled and stepwise fashion. The general transcription factor TFIIB associates with TFIID bound to DNA in the TATA box region of the type II promoter (Sharp, 1991). This binding permits entry of RNA polymerase II into the complex, subsequent association of other general factors, and initiation of transcription (Hawley, 1991).

The sequence of human TFIIB (Ha et al., 1991; Malik et al., 1991) is similar to that of rat (99% homology; Tsuboi et al., 1992), Xenopus (94% homology; Hisatake et al., 1991), Drosophila (79% homology; Yamashita et al., 1992; Wampler and Kadanoga, unpublished data), and yeast (34% homology; Pinto et al., 1992). Archaebacterial transcription factors have so far not been identified. However, the inability of purified archaebacterial RNA polymerase to initiate transcription correctly can be restored by addition of a cell extract. This fact is circumstantial evidence for the existence of mechanisms regulating precise initiation of transcription in archaebacteria (Thomm et al., 1989; Frey et al., 1990). Taken together with the presence of TATA-like promoter sequences in archaebacteria (Reiter et al., 1990), this suggests a common origin of the transcriptional machinery of archaebacteria and eukaryotes.

After searching nucleotide and protein sequence data bases for TFIIB homologs (Pearson and Lipman, 1988), we have now found that a previously undetected partial open reading frame in P. woesei (Creti et al., 1991) is significantly similar to eukaryotic TFIIBs (Figure 1). This open reading frame is downstream of the operon for elongation factor EF-1a, ribosomal protein S10, and $tRNA_{Ser}$

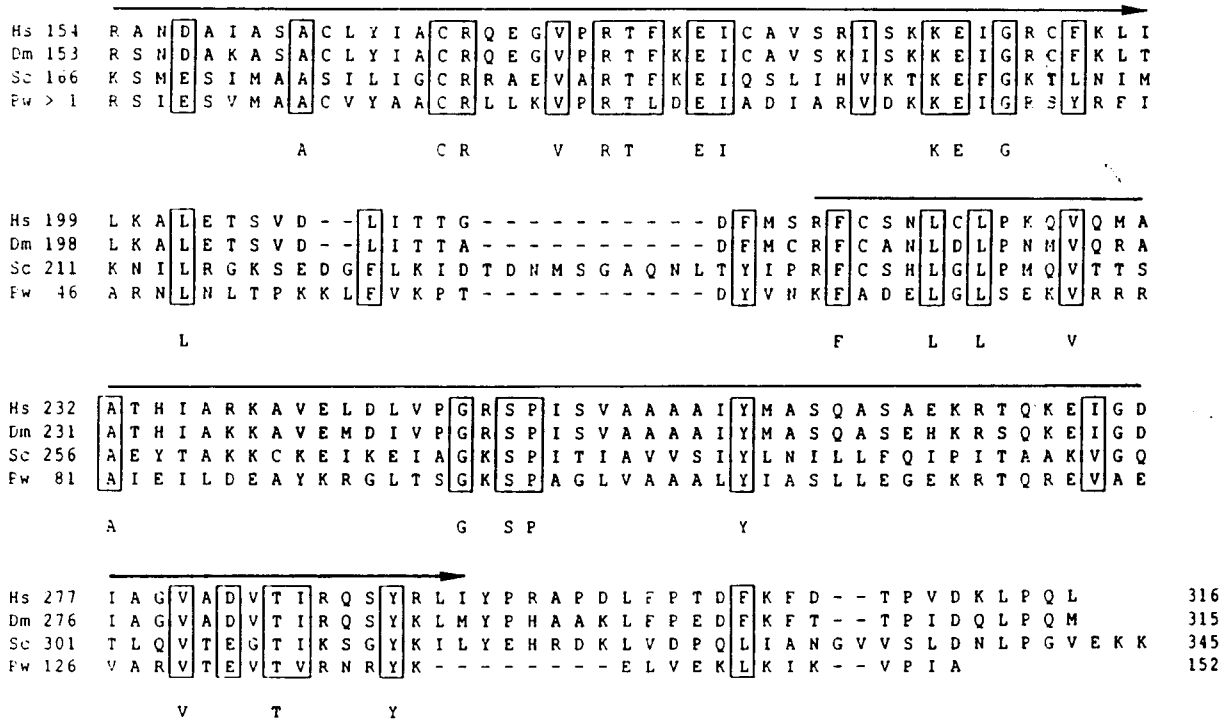

Figure 1. Aligned Representative Sequences from the TFIIB Family

Hs, human; Dm, Drosophila; Sc, yeast; Pw, archaebacterium P. woesei. Sequence positions are noted at the N- and C-termini; those for P. woesei are arbitrary, as the sequence is incomplete at its N-terminus. Sequences of the rat and Xenopus homologs are not shown, but residues conserved in all six sequences are in the bottom line. The (incomplete) first copy and the second copy of the putative internal repeat are indicated by arrows. Sequence references with data base accession numbers in parentheses follow. Human sequences: Ha et al., 1991 (X59268); Malik et al., 1991 (M76766). Drosophila sequences: Yamashita et al., 1992 (M88164); Wampler and Kadonaga, unpublished data (M91081). Yeast sequences: Pinto et al., 1992 (M81380); Hahn and Colbert, unpublished data (M91073).

on the complementary strand. The partial sequence represents the C-terminal 152 amino acids of an archaebacterial factor that probably has a total length of about 330 residues.

The archaebacterial TFIIB appears to be about equidistant from its human, Drosophila, and yeast cousins (32%–36% identity over a length of more than 150 amino acids). This degree of sequence similarity is highly significant, as it is 7–11 percentage points above the threshold for structural homology established by Sander and Schneider (1991). The conservation implies a similar three-dimensional structure and is suggestive of a related function in a transcriptional initiation complex.

The extended family alignment (Figure 1) defines a small set of strictly conserved residues in the TFIIB family. Most of these are likely to be conserved because they have a direct functional role and are prime candidates for molecular genetic analysis. Positions that differ characteristically in the Pyrococcus sequence may be related to the extreme thermophilic nature of this archaebacterium.

What are the evolutionary implications of the existence of TFIIB in archaebacteria? It is by now well established that archaebacteria are more closely related to eukaryotes than eubacteria, as a result of sequence comparisons of primordial molecules such as rRNA, elongation factors, and DNA-dependent RNA polymerases (Woese et al., 1983; Cammarano et al., 1992; Pühler et al., 1989). TFIIB provides an additional strong link: both eukaryotes and archaebacteria have this type of transcription factor, while eubacteria apparently do not. The fact that both RNA polymerase and at least one general transcription factor are conserved between archaebacteria and eukaryotes suggests that there is homology not only between the individual molecules but also between the overall organization of the transcriptional apparatus. This immediately leads to the prediction that TFIID as well as TFIIB is present in archaebacteria.

The observation that TFIIB is conserved between eukaryotes and archaebacteria may also be useful in the context of the recent debate about the appropriate classification of sulfur-dependent thermophilic archaebacteria, sometimes called "eocytes." Do these belong with the domain of Archaea (Woese et al., 1990) or are they in a separate domain on the way from Archaea to Eucarya (Rivera and Lake, 1992)? Comparison of the sequences of TFIIB from these organisms would help resolve this issue with regard to the evolution of the mechanism of transcription.

Irrespective of the outcome of the classification issue, it is already evident that essential aspects of the eukaryotic transcriptional machinery were in place before the divergence of archaebacteria and eukaryotes—much earlier in evolution than previously thought.

Christos Ouzounis and Chris Sander
Protein Design Group
European Molecular Biology Laboratory
Heidelberg, D-6900
Germany

References

Cammarano, P., Palm, P., Creti, R., Ceccarelli, E., Sanangelantoni, A. M., and Tiboni, O. (1992). J. Mol. Evol. *34*, 396–405.

Creti, R., Citarella, F., Tiboni, O., Sanangelantoni, A., Palm, P., and Cammarano, P. (1991). J. Mol. Evol. *33*, 332–342.

Frey, G., Thomm, M., Brudigam, B., Gohl, H. P., and Hausner, W. (1990). Nucl. Acids Res. *18*, 1361–1367.

Ha, I., Lane, W. S., and Reinberg, D. (1991). Nature *352*, 689–695.

Hawley, D. (1991). Trends Biochem. Sci. *16*, 317–318.

Hisatake, K., Malik, S., Roeder, R. G., and Horikoshi, M. (1991). Nucl. Acids Res. *19*, 6639.

Malik, S., Hisatake, K., Sumimoto, H., Horikoshi, M., and Roeder, R. G. (1991). Proc. Natl. Acad. Sci. USA *88*, 9553–9557.

Parvin, J. D., Timmers, H. T. M., and Sharp, P. A. (1992). Cell *68*, 1135–1144.

Pearson, W. R., and Lipman, D. J. (1988). Proc. Natl. Acad. Sci. USA *85*, 2444–2448.

Pinto, I., Ware, D. E., and Hampsey, M. (1992). Cell *68*, 977–988.

Pühler, G., Leffers, H., Gropp, F., Palm, P., Klenk, H.-P., Lottspeich, F., Garrett, R. A., and Zillig, W. (1989). Proc. Natl. Acad. Sci. USA *86*, 4569–4573.

Reiter, W. D., Hudepohl, U., and Zillig, W. (1990). Proc. Natl. Acad. Sci. USA *87*, 9509–9513.

Rivera, M. C., and Lake, J. A. (1992). Science *257*, 74–76.

Sander, C., and Schneider, R. (1991). Proteins *9*, 56–68.

Sawadogo, M., and Sentenac, A. (1990). Annu. Rev. Biochem. *59*, 711–754.

Sharp, P. A. (1991). Nature *351*, 16–18.

Thomm, M., Wich, G., Brown, J. W., Frey, G., Scherf, B. A., and Beckler, G. S. (1989). Can. J. Microbiol. *35*, 30–35.

Tsuboi, A., Conger, K., Garrett, K. P., Conaway, R. C., and Arai, N. (1992). Nucl. Acids Res. *20*, 3250.

Woese, C. R., Gutell, R., Gupta, R., and Noller, H. F. (1983). Microbiol. Rev. *47*, 621–669.

Woese, C. R., Kandler, O., and Wheelis, M. L. (1990). Proc. Natl. Acad. Sci. USA *87*, 4576–4579.

Yamashita, S., Wada, K., Horikoshi, M., Gong, D.-W., Kokubo, T., Hisatake, K., Yokotani, N., Malik, S., Roeder, R. G., and Nakatani, Y. (1992). Proc. Natl. Acad. Sci. USA *89*, 2839–2843.