Proc. Natl. Acad. Sci. USA
Vol. 92, pp. 5768–5772, June 1995
Evolution

Transcription in Archaea: Similarity to that in Eucarya

DORIS LANGER*, JOHANNES HAIN*, PIERRE THURIAUX†, AND WOLFRAM ZILLIG*

*Max-Planck-Institut für Biochemie, D-82152 Martinsried, Germany; and †Service de Biochimie et Genetique Moleculaire, Commissariat à l'Energie Atomique, Centre d'Etudes de Saclay, F-91191 cedex Gif sur Yvette, France

Communicated by Carl R. Woese, University of Illinois, Urbana, IL, February 14, 1995

ABSTRACT **We present homologies between archaeal and eucaryal DNA-dependent RNA polymerase (RNAP) subunits and transcription factors. The sequences of the *Sulfolobus acidocaldarius* subunits D, E, and N and alignments with eucaryal homologs are presented here. The similarities between archaeal transcription factors and their eucaryal homologs TFIIB and TBP have been established in other laboratories. The archaeal RNAP subunits H, K, and N, respectively, show high sequence similarity to ABC27, ABC23, and ABC10β (found in all three eucaryal RNAPs); subunit D, to AC40 (common to polymerase II and polymerase III) and B44 (polymerase II); and subunit L, to AC19 and B12.5. The similarity of subunit D and its eucaryal homologs to bacterial α is limited to the "α-motif," which is also present in subunit L and its eucaryal homologs. Genes encoding homologs of the related eucaryal RNAP subunits A12.2/B12.6 and also homologs of eucaryal transcription elongation factors of the TFIIS family have been detected in *Sulfolobus acidocaldarius* and *Thermococcus celer*. In archaea, the protein is not an RNAP subunit. Together with the sequence similarities between archaeal box A-containing and eucaryal TATA box-containing promoters, this shows that the archaeal and eucaryal transcription systems are truly homologous and that they differ structurally and functionally from the bacterial transcription machinery. In contrast, however, a number of genes for the archaeal transcription apparatus are organized in clusters resembling the clusters of transcription-associated genes in Bacteria.**

It is now accepted that the living world is divided into three domains: Bacteria, Archaea, and Eucarya (1, 2). There are a sufficient number of molecular features specifically shared between Archaea and Eucarya to suggest a common ancestry, apart from the Bacteria. This is most clearly documented in their transcription systems. Although the large components of the RNA polymerases (RNAPs) are homologous among all domains, a much higher similarity exists between the archaeal and eucaryal versions than between either of these and the (eu)bacterial version (3, 4). The canonical archaeal transcription promoter closely resembles the eucaryal TATA-box-containing [RNA] polymerase (pol) II promoters (5). Sequences of all but one subunit of the RNAP of the extremely thermophilic archaeon *Sulfolobus acidocaldarius* have now been determined, thus allowing a comprehensive comparison of RNAPs among the three domains.

In the Bacteria, transcription involves a single RNAP with only four different basic subunits: β, β′, α, and σ. In cyanobacteria and chloroplasts, the β′ component is replaced by two fragments of about equal size (6, 7). In certain species, additional components have been reported, some of which, at least, effect specific initiation of transcription (8–10).

In contrast, the nuclei of Eucarya harbor three specialized RNAPs—pol I (or A), pol II (or B), and pol III (or C)—which have been well characterized in *Saccharomyces cerevisiae* (11). The two largest subunits in each case are homologous to the bacterial components β and β′, and there is some structural and functional resemblance between the bacterial α subunit and the eucaryal AC40 and B44 subunits (12–14).

Limited similarity also has been claimed between the bacterial α subunit and the eucaryal AC19 and B12.5 subunits (14–16). Claimed homologies between the bacterial transcription initiation factor σ and certain eucaryal RNAP subunits (17) or transcription factors (18) are tenuous. No bacterial homologs of the five small subunits shared by the three eucaryal RNAPs—i.e., ABC27, ABC23, ABC14.5, ABC10α, and ABC10β—and their archaeal counterparts and of any of the unique components of yeast pol I, pol II, or pol III have been identified. Eucarya utilize various factors, either shared among all three polymerases or specific for one or two of them, or specific for particular promoters. Whereas in Bacteria the RNAP finds different promoters by means of specifying RNAP-bound σ factors, in Eucarya the specifying factors are bound to the corresponding promoters (19).

This overview reports the results of sequence comparison between all but one of the RNAP subunits from *Sulfolobus acidocaldarius*, including the sequences of subunits D, E, and N‡ and their homologs in *S. cerevisiae*. In conclusion we show that, despite the Archaea being prokaryotic in cell type, their transcription system resembles that of Eucarya and is thus different from that of Bacteria.

Homology of Eucaryal and Archaeal Small RNAP Subunits

Like their eucaryal counterparts, the archaeal RNAPs show a high complexity. The RNAP of *Sulfolobus acidocaldarius* comprises 13 different single-copy subunits. A semiquantitative immunoblotting approach (3, 4) showed that the three (or four; see below) largest archaeal subunits are homologs of the two largest eucaryal subunits and, therefore, are also related (in a more distant way) to the β and β′ subunits in Bacteria (3, 4). These conclusions have been substantiated by cloning and sequencing the corresponding gene clusters, completely in the case of six archaea and partially for another two (20, 21).

The largest eucaryal RNAP subunit (and the bacterial β′ subunit) are replaced by two subunits in Archaea, A′, homologous to roughly the N-terminal two-thirds, and A″, homologous to roughly the C-terminal third of the eucaryal and bacterial subunits. In the methanogens and extreme halophiles the homolog of the second largest eucaryal subunit and the bacterial β subunit is also replaced by two subunits, B″, corresponding to roughly the N-terminal half, and B′, corresponding to roughly the C-terminal half of the eucaryal and the bacterial versions. Sequence similarities are highest between the large archaeal subunits and their corresponding pol II and pol III homologs, somewhat lower between the homologous pol II and pol III subunits themselves, lower still between the archaeal subunits and their pol I homologs, and lowest of all

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: RNAP, RNA polymerase; pol, [RNA] polymerase; ORF, open reading frame.

‡The sequences reported in this paper have been deposited in the GenBank data base (X80194 for subunits D and N and X75411 for subunit E).

between the archaeal or eucaryal subunits on the one hand and their bacterial counterparts on the other.

In addition to the subunits B, A′, and A″, the RNAP of *Sulfolobus acidocaldarius* contains the following smaller components: D, E, F, G, H, I, K, L, N (15), and possibly M (see Table 1).

The genes for all of these (22–25), except I, were cloned by using oligonucleotides designed on the basis of partial amino acid sequences as hybridization probes. All but the I and M subunit genes were completely sequenced.

As shown in Table 1 and by comparison of aligned sequences (see refs. 24 and 25) the *Sulfolobus acidocaldarius* RNAP contains homologs of three of the five universal eucaryal small subunits—namely, H, a homolog of yeast RNAP ABC27 (24); K, a homolog of ABC23 (25); and N, which corresponds to ABC10β (Fig. 1). Krömer and Arndt (27) have sequenced an operon from the archaeon *Haloarcula marismortui* containing genes for the ribosomal proteins L13 and S9. We have identified two adjacent open reading frames (ORFs) downstream of these genes as homologous to the genes of *Sulfolobus* RNAP subunits N and K (in that order). In *Sulfolobus acidocaldarius*, only the gene encoding subunit N is part of a gene cluster, which links modified equivalents of the "α operon" and the "S9 operon" of *H. marismortui*.

Subunit L of *Sulfolobus acidocaldarius* is the homolog of the eucaryal subunits AC19 and B12.5 (30) (Table 1, Fig. 1). All of these share part of the so-called "α motif" highly conserved in the α subunits of (eu)bacterial RNAPs (25).

Subunit D of *Sulfolobus acidocaldarius* corresponds to the eukaryotic AC40 and B44 subunits (see Fig. 2) and like the latter shows limited homology to the (eu)bacterial α subunit, mainly in a well-conserved α motif. Scholzen and Arndt (32) have reported an ORF in a gene cluster of *H. marismortui* that shows high similarity to *Sulfolobus* D and the eucaryal AC40/B44 subunits. The *H. marismortui* version is contained in a gene cluster resembling the α operon of *E. coli* (32). During the assembly of pol I and pol III the eucaryal subunits AC40 and AC19 appear to be associated in a "heterodimeric" structure (14). The homologous *Sulfolobus* subunits D and L form a stable complex with each other (25).

Sulfolobus subunit E (Fig. 1) shows homology to the *S. cerevisiae* pol II subunit B16 (33) and to the product of the ORF YKL144c of chromosome XI in *S. cerevisiae*, which appears to be the 25-kDa subunit of pol III (29, 34) (Table 1). No equivalent is known in pol I. The S1 motif present in subunit E (22), is responsible for RNA binding in several contexts in both Bacteria and Eucarya and is thought to be involved in resolving or rearranging helical structures in nucleic acids. Subunit E might act in this manner during transcription elongation (22).

Six smaller subunits of *Sulfolobus acidocaldarius* RNAP thus have homologs in eucaryal RNAPs. In three cases the component is common to all three eucaryal RNAPs, in two only to pol I and pol III (A and C), with a homolog in pol II (B). In the remaining case a homolog exists in pol II and presumably also pol III (Table 1). The sequence similarities of all these subunits from *S. cerevisiae* and *E. coli* are shown in Table 2.

The archaeal RNAP components M and I remain candidates for homology with further eucaryal subunits. The sequence of I is unknown. In SDS/PAGE component M appears substoichiometric. Its N-terminal amino acid sequence is the same as in subunit K, suggesting it to be a fragment of K rather than a distinct subunit.

No eucaryal homologs have been found for the archaeal RNAP components G and F. Component F is possibly not a true subunit, because the polymerase devoid of F is still able to correctly initiate transcription *in vitro* (25). Furthermore, F is not bound as tightly as the other components to the RNAP complex and is strikingly polar, with an excess of negative charge, which could account for its binding to the RNAP.

Homology of Archaeal and Eucaryal Transcription Factors

The canonical archaeal transcription promoter contains a "box A" sequence, TTTAWA (5), reminiscent of the TATA box of certain eucaryal pol II promoters. The center of box A is located 27 ± 4 nucleotides from the start of transcription on a YR sequence (35). This distance is similar to that in many pol II promoters. Purified *Sulfolobus* RNAP is able to start transcription at the "normal" start site even when insertions or deletions have been introduced between box A and this site, or when box A has been completely deleted. However, efficient initiation of transcription *in vitro* requires both a promoter containing the box A sequence and a protein fraction (25, 36). In this case, however, transcription initiation occurs on the R of a YR sequence that is 27 ± 4 nucleotides downstream from the center of box A rather than at the "normal" start site and thus appears directed by box A. Using RNAP from *Methanococcus* sp., Frey *et al.* (37) and Hausner and Thomm (38) have been able to replace the undefined protein fraction by two purified proteins, called aTFA and aTFB.

These results suggest that the archaeal RNAP has a dual interaction with the promoter, first directly at an unidentified signal including the YR start, and second mediated by transcription factors at box A. In Eucarya, several proteins are involved in the initiation of transcription, but two of them, TBP

Table 1. Subunits of the DNA-dependent RNAP of *Sulfolobus acidocaldarius*, their lengths, their molecular masses, and their homologs in yeast and *Escherichia coli*

RNAP subunit of *Sulfolobus acidocaldarius*	Mass, kDa	Length, amino acid residues	Homologous RNAP subunits: Bacterial (*E. coli*)	Homologous RNAP subunits: Eucaryal (*S. cerevisiae*)	Ref.	Accession no. for *Sulfolobus*
B	122	1126	β	A135, B150, C128	26	X14818
A′	101	880	β′	A190, B220, C160 (N-terminal 2/3)	26	X14818
A″	44	392	β′	A190, B220, C160 (C-terminal 1/3)	26	X14818
D	30	264	α	AC40, B44		X80194
E	27	248		B16, C25		X75411
F	12	105				
G	13.8	121			23	
H	11.8	88		ABC27 (C-terminal 1/3)	24	X14818
I	9.7*	83*				
K	9.7	83		ABC23 (C-terminal 1/2)	25	X80753
L	10	90		AC19, B12.5	25	X70805
M	5.5*	50*				
N	7.5	66		ABC10β	25	X80194

*Data derived only from mobility in SDS/PAGE.

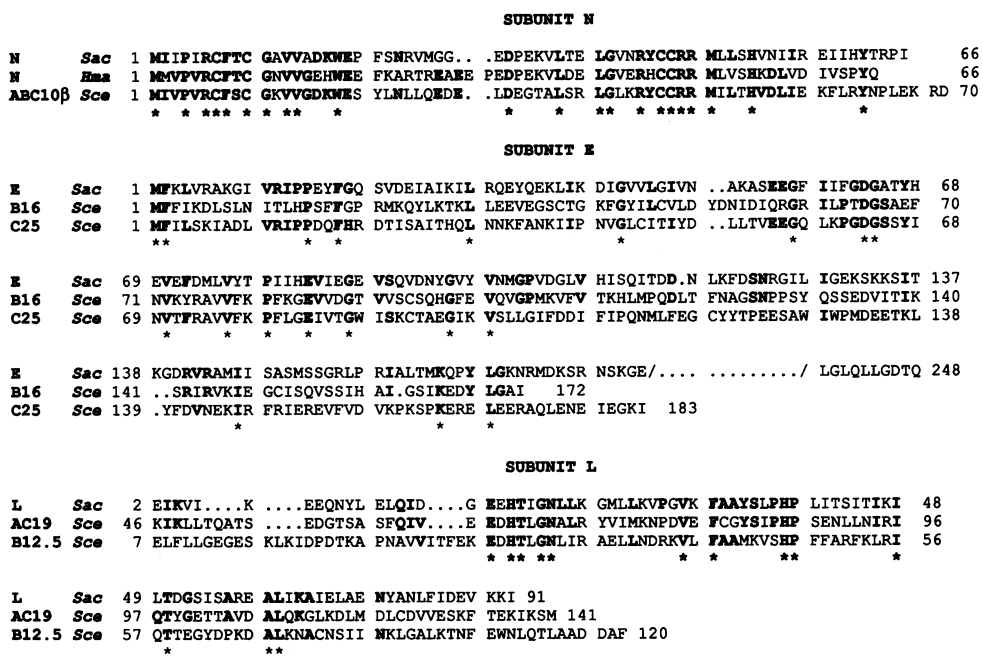

FIG. 1. Sequence alignments of subunits N, E, and L. Sequence alignments are shown for archaeal subunit N of the *Sulfolobus acidocaldarius* (*Sac*) and *H. marismortui* (*Hma*) (27) RNAPs with subunits ABC10β of the RNAP of *S. cerevisiae* (*Sce*) (14). In boldface letters are symbols for amino acids that are conserved in at least one archaeal and one eucaryal subunit. Next are sequence alignments of archaeal subunit E of the *Sulfolobus acidocaldarius* (*Sac*) with subunits B16 (25, 28) and C25 (29) of the pol II of *S. cerevisiae* (*Sce*). At the bottom are sequence alignments of archaeal subunit L with subunits AC19 and B12.5 of the RNAP of *S. cerevisiae* (*Sce*) (15, 16, 25).

and TFIIB, appear to be sufficient to effect specific initiation by pol II in a minimal system on a supercoiled template (39). TBP, a part of the TFIID complex, is common to all three eucaryal RNAPs (40), while TFIIB mediates the binding of the RNAP to the promoter-bound TFIID in the case of pol II.

ORFs with homologies to eucaryal TBP genes have recently been cloned from the two archaea *Thermococcus celer* (41) and *Pyrococcus woesei* (42). An ORF in *P. woesei* encoding a putative homolog of TFIIB was cloned and sequenced by Creti *et al.* (43), and identified by Ouzounis and Sander (44).

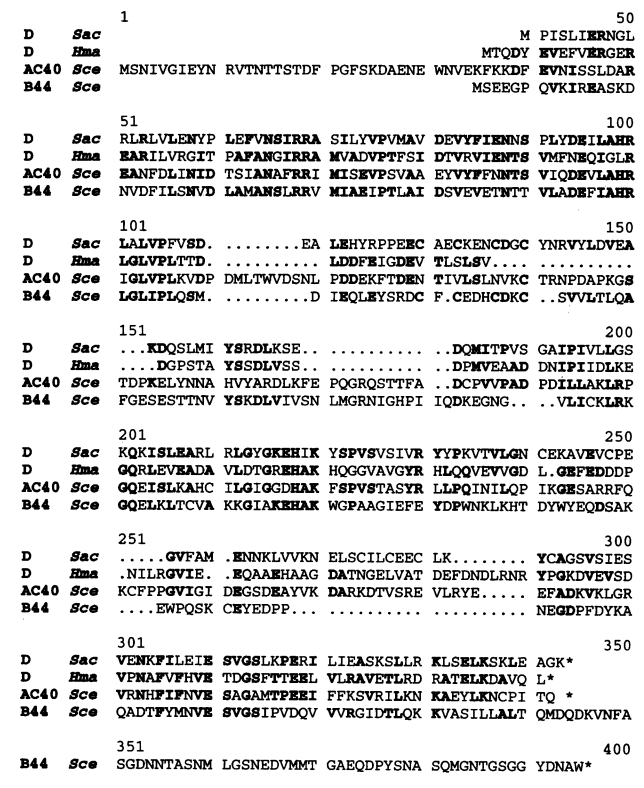

FIG. 2. Sequence alignment of archaeal subunits D of the *Sulfolobus acidocaldarius* (*Sac*) and *H. marismortui* (*Hma*) RNAP with the subunits AC40 (31) and B44 (13) of the *S. cerevisiae* (*Sce*) RNAP. In boldface letters are symbols for amino acids that are conserved in at least one archaeal and one eucaryal subunit.

Both Bacteria and Eucarya possess factors that resolve "jams" in transcription elongation: the greA and greB proteins in *E. coli* (45) and the unrelated TFIIS factor, associated with eucaryal pol II (46). In *Sulfolobus acidocaldarius*, downstream of the gene for RNAP subunit L we have located an ORF encoding a protein with a C-terminal domain which shows high sequence similarity to that of TFIIS factors (47). A similar ORF has also been reported in *T. celer* (48). The archaeal proteins would be roughly 100 amino acids long, although their homology with TFIIS is confined to about 40 C-terminal amino acids including a zinc finger (47, 48). Two yeast RNAP subunits, A12.2 and B12.6 (49, 50), however, each share two zinc finger domains with the putative archaeal proteins and appear more similar to the latter than the TFIIS factors. But in contrast to the situation in Eucarya, in *Sulfolobus* the putative protein is not an RNAP subunit.

In conclusion, archaea utilize two transcription initiation factors and a possible transcription elongation factor which are clearly homologous to eucaryal rather than bacterial proteins. The eucaryal homologs of these archaeal factors constitute a minimal set just sufficient for specific transcription from certain promoters (39). No factors involved in promoter selection and thus transcription specificity have so far been identified in archaea.

Gene Organization

The genes for the largest RNAP subunits of archaea (Fig. 3*A*) are organized in a cluster resembling the *rpoBC* operon in *E. coli*. The archaeal cluster contains, however, in addition the gene for subunit H, having a homolog in eucarya but not in bacteria, immediately upstream of *rpoB* but downstream of the common promoter. There are several other examples (especially involving ribosomal protein genes) where archaeal gene clusters resemble their (eu)bacterial counterparts in gene order but harbor in addition related genes found only in archaea and eucarya—e.g., the above-cited operonal organization of *rpoN* and *rpoK*. The existence of such gene clusters appears to be an ancestral feature which has been lost in the eucaryal lineage.

The large RNAP subunit gene clusters in Archaea and Bacteria are in quite different genomic context. In (eu)bacte-

Table 2. Quantification of sequence similarities of *Sulfolobus acidocaldarius* to corresponding RNAP subunits of *S. cerevisiae* and *E. coli* RNAPs

Sulfolobus acidocaldarius RNAP subunit	Identity/similarity to RNAP subunits							
	S. cerevisiae						*E. coli*	
	pol I subunit	%/%	pol II subunit	%/%	pol III subunit	%/%	RNAP subunit	%/%
A′	A190	36/59	B220	43/63	C160	43/64	β′	30/51
A″	A190	32/54	B220	33/59	C160	34/59	β′	25/47
B	A135	30/56	B150	44/65	C128	38/61	β	25/49
D	AC40	34/57	B44	25/49	AC40	34/57	α	22/46
E		—	B16	23/46	C25	23/44		—
H	ABC27	40/55	ABC27	40/55	ABC27	40/55		—
K	ABC23	38/61	ABC23	38/61	ABC23	38/61		—
L	AC19	35/62	B12.5	26/52	AC19	35/62		—
N	ABC10β	50/73	ABC10β	50/73	ABC10β	50/73		—

The first number for each subunit is the percent identity with the *Sulfolobus* subunit and the second is the percent similarity according to the Dayhoff matrix.

ria, genes for ribosomal proteins L11, L1, L10, and L12 are situated immediately upstream of *rpoB*. Archaea possess the same gene cluster, though unlinked to the large component genes. In *E. coli* the "streptomycin-operon" is situated in a different map position than the corresponding gene cluster in archaea, which closely follows the large RNAP subunit gene cluster (20). In the archaeon *H. marismortui*, the α operon (32) and the S9 operon (27) are immediately adjacent to each other; in the (eu)bacterium *E. coli*, they are separated.

We found a gene cluster in *Sulfolobus acidocaldarius* containing the genes for RNAP subunits D and N, and for six ribosomal proteins and an ORF for a putative nucleic acid-binding protein, which resembles in its composition and largely in its gene order the adjacent α (32) and S9 operon (27) of *H. marismortui* (Fig. 3*B*). This and the neighborhood of *rpoL* and the gene encoding the A12.2/B12.6 homolog in *Sulfolobus acidocaldarius* (Fig. 3*C*) show that elements of the transcription machinery of archaea are also organized in clusters.

Evolutionary Implications

The strong sequence divergences between the subunits in the bacterial transcription system on the one hand, and their archaeal/eucaryal homologs on the other, plus the fact that many components common to the archaeal and eucaryal systems do not find homologs in the bacterial system, suggest functional differences in transcription.

Bacterial promoters are thought to be freely accessible to RNAP, and specialized σ subunits may thus suffice for promoter recognition and transcriptional initiation. In contrast, the nucleosomal structure of the eucaryal chromosome hinders the recognition of transcription promoters, which seems then to require a local "resolving" of the chromatin structure by factors including TBP and TFIIB. The anchoring of these factors to the DNA presumably allows successive rounds of transcription to occur from preinitiation complexes, as opposed to the complete recycling of polymerase and σ factors

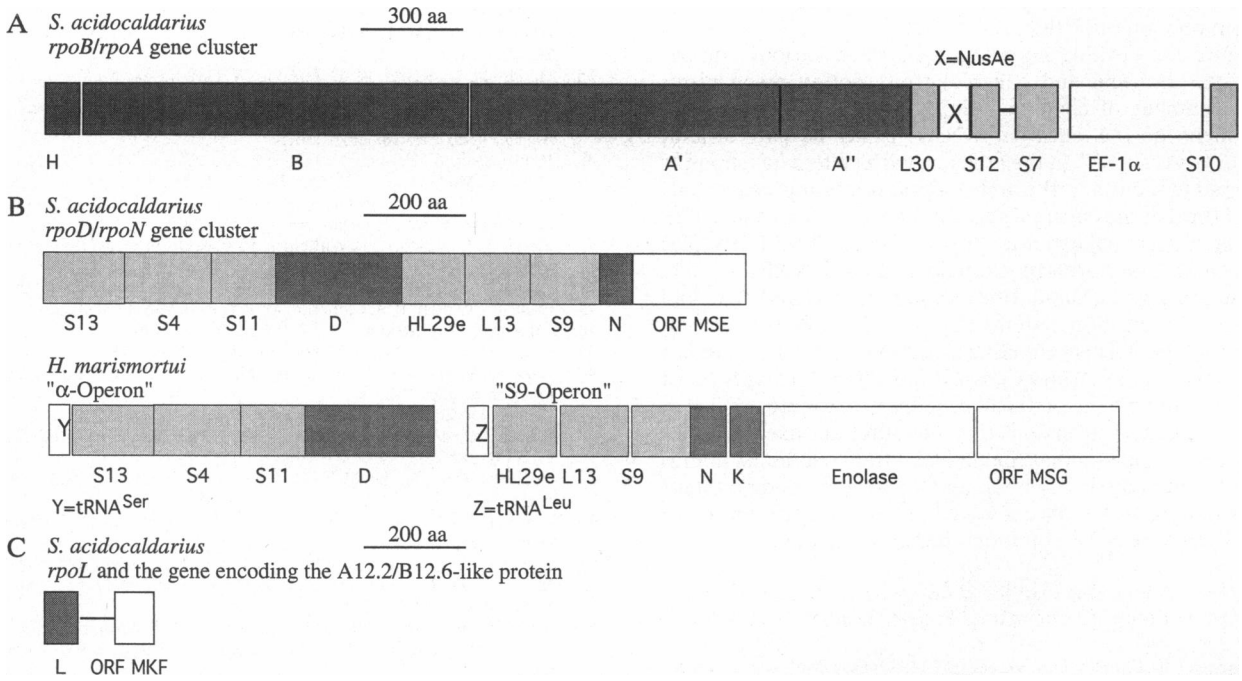

FIG. 3. Gene clusters containing RNAP subunit genes from archaea (dark gray) and in three of the example genes for ribosomal proteins (light gray). The names of the gene products are written below. (*A*) *rpoB/rpoA′/rpoA″* gene cluster from *Sulfolobus acidocaldarius*. NusAE, homolog of the antitermination factor NusA. (*B*) *rpoD/rpoN* gene cluster from *Sulfolobus acidocaldarius* and the two corresponding operons from *H. marismortui*. ORFs MSE (D.L., unpublished work) and MSG (see ref. 27) are discussed elsewhere. (*C*) *rpoL* and the adjacent ORF MKF, which encodes the A12.2/B12.6-like protein.

characterizing bacterial transcription (51). The existence of archaeal homologs for TBP and TFIIB indicates that eucaryote-like preinitiation complexes also function in archaea, suggesting that the archaeal genome has chromatin-like structure as well (52).

A key evolutionary question is whether bacterial RNAPs are simpler in structure than archaeal and eucaryal RNAPs because they have lost a number of subunits present in the universal ancestor or because they have never evolved them.

The fact that the Archaea and Bacteria both contain similar gene clusters implies the presence of such clusters in the common ancestor, whereas the domain-specific differences of the arrangement of such clusters in bacteria and archaea indicate that their linkage into coherent genomes occurred twice independently, in the bacteria and in the archaea (53). This is in line with the proposal of Woese that the ancestral progenote preceding the common ancestor of the three domains of life did not have a coherent genome yet (53, 54). The archaeal gene clusters contain genes absent from their bacterial counterparts—e.g., for RNAP subunits that have homologs in eucarya but not in bacteria. It seems more probable that this resulted by loss of such genes from the ancestral gene cluster in the bacterial lineage, rather than from their introduction into "suitable" clusters in archaea, for which mechanisms of choice are hard to imagine. This argument is supported by the finding that in archaea such clusters have been found to be interrupted by promoters and terminators and thus do not appear to originally have constituted regulatory units. They might rather have constituted packages to ensure the joint transfer of genes encoding cooperating gene products between individuals in ancestral populations in a stage of evolution where the separation of lineages had not yet occurred (54). The genes in such clusters should then be ancestral rather than invented after the separation of bacteria on one hand and archaea and eucarya on the other. Another argument for this assumption is furnished by the notion that in phylogenetic trees of certain molecules—e.g., DNA-dependent RNAPs—the bacterial lineage appears to be longer than the archaeal, indicating the bacteria to have diverged further from the common ancestor than the archaea.

Despite the striking homologies between various components, the archaeal and eucaryal transcription mechanisms show a number of characteristic differences: e.g., the large component of the eucaryal RNAP exists as two disjoint fragments, A′ and A″, in the archaea; each system has subunits not found in the other; the archaeal subunit H represents only the last third of the eucaryal subunit ABC27, the archaeal TBP lacks the N-terminal extension found in eukaryotic TBPs (39); and eucarya have three functionally distinct RNAPs.

In summary, the similarities shown by the archaeal and eucaryal transcription systems strongly reinforce the idea (1, 55, 56) that the Archaea and Eucarya are relatives. In a number of cases where the eucarya possess more than one version of an RNAP subunit the archaeal sequences are more similar to each of the eucaryal versions than the latter are to each other. It, therefore, appears that the archaeal transcription system is not only simpler (in number and/or size of components/factors required) than its eucaryal homolog but also has more faithfully retained the common ancestral function.

We thank Andre Sentenac for many discussions. This work was supported by European Community Project Grant SC1-CT91-0702.

1. Woese, C. R., Kandler, O. & Wheelis, M. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 4576–4579.
2. Woese, C. R. (1994) *Microbiol. Rev.* **58,** 1–9.
3. Huet, J., Schnabel, R., Sentenac, A. & Zillig, W. (1983) *EMBO J.* **2,** 1291–1294.
4. Schnabel, R., Thomm, M., Gerardy-Schahn, R., Zillig, W., Stetter, K. O. & Huet, J. (1983) *EMBO J.* **2,** 751–755.
5. Reiter, W.-D., Hüdepohl, U. & Zillig, W. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 9509–9513.
6. Schneider, G. J. & Haselkorn, R. (1988) *J. Bacteriol.* **170,** 4136–4140.
7. Ozeki, H., Ohyama, K., Inokuchi, H., Fukuzawa, H., Kohchi, T., Sano, T., Nakahigashi, K. & Umesono, K. (1987) *Cold Spring Harbor Symp. Quant. Biol.* **52,** 791–804.
8. Stetter, K. O. & Zillig, W. (1974) *Eur. J. Biochem.* **48,** 527–540.
9. Gierl, A., Zillig, W. & Stetter, K. O. (1982) *Eur. J. Biochem.* **125,** 41–47.
10. Pero, J., Nelson, J. & Fox, T. D. (1975) *Proc. Natl. Acad. Sci. USA* **72,** 1589–1593.
11. Thuriaux, P. & Sentenac, A. (1992) in *The Molecular and Cellular Biology of the Yeast Saccharomyces*, eds. Jones, E. W., Pringle, J. R. & Broach, J. R. (Cold Spring Harbor Lab. Press, Plainview, NY), Vol. 2, 1–45.
12. Martindale, D. W. (1990) *Nucleic Acids Res.* **18,** 2953–2959.
13. Kolodzej, P. & Young, R. A. (1989) *Mol. Cell. Biol.* **9,** 5387–5394.
14. Lalo, D., Carles, C., Sentenac, A. & Thuriaux, P. (1992) *Proc. Natl. Acad. Sci. USA* **90,** 5524–5528.
15. Woychik, N. A., McKune, K., Lane, W. S. & Young, R. A. (1993) *Gene Exp.* **3,** 77–82.
16. Dequard-Chablat, M., Riva, M., Carles, C. & Sentenac, A. (1991) *J. Biol. Chem.* **266,** 15300–15307.
17. Woychik, N. A. & Young, R. A. (1989) *Mol. Cell. Biol.* **9,** 2854–2859.
18. Sopta, M., Burton, Z. F. & Greenblatt, J. (1989) *Nature (London)* **341,** 410–415.
19. Burgess, R. R. (1971) *Annu. Rev. Biochem.* **40,** 711–738.
20. Zillig, W., Palm, P., Klenk, H.-P., Langer, D., Hüdepohl, U., Hain, J., Lanzendörfer, M. & Holz, I. (1993) in *The Biochemistry of Archaea (Archaebacteria)*, eds. Kates, M., Kushner, D. J. & Matheson, A. T. (Elsevier, New York), pp. 367–391.
21. Klenk, H.-P., Palm, P. & Zillig, W. (1993) *Syst. Appl. Microbiol.* **16,** 138–147.
22. Langer, D., Lottspeich, F. & Zillig, W. (1994) *Nucleic Acids Res.* **22,** 694.
23. Langer, D. (1991) Thesis (Ludwig-Maximilians-Universität, Munich).
24. Klenk, H.-P., Palm, P., Lottspeich, F. & Zillig, W. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 407–410.
25. Lanzendörfer, M., Langer, D., Hain, J., Klenk, H.-P., Holz, I., Arnold-Ammer, I. & Zillig, W. (1994) *Syst. Appl. Microbiol.* **16,** 156–164.
26. Pühler, G., Lottspeich, F. & Zillig, W. (1989) *Nucleic Acids Res.* **17,** 4517–4534.
27. Krömer, W. J. & Arndt, E. (1991) *J. Biol. Chem.* **266,** 24573–24579.
28. McKune, K., Richards, K. L., Edwards, A. M., Young, R. A. & Woychik, N. A. (1993) *Yeast* **9,** 295–299.
29. Sadhale, P. P. & Woychik, N. A. (1994) *Mol. Cell. Biol.* **14,** 6164–6170.
30. Woychik, N. A., McKune, K., Lane, W. S. & Young, R. A. (1993) *Gene Exp.* **3,** 77–82.
31. Mann, C., Buhler, J.-M., Treich, I. & Sentenac, A. (1987) *Cell* **48,** 627–637.
32. Scholzen, T. & Arndt, E. (1992) *J. Biol. Chem.* **267,** 12123–12130.
33. McKune, K., Richards, K. L., Edwards, A. M., Young, R. A. & Woychik, N. A. (1993) *Yeast* **9,** 295–299.
34. Huet, J., Riva, M., Sentenac, A. & Fromageot, P. (1985) *J. Biol. Chem.* **260,** 15304–15310.
35. Hain, J., Reiter, W.-D., Hüdepohl, U. & Zillig, W. (1992) *Nucleic Acids Res.* **20,** 5423–5428.
36. Hüdepohl, U., Reiter, W.-D. & Zillig, W. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 5851–5855.
37. Frey, G., Thomm, M., Brüdigam, B., Gohl, H. P. & Hausner, W. (1990) *Nucleic Acids Res.* **18,** 1361–1367.
38. Hausner, W. & Thomm, M. (1993) *J. Biol. Chem.* **268,** 24047–24052.
39. Tyree, C. M., George, C. P., Lira-De Vito, L. M., Wampler, S. L., Dahmus, M. E., Zawel, L. & Kadonaga, J. T. (1993) *Genes Dev.* **7,** 1254–1265.
40. Hernandez, N. (1993) *Genes Dev.* **7,** 1291–1308.
41. Marsh, T. L., Reich, C. I., Whitelock, R. B. & Olsen, G. J. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 4180–4184.
42. Rowlands, T., Baumann, P. & Jackson, S. P. (1994) *Science* **264,** 1326–1329.
43. Creti, R., Londei, P. & Cammarano, P. (1993) *Nucleic Acids Res.* **21,** 2942.
44. Ouzounis, C. & Sander, C. (1992) *Cell* **71,** 189–190.
45. Sparkowski, J. & Das, A. (1990) *Nucleic Acids Res.* **18,** 6443.
46. Agarwal, K., Baek, K., Jeon, C., Miyamoto, K., Ueno, A. & Yoon, H. (1991) *Biochemistry* **30,** 7842.
47. Langer, D. & Zillig, W. (1993) *Nucleic Acids Res.* **21,** 2251.
48. Kaine, B. P., Mehr, I. J. & Woese, C. R. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 3854–3856.
49. Nogi, Y., Yano, R., Dodd, J., Carles, C. & Nomura, M. (1993) *Mol. Cell. Biol.* **13,** 114–122.
50. Woychik, N. A., Lane, W. S. & Young, R. A. (1991) *J. Biol. Chem.* **266,** 19053–19055.
51. Travers, A. A. & Burgess, R. R. (1969) *Nature (London)* **222,** 537–540.
52. Musgrave, D. R., Sandmann, K. M. & Reeve, J. N. (1992) *Proc. Natl. Acad. Sci. USA* **88,** 10397–10401.
53. Zillig, W., Palm, P. & Klenk, H.-P. (1992) in *The Origin and Evolution of the Cell*, eds. Hartmann, H. & Matsuno, K. (World Scientific, Singapore), pp. 163–182.
54. Woese, C. R. (1987) *Microb. Rev.* **51,** 221–271.
55. Gogarten, J. P., Kibak, H., Dittrich, P., Taiz, P., Bowman, E. J., Manolson, M. F., Poole, P. J., Date, T., Oshima, T., Konishi, J., Denda, K. & Yoshida, M. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 9355–9359.
56. Iwabe, N., Kuma, K., Hasegawa, M., Osawa, S. & Miyata, T. (1989) *J. Mol. Evol.* **32,** 70–78.