

Putative *tflls* gene of *Sulfolobus acidocaldarius* encoding an archaeal transcription elongation factor is situated directly downstream of the gene for a small subunit of DNA-dependent RNA polymerase

Doris Langer and Wolfram Zillig*

Max-Planck-Institut für Biochemie, Am Klopferspitz 18, 8033 Martinsried, Germany

Received March 18, 1992; Accepted March 24, 1993

EMBL accession no. X70805

Transcription in Archaea resembles in several respects that in Eucarya rather than that in Bacteria both regarding the components of the transcription machinery as well as the promoter structure. The subunit pattern of the DNA-dependent RNA polymerase (RNAP) of Archaea is of similar complexity as that of Eucarya: so far seven RNAP components (B, A', A'', H, K, L and N) were found to be homologous to eucaryal RNAP subunits (1, 2, 3). The sequence and the arrangement of the promoter elements, especially the distance of the RNA start site, resembles that of the eucaryotic TATA box containing polymerase II promoters (4). Finally an archaeal homolog to the eucaryal transcription initiation factor TFIIB has recently been identified (5).

Here we show that the archaeon *Sulfolobus acidocaldarius* also possesses a gene with significant similarities to members of the TFIIS family of transcription elongation factors which so far was only represented in eucarya (6, 7, 8). In some of these cases TFIIS was found to be part of both the initiation and the elongation complex (9), or was shown to help the RNAP to overcome pausing during transcription and it was suggested that TFIIS plays a role in the regulation of elongation or termination of heterogeneous nuclear RNA (10).

In *S.acidocaldarius* we found an ORF encoding a 23 kD protein with significant homology to eucaryal TFIIS factors directly downstream of the gene encoding subunit L of the RNA polymerase. The sequence similarity is most pronounced in the C-terminal region (see Figure 1) and especially in the last forty amino acids which are also best conserved within the TFIIS family (50% identity). This highly conserved sequence contains a zinc finger motif with the structure Cys-X₂-Cys-X₂₄-Cys-X₂-Cys, which was shown to be involved in the stimulation of transcription elongation (11). Eucaryal TFIIS factors contain two conserved domains separated by a variable serine-rich region. The putative factor of *S.acidocaldarius* and the homolog that is found in the eucaryal virus ASF (12) both lack the N-terminal domain and the serine-rich region and are therefore about 30% shorter.

The conserved C-terminal domain is also found in the integral subunit rpo30 of vaccinia virus RNAP, which, however, differs from the TFIIS-factors by a unique C-terminal extension (13, 14). As in eucarya, the transcription factor of *S.acidocaldarius*

is not a component of purified RNA polymerase, suggesting that as in eucarya it is only present in the transcription complex.

One remarkable feature in which the *S.acidocaldarius* factor differs from its homologs is a second presumable zinc binding domain (Cys-X₂-Cys-X₁₆-Cys-X₂-Cys), 44 amino acids upstream of the conserved zinc finger motif present in all members of the TFIIS family.

The *S.acidocaldarius* *tflls* gene is located directly downstream of the gene for RNAP subunit L, even with an overlap of four nucleotides. It lacks a promoter sequence of its own and is thus probably cotranscribed with the gene for the RNAP subunit L. This localization is a further indication for the factor being a true component of the transcription machinery.

REFERENCES

- Pühler, G., et al. (1989) *Nucleic Acids Res.* **17**, 4517–4534.
- Klenk, H.-P., et al. (1992) *Nucleic Acids Res.* **20**, 5226.
- Langer, D., et al. (1993) in preparation.
- Reiter, W.-D., et al. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 9509–9513.
- Ouzounis, C. and Sander, C. (1992) *Cell* **71**, 189–190.
- Hubert, J.-C., et al. (1983) *EMBO J.* **2**, 2071–2073.
- Hirashima, S., et al. (1988) *J. Biol. Chem.* **263**, 3858–3862.
- Marshall, T.K., et al. (1990) *Nucleic Acids Res.* **18**, 6393–6398.
- Horikoshi, M., et al. (1982) *J. Biol. Chem.* **257**, 2719–2721.
- Reinberg, D. and Roeder, R.G. (1987) *J. Biol. Chem.* **262**, 3331–3337.
- Agarwal, K., et al. (1991), *EMBO J.* (in press).
- Rodriguez, J.M., et al. (1992) *Virology* **186**, 40–52.
- Ahn, B.-Y., et al. (1990) *Mol. Cell. Biol.* **10**, 5433–5441.
- Broyles, S. and Pennington, M. (1990) *J. Virology* **64**, 5376–5382.
- Yoo, O.J., et al. (1991) *Nucleic Acids Res.* **19**, 1073–1079.

```

S.a. 143 KTLVILASDMP NPGVQLTRGI SCPSCGNDEA YFWILOTRSA DEPATRFVYC TKGCKVWREY.E* 204
D.m. 253 KEAINDAQLA TVGGTKTDLI KCAKCKKRNK TYNLOTRSA DEPMTFVVC NECGNRWRFK* 313
H.s. 220 KEAIREHQMA KTGQTQDILF TCGKCKKRNK TYTQVOTRSA DEPMTFVVC NECGNRWRFK* 280
S.c. 249 KQNLVNAQGA TIERSTDRF TCGKCKEKKV SYVLOTRSA DEPLTFCTC EACGNRWRFK* 309
ASF 1243L 184 KTEITLRSQQ KVAEKTSQLY KCPNCKQRMK TYREVOTRAL DEPSTIFCTC KKGHEFIG* 243
Vac Rpo30 137 KIKDVSYNDY FHWLDEKYNF PCFNCKSRNT FEMMIQTRAA DEPLLVRHAC RDCQKHFKEP K 197

```

Figure 1. Sequence alignment of the most conserved region of *tflls*-like genes of different species; S.a.: *S.acidocaldarius*, D.m.: *D.melanogaster* (8), H.s.: *H.sapiens* (15), S.c.: *S.cerevisiae* (6) and the homologous part of ASF 1243L: ORF 1243: of *African swine virus* (12) and the Vac Rpo30: polymerase subunit of *vaccinia virus* (13, 14). The asterisks above the sequences indicate the putative zinc binding domains.

* To whom correspondence should be addressed