THE EVOLUTION OF THE TRANSCRIPTION APPARATUS

W. Zillig, R. Schnabel, F. Gropp and W. D. Reiter

Max-Planck-Institut für Biochemie
Martinsried, Federal Republic of Germany

K. Stetter and M. Thomm

Institut für Biochemie, Genetik und Mikrobiologie
Universität Regensburg
Regensburg, Federal Republic of Germany

1. INTRODUCTION

The division of the prokaryotic domain into the urkingdoms of the Eubacteria and the Archaebacteria by Carl Woese and his collaborators (Woese and Fox, 1977; Woese et al., 1978) and in consequence most of our present knowledge of the natural systems of the two novel urkingdoms (Fox et al., 1980) rest on the comparison of sequences of T1 RNAase oligonucleotide catalogs of 16S rRNAs of a large number of species (Fox et al., 1977). An increasing body of additional evidence including comparative analyses of other parts of the translation apparatus (Hori et al., 1982; Fox et al., 1982; Matheson and Yaguchi, 1982; Yaguchi et al., 1982; Böck et al., 1984; Gupta, 1984; Kessel and Klink, 1982), of the transcription machinery (Zillig et al., 1982; Schnabel et al., 1983; Huet et al., 1983), of cell envelopes (Kandler, 1982) and membranes (review by Langworthy et al., 1982), of responses to antibiotics (review by Böck and Kandler, in the press) and of certain aspects of metabolism prove these revolutionizing ideas beyond reasonable doubt.

The error margin of the original S_{AB} method, which led to the recognition of the three entirely separate urkingdoms of life, the Eubacteria, the Archaebacteria and the "eukaryotic cytoplasm" or Eucyta (defined as the truly eukaryotic com-

EVOLUTION OF
PROKARYOTES

ISBN 0-12-625130-4

partments of eukaryotic organisms), does not allow a significant estimation of S_{AB} values and thus phylogenetic distance between the kingdoms, so that the early branching sequence in the evolution of the living world remains undetermined.

As more evidence becomes available, the problem acquires new dimensions.

(1) Comparative sequence analysis of homologous macromolecules measures the distances between the present ends of lineages, not, as would be required, between their origins i.e. the early branching points. Many more sequences have to be determined such that "ur-sequences" of lineages can be derived.

(2) Evolution might not solely proceed via the separation of lineages but, to an unknown, though possibly small extent, also via horizontal gene exchange or recombination between different lineages. This has for example certainly occurred between the urmitochondrial and the eucytic compartments in early eukaryotes. Examples of such processes between distant prokaryotes are as yet unknown.

(3) The evolution rates of different genes within a lineage are not necessarily parallel to each other. This is the reason why comparative analysis of different macromolecules has led to somewhat different phylogenetic trees. That derived from 5S rRNA sequence (Fox et al., 1982; Hori et al., 1982) is for example differing in details from that obtained by comparison of T1 RNAase oligonucleotide catalogs of 16S rRNAs. This review will give other examples.

In the following, we will show that a comparison of the component patterns (complexity, homologies) and antibiotic responses of the DNA dependent RNA polymerases of Archaebacteria, Eubacteria and eukaryotes,

(1) is independent evidence for the large phylogenetic distance between Eubacteria and Archaebacteria;
(2) testifies for the deep division between the two major branches of the urkingdom of the Archaebacteria, together with other feature designs;
(3) suggests a specific relation of Archaebacteria, especially of the sulfur dependent branch, with Eucyta which remains to be defined;
(4) suggests a specific relation of Methanobacteriales and

Methanococcales to Eubacteria, which remains to be defined; (5) exemplifies differences in the rates of evolution of different genes.

2. FACTS

2.1 Signal structures and transcription mechanisms

Except for the presence of certain "boxes", e.g. the TATAAT box plus the -35 box in eubacterial promoters, the TATA box in eukaryotic RNA polymerase II promoters, and certain downstream structures in RNA polymerase III initiation sites, the sequences interacting with RNA polymerases in transcription initiaton are ill defined. RNA polymerases appear to recognize the specificity of promoters better than the human mind or the computer. Accordingly, the few known archaebacterial sequences containing transcription starts (DasSarma et al., 1984; Klein, H., personal communication) do not allow the derivation of general sequences or structures of archaebacterial promoters. However, mapping of transcription starts has revealed that sequences acting as promoters for E. coli RNA polymerase are inactive in certain Archaebacteria and vice versa (J. Konisky, personal communication). In vitro investigations strengthen the conclusion that archaebacterial promoters differ from their eubacterial counterparts (Prangishvilli et al., 1982).

Like in many cases in Eubacteria, two different genes, encoding distinct components of methyl CoM reductase, are linked within one operon in Methanococcus voltae (Konheiser et al., 1984).

Since in vitro systems for cyclic, asymmetric, signal specific transcription by archaebacterial RNA polymerases are not yet available, details of the transcription process in Archaebacteria e.g. binding of the polymerase to the DNA, specificity, role of components of the polymerases are unknown. Not much more is known about transcription by the eukaryotic RNA polymerases, which so closely resemble those of Archaebacteria.

So far, transcription by isolated archaebacterial polymerases differs from that by eubacterial enzymes mainly in that they are, by themselves, unable to transcribe free native DNA specifically.

Thus, the present status of ignorance of transcription signal structures and mechanisms not only in Archaebacteria, but also in eukaryotes and, less so, in Eubacteria, prohibits a meaningful comparative analysis aiming at understanding the phylogeny of the transcription systems. Therefore, we concentrate here on the much more lucid situation pertaining to the comparison of structures (compositions) and component homologies of the "normal" DNA dependent RNA polymerases.

2.2 The DNA dependent RNA polymerases of the Eubacteria

are all of rather low complexity, with only four types of "true" components (not considering "binding proteins" like ω and τ), are almost all, in their wild type versions, sensitive to the antibiotics rifampicin, which specifically inhibits initiation, and streptolydigin, which specifically inhibits elongation of transcription.

They are all of the type $\beta'\beta\sigma\alpha_2$ (see Fig. 1), in which β' and β are two large peptides with molecular weights above or close to 100.000, α is present twice per enzyme monomer and σ is an initiation factor involved in promoter recognition and, at least in some cases, released after initiation. β' is involved in DNA binding, β appears to harbour important active sites (reviews: Zillig et al., 1976; Burgess, 1976).

Differences between different groups concern the relative sizes of β', β and σ. In gramnegative Eubacteria, β' is usually larger than β, in grampositive Eubacteria, it is the other way round. Gramnegative Eubacteria possess, in average, large σ components. In grampositive Eubacteria like Bacillus σ is smaller, but an additional component, δ in Bacillus, y in Lactobacillus, appears to be functionally homologous to a portion of the large σ of the gramnegatives (Dickel et al., 1980; Achberger and Whitely, 1981; Gierl et al., 1982), suggesting the division of the large polyfunctional σ of the gramnegatives into two peptides of different function, one, σ , concerned with promoter recognition, the other, δ or y, with the suppression of unspecific binding and/or initiaton.

In some phyla like the blue green algae, σ is a stoichiometric component (Herzfeld and Kiper, 1976), in others, like E. coli, it is a substoichiometric factor only bound to free core enzyme. In Bacillus and Lactobacillus, it

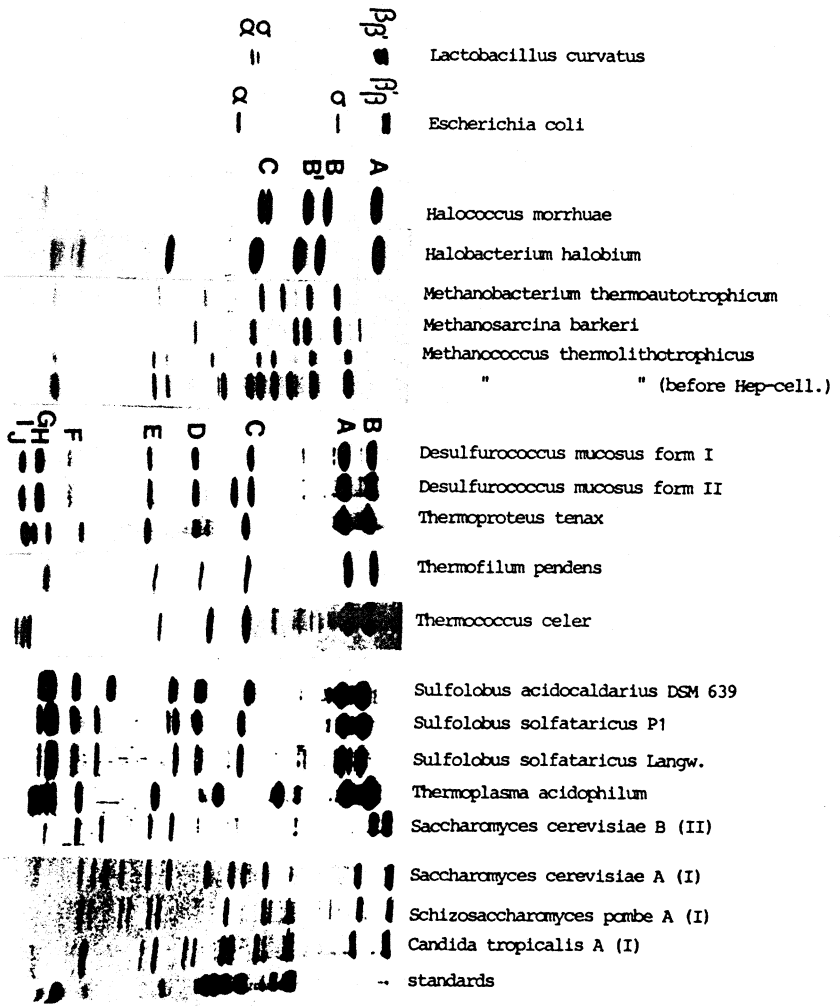

Fig. 1 SDS polyacrylamide gel electrophoresis patterns of DNA dependent RNA polymerase of eubacteria (top), methanogenic and halophic and sulfur dependent Archaebacteria (the latter including T. acidophilum) (middle) and Eucyta (bottom). Compiled and aligned from different runs.

replaces δ or y in the preinitiation complex and is released upon initiation (Gierl et al., 1982).

The homology of the components of RNA polymerases of different orders or families is evident from Western blot analysis employing antibodies against single components of one or the other polymerase, and also, in some cases, from functional tests, employing mutant components in reconstitution experiments. However, RNA polymerases of different families, e.g. grampositive and gramnegative Eubacteria, do not crossreact in the Ouchterlony in-gel-immunodiffusion test which is based on immunoprecipitation, thus allowing to determine phylogenetic distance in a qualitative manner. A discussion of the role of RNA polymerase binding proteins like ω and τ and of additional σ factors of different promoter specificities as found in B. subtilis, where they appear to be involved in sporulation, is beyond the scope of this review.

The RNA polymerases of organelles of eukaryotes, plastids and mitochondria, which should be derived from those of Eubacteria, are not discussed here because relevant evidence is completely lacking.

2.3 The DNA dependent RNA polymerases of the Eucytes i.e. of the nuclear compartment of the eukaryotes

are of much higher complexity, containing 10 or more components per enzyme monomer (see Fig. 1). A component is normally tightly and stoichiometrically bound to the enzyme particle, thus copurifying with other components and with the activity. Little is known about the requirement of certain components for enzymatic activity and even less about their specific role in the transcription process.

Even in primitive eukaryotes like yeast three types of nuclear RNA polymerases exist, differing in their functions and in their specific component patterns, but resembling each other in their general composition:

RNA polymerase I (or A) transcribes rRNA structural genes (including 5.8S rRNA) and is insensitive to the fungal poison α amanitin.

RNA polymerase II (or B) transcribes mRNA (protein) structural genes and is highly sensitive to α amanitin.

RNA polymerase III (or C) transcribes tRNA and 5S rRNA structural genes. Its sensitivity to α amanitin is intermediate.

Their component patterns each consist of two large subunits, with sizes around 100 kdaltons, and 8 to 12 smaller components (Roeder, 1976; Huet et al., 1982). However, only three of these, in Saccharomyces AB C27 (I,II,III 27), ABC 23 (I,II,III 23) and ABC 14.5 (I,II,III 14.5) are shared by all three enzymes. In addition, RNA polymerases I and III share AC 40 (I,III 40) and AC 19 (I,III 19) (Buhler et al., 1980).

2.4 The DNA dependent RNA polymerases of the Archaebacteria

are insensitive to rifampicin and α amanitin. Though forms lacking one or the other component (Sturm et al., 1980; Madon et al., 1983), in some cases inactive, and, sometimes forms with additional components (Prangishvilli et al., 1982) have been isolated, it appears that different types of enzymes as in eukaryotic nuclei do not exist.

The enzymes are of two types, one of high, the other of intermediate complexity. Those of the sulfur dependent Archaebacteria, which appear strikingly similar to yeast polymerase I, contain four heavy components, B,A,C and D, and more than six smaller components (E to J), those of the methanogens and extreme halophiles, which show a novel type of component pattern, contain five heavy components, A , B', B'', C and D, and few (about three) smaller components (see Fig. 1 and Table 1).

Their structures (and functions of which little is known) have been discussed in detail in previous reviews (e.g. Zillig et al., 1982).

2.5 Homologies of archaebacterial, eukaryotic and eubacterial DNA dependent RNA polymerases

2.5.1 The Western blotting technique

For the analysis of homologies between distinct components of corresponding multicomponent complexes of different origin, as for example the DNA dependent RNA polymerases of different organisms, the so-called Western blotting tech-

Table 1 Molecular weights (in kilodaltons) of components of archaebacterial DNA dependent RNA polymerases

Sub-unit	Methano-bacterium thermoauto-trophicum M	Methano-sarcina barkeri	Methano-coccus thermo-litho-trophicus	Methano-lobus volcani	Halo-bacterium halobium	Sulfo-lobus acido-caldarius	Sulfo-lobus solfa-taricus	Sulfo-lobus brierleyi	Thermo-plasma acido-philum	Thermo-proteus tenax	Thermo-filum pendens	Desul-furo-coccus mucosus	Thermo-coccus celer
A	119	97	109	102	148	101	101	106	108	104	102	106	105
B	-	-	-	-	-	122	121	130	135	130	132	135	125
B'	78	72	78	70	84	-	-	-	-	-	-	-	-
B''	56	66	53	73	69	-	-	-	-	-	-	-	-
C_1	42	46	49	49	49	44	41.5	42.5	56	42.5	45	44.5	44
C_2	-	-	-	42	46	-	-	-	-	-	-	(39.5)	-
D_1	33	30	25	31.5	23.5 (ε)	33	30	28	33	34	33	31.5	35
D_2	-	-	-	-	-	32	29.5	-	-	33	-	-	-
E	24.5	?	22.5	28	-	26	23	23	22	23	24	23	25
F	-	11	-	-	13	17.5	15.5	-	13.5	15.2	14	14.4	13
G	10.5	10.2	12	12	12.5	13.8	13.9	14.8	11.7	12.5	12	12.8	10.5
H	-	-	-	-	11.8	(4x) 11.8	(4x) 11.7	(4x)12.9	11.4	(2x)11.7	11.7	11.9	10.2
I	-	-	-	-	11.3	11.1	12.2	13.1	10.5	(2x)11.3	11.0	11.1	10.0
J	-	-	-	-	-	10.8	11.0	10.8	-	10.8	-	10.7	-

The designations of the smaller components, especially of the RNA polymerases of Halobacterium and of methanogenic bacteria remain tentative until homology has been proved. C_2 in RNA polymerase from D. mucosus is an additional component of one form of this enzyme. D_1 and D_2 often form a doublet of unsharp bands with characteristic tinges.

nique of probing for immunochemical crossreactivity offers a short cut to nucleotide sequence comparison (Prager et al., 1980; Huet et al., 1982). Though only semiquantitative and by far not on the same high level of lucidity as sequence analysis, it yields valuable information speedily and with ease. The method consists of challenging blotted patterns of separated components of the complex, e.g. SDS polyacrylamide gel component patterns of RNA polymerases transferred to nitrocellulose sheets, with antibodies against single components. Crossreaction is visualized e.g. employing protein A of Staphylococcus aureus labelled with ^{125}J or anti-antibody covalently coupled to peroxidase, thus allowing detection by autoradiography or by a colour reaction. Positive reaction indicates that antigenic groups are shared between the immunogen and the reacting component. The strength of the reaction depends both on the number of shared antigenic determinants and on their affinity to the antibody which might be decreased by structural changes in evolution. For the purpose of phylogenetic studies, the antibodies employed for probing should be polyclonal i.e. directed against as many as possible of the antigenic determinants of the completely denatured polypeptides. The probability of crossreaction is determined by structural correspondence depending on phylogenetic distance but also by the size of the peptide such that the method is normally not suitable to prove homologies between small peptides (below 20 kdaltons) even if phylogenetically rather closely related.

2.5.2 Homologies

We collaborated with J. Huet and A. Sentenac in probing archaebacterial, eubacterial and eukaryotic RNA polymerases, both in spot tests and by Western blotting of SDS polyacrylamide gel electrophoresis component patterns, with antibodies against single components of yeast (Saccharomyces) RNA polymerases I (A) and II (B) (Huet et al., 1983). The homologies between different archaebacterial RNA polymerases were established employing antibodies against single components of the RNA polymerases of Sulfolobus acidocaldarius (R. Schnabel) and Methanobacterium thermoautotrophicum (M. Thomm), which were also used to challenge patterns of eukaryotic and eubacterial enzymes (Schnabel et al., 1983). Finally, antibodies against single subunits of E. coli RNA polymerase were used.

The RNA polymerases were from Escherichia coli, in few instances Bacillus cereus (Rexer et al., 1975) and Lactobacillus curvatus (Stetter and Zillig, 1974), representing the Eubacteria, Methanobacterium thermoautotrophicum (Stetter et al., 1980), Methanococcus thermolithotrophicus, Methanosarcina barkeri, Methanolobus volcani (all Thomm, 1983) and Halobacterium halobium (Zillig et al., 1978; Madon and Zillig, 1983), representing the methanogenic and extremely halophilic branch of Archaebacteria, Desulfurococcus mucosus, Thermoproteus tenax (both Prangishvilli et al., 1982), Thermofilum pendens (Zillig et al., 1983a), Thermococcus celer (Zillig et al., 1983b), and several Sulfolobus species (Zillig et al., 1980) representing the sulfur dependent branch of the Archaebacteria, Thermoplasma acidophilum (Sturm et al., 1980) as an isolated genus, and Saccharomyces cerevisiae, forms I(A) and II(B) Buhler et al., 1980) representing the Eucyta. The results are schematically shown in Fig. 2.

With exceptions, the crossreactions between corresponding components of Archaebacteria of both branches were strong. The crossreactions between corresponding components of archaebacterial and eucytic polymerases were of similar strength, often significantly stronger than between corresponding components of the eukaryotic polymerases I (A) and II (B) themselves. The observed crossreactions between homologous components of eubacterial RNA polymerases on the one hand and archaebacterial and eukaryotic RNA polymerases on the other were significant, but usually weak (Fig. 3).

For the purpose of discussion of the correspondence of the component patterns of different RNA polymerase types, we have proposed a nomenclature, in which the heaviest and the second components of the nuclear yeast RNA polymerases I and II are called A and B (Schnabel et al., 1983). In the RNA polymerases of all sulfur dependent Archaebacteria and of Thermoplasma acidophilum, the corresponding components appear in reversed succession, i.e. B with the largest apparent size followed by A. In contrast, in the enzymes of the methanogenic and extremely halophilic Archaebacteria, the heaviest component corresponds again to the heaviest of yeast and the second of Sulfolobus and has therefore to be termed A. The second and the third components of these enzymes both react with Sulfolobus B antibody but do not crossreact with each other. Thus they each correspond to a different portion of the eukaryotic or sulfur dependent component B. To account for their equivalence, we propose to

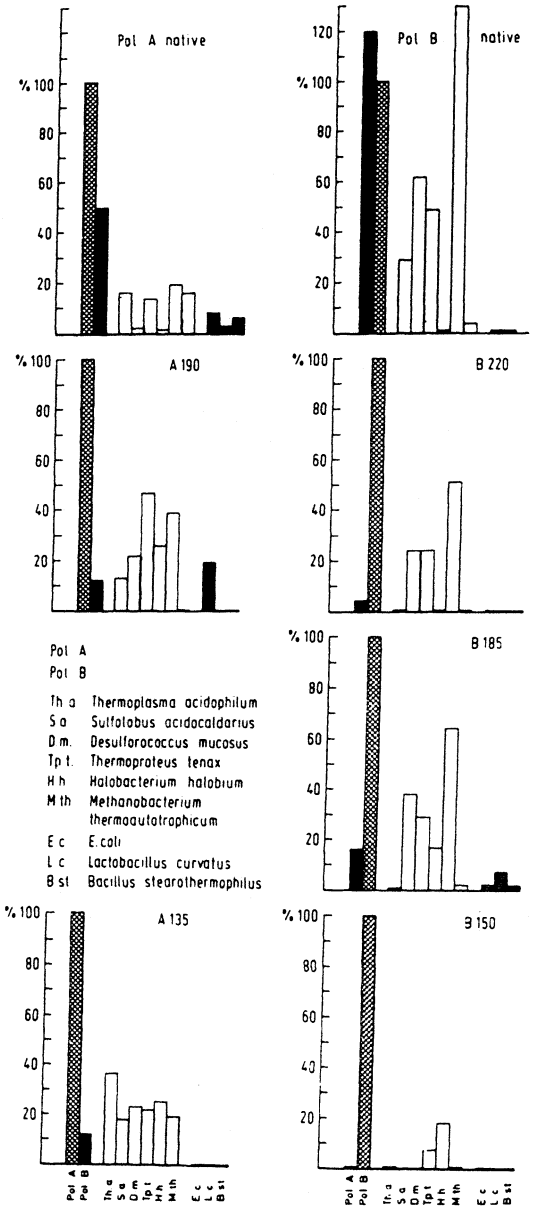

Fig. 3 Quantification of immunological crossreaction of different DNA dependent RNA polymerases with antibodies raised against native polymerase I (Pol A) and II (Pol B) of yeast and against the large components of these enzymes.

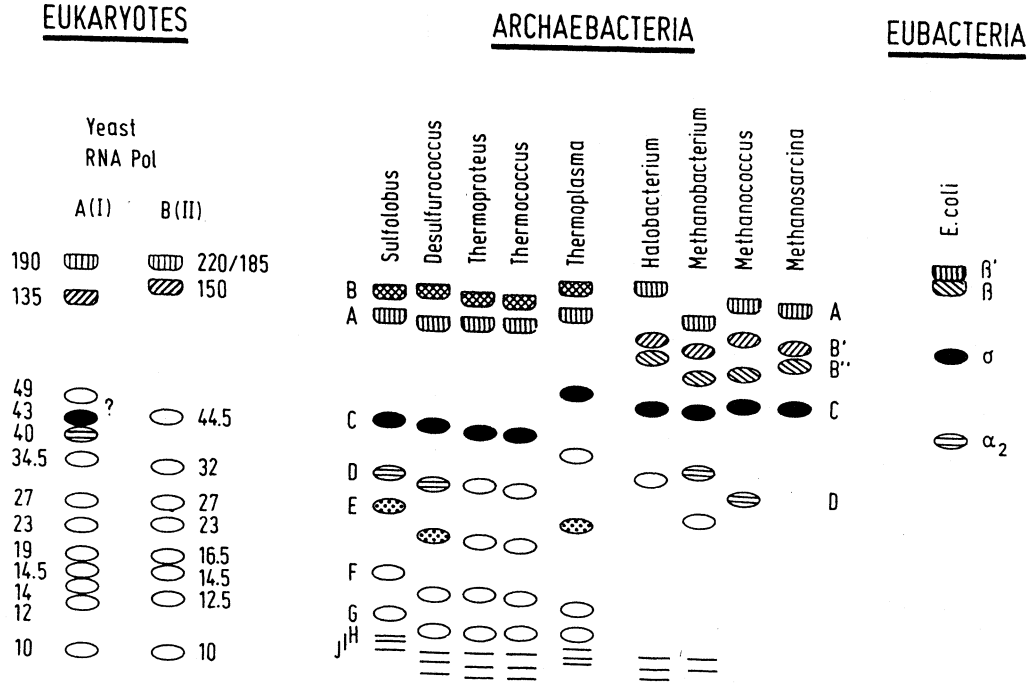

Fig. 2 Homologies between components in schematically drawn SDS polyacrylamide gel electrophoresis patterns as indicated by equal designs of corresponding bands.

substitute their previous designations B and B' by B' and B'', respectively, in the order of decreasing apparent size. The large B subunit of the sulfur dependent Archaebacteria and Eucyta is either a fusion product of B' and B'', or B' and B'' are splitting products of B, depending on the direction of evolution, whether on the levels of transcription or translation, or posttranslationally, remains to be determined.

Antibodies against the B components of yeast RNA polymerases I and II recognize only the B', not the B'' subunit of H. halobium polymerase. The antibody against component B' of the RNA polymerase of Methanobacterium recognizes the B components of sulfur dependent Archaebacteria and the B' components of polymerases from methanogens and halophiles exclusively.

The antibody against component B'' of Methanobacterium recognizes the B components from sulfur dependent Archaebacteria and the B'', but not the B' components from Methanobacterium and Methanococcus. In the pattern from Methanolobus volcani, a representative of the Methanomicrobiales, it reacts strongly with B'', yet significantly though weakly also with component B'. In the pattern from H. halobium, it reacts almost as strongly with B' as with B'' (Fig. 4). Thus, in the Methanococcales, the split separates the antigenic sites in the same way as in the Methanobacteriales. In the Methanomicrobiales, it is in a slightly different position, in the Halobacteriales in a quite different position. The translocation of the B → B'+B'' split in the Methanomicrobiales might be the reason for the smaller size difference of B' and B'' as compared to that seen in other orders of this branch. In conclusion, the B' and the B'' components of the Methanococcales and Methanobacteriales represent different portions of the large B component from those of the Methanomicrobiales and the Halobacteriales.

The third components in size, C, of the enzymes of sulfur dependent Archaebacteria and Thermoplasma are homologous to the fourth components of methanogenic and extremely halophilic polymerases, which thus also have to be termed C. In the same manner, the homology has also been proved for the fourth component of the sulfur dependent Archaebacteria, D, and, in several cases, the fifth component, E. For smaller components the limitations of the method do not allow recognition of correspondence.

Though the crossreactions between components of eukaryotic and archaebacterial RNA polymerases on the one and eubacterial RNA polymerases on the other hand are weak, they suffice to elucidate interkingdom homologies: E. coli β' corresponds to component A of yeast and Archaebacteria. E. coli β antibody crossreacts with component B of yeast and Sulfolobus, exclusively with component B'' of Methanobacterium and Methanococcus, mainly with component B'' but weakly with B' of Methanolobus (Methanomicrobiales), and almost as strongly with component B' as with B'' of Halobacterium (fig. 4). In spite of the generally lower reactivity, it is particularly striking that the ratio of the responses of the B' and B'' components of various representatives of the methanogenic and extremely halophilic Archaebacteria to E. coli β antibody is very similar to the ratio of the responses of the same components to Methanobacterium B'' antibody, indicating correspondence specifically between Methanobacterium (and Methanococcus) B'' and E. coli β.

Antibody against component C from Sulfolobus reacts with either component A 43 or A 49 of yeast RNA polymerase I. A more precise assignment would require better resolution. In the E. coli polymerase pattern, it reacts strongly with the σ factor (Fig. 4). The other way round, E. coli σ -antibody reacts weakly but also significantly with the C components of both Methanolobus and Sulfolobus, i.e. of members of both branches of the archaebacterial kingdom (Fig. 4). Thus, the eubacterial σ corresponds to the archaebacterial C subunit. In yeast RNA polymerase I, C is probably component A43 or component A49.

E. coli α antibody reacts weakly with component A40. Since Sulfolobus D antibody also reacts with A40, it is highly probable that α corresponds to A40 and to the archaebacterial D component, though a direct crossreaction of α and D has not been observed.

Component E of Sulfolobus seems to correspond to yeast A27. A crossreaction with an E. coli subunit has not been observed. E is apparently involved in the interaction of the archaebacterial enzymes with the DNA (Schnabel et al., 1982). For smaller components, interkingdom homologies could not be established by immunochemical crossreaction.

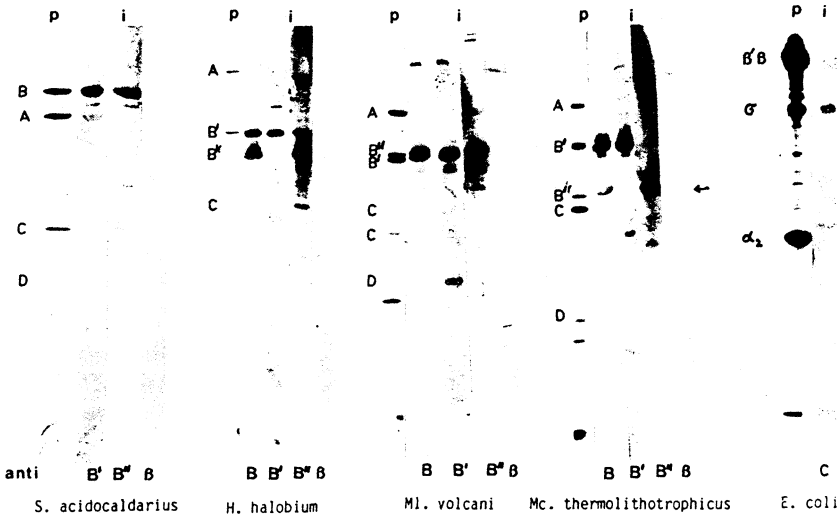

Fig. 4 Immunological crossreactions (Western blots) of components of archaebacterial and eubacterial RNA polymerases with antibodies raised against components B and C of S. acidocaldarius RNA polymerase, B' and B'' of Methanobacterium thermoautotrophicum RNA polymerase and β of E. coli RNA polymerase.

RNA polymerases, separated on SDS polyacrylamide gels and transferred to nitrocellulose sheets by electroblotting procedure, were challenged with antibodies.

Tracks labelled p show the component pattern visualized by Coomassie blue staining. Those labelled i show bound antibodies visualized with peroxidase-coupled antibodies.

3. Conclusions

(1) The large components A, B, C and D of the three types of eukaryotic nuclear RNA polymerases closely correspond to each other, to the components A, B, C and D of the polymerases of the sulfur dependent Archaebacteria and of Thermoplasma, to the components A, B', B'', C and D of the enzymes of the methanogenic and extremely halophilic Archaebacteria and, clearly though less closely to the subunits β', β, σ and α of the eubacterial RNA polymerases (Fig. 2).

This is convincing evidence for a common origin of all "normal" RNA polymerases and suggests a certain branching pattern in early evolution which shall be discussed below.

(2) Archaebacterial RNA polymerases differ from eukaryotic nuclear polymerases: The methanogens and halophiles by the B → B' + B'' division and the sulfur dependent Archaebacteria and Thermoplasma by a reversed size order of the components A and B. Both are insensitive to α amanitin. Archaebacterial differ from eubacterial RNA polymerases, as well in the considerably greater complexity of their component patterns, as in their insensitivity to rifampicin and streptolydigin.

(3) The more striking observation, however, is that they do not represent a uniform type, but rather clearly distinct types. The BACD... type realized in all RNA polymerases from sulfur dependent Archaebacteria and Thermoplasma differs from the ABCD... type represented by the three eukaryotic RNA polymerases only in the reversed size order of components A and B.

The AB'B''CD... types of the methanogenic and extremely halophilic Archaebacteria have apparently resulted from transcriptional, translational or posttranslational splits of the large component B of the sulfur dependent Archaebacteria and Thermoplasma.

In Methanolobus, representing the Methanomicrobiales, and more significantly in Halobacterium, this split leaves antigenic determinants linked in the respective B''s, which are separated in Methanobacterium B''. The split should therefore either have occurred independently (at least in its final consequence) several times in evolution, or it was specifically translocated by recombination processes after

the separation of the respective lineages. In the case of the opposite direction of evolution, these conclusions would have to be reversed: the B B'+B'' split would have been a B'+B'' → B fusion. Recombination as a cause for the differences between the B's and B''s each of the Methanobacteriales and Methanococcales, the Methanomicrobiales and the Halobacteriales would in this case have been a necessity. This is a strong argument against this sequence of events.

(4) The homologies between corresponding components of Eucyta and Archaebacteria (of both branches) are quantitatively much stronger than between the latter and eubacterial components (Fig. 4). They appear even stronger than between different types of eukaryotic nuclear RNA polymerase components: evidence for the large phylogenetic distance, which separates the present eukaryotes from their one-polymerase-ancestors or, the other way round, for the primitive nature of the present Archaebacteria.

(5) The failure of the E. coli β antibody to react with the B'-components of Methanobacterium and Methanococcus and the striking resemblance of the reactivity ratios of the E. coli β and Methanobacterium B'' antibodies with the Methanolobus and Halobacterium B' and B'' components respectively suggests that E. coli β corresponds to B'' of Methanobacterium. Thus E. coli polymerase which exhibits the same B → B'+B'' split as Methanobacterium (at least as revealed by immunochemical homology) could have arisen from the same ancestor by loss of the B' component and additional streamlining (Fig. 5).

(6) The B'' component of Methanobacterium, which corresponds to only part of the B'' of Halobacterium, crossreacts with E. coli β, but not with the B component of yeast RNA polymerases I and II. In contrast, the B' component of Halobacterium, which corresponds to only part of B' of Methanobacterium, crossreacts with the B component of the yeast polymerases, but not with E. coli β. The B component of sulfur dependent Archaebacteria crossreacts with both Methanobacterium B'' and Halobacterium B'. This situation is visualized schematically in Fig. 5. Eubacteria and Eucyta have conserved different antigenic determinants of the B component of the sulfur dependent Archaebacteria, besides a core of common determinants. This opposite trimming is a strong argument for the ancestral nature of the sulfur dependent B and more generally of the Archaebacteria.

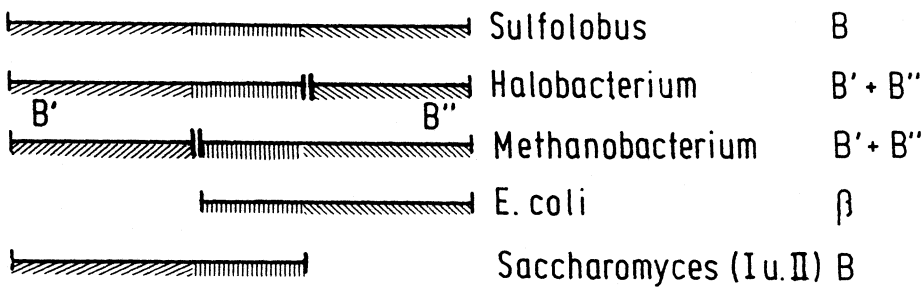

Fig. 5 Schematic representation of correspondence of E. coli β and Saccharomyces POL I B with B components of Halobacterium, Methanobacterium and Sulfolobus. Vertical lines connect corresponding structural elements i.e. antigenic sites. Horizontal lines do not give size of peptides but alignments of determinants.

(7) These data suggest that the original (ancestral) polymerase was that conserved in the sulfur dependent Archaebacteria and Thermoplasma (see Fig. 6).

The BACD type polymerase would then have given rise to the urkaryotic ABCD enzyme and to the (different!) AB'B''CD polymerases of the Halobacteriales and the methanogens, the first directly, the latter by a number of successive independent B → B'+B'' splits. Alternatively, different lineages could have separated after one original B → B'+B'' split, followed by its lineage specific recombinative translocation as discussed above. In any case, the eubacterial RNA polymerases appear derived from the same ancestral type as those of the Methanobacteriales and the Methanococcales, which are particularly close to each other, whereas the Methanomicrobiales are signifanctly more distant (Fig. 6). The big difference in the B split in H. halobium could be a consequence of the extreme genome instability of this organism but separates the Halobacteriales clearly from the other orders of the branch.

The events leading from an RNA polymerase of the Methanobacterium design to a typically eubacterial enzyme then were the loss of B' and additional streamlining by the loss of all (or most) components smaller than D (which corresponds to α).

It is not impossible though less probable that the eukaryotic polymerases were derived from the eubacterial enzymes by the opposite order of events: first acquisition of B' and of additional small components, then fusion of B'' and B'. A final decision between these two extreme trees has not been reached. The sequence of events, in one or the other direction, appears, however, logical and even cogent (Fig. 6).

(8) The RNA polymerases of Archaebacteria, especially of the sulfur dependent branch, appear strikingly "eukaryotic" like a number of other feature designs in Archaebacteria, e.g. the ADP ribosylation of EF2 of both branches by diphtheria toxin (Kessel and Klink, 1982), the existence, again in both branches, of replicating DNA polymerases sensitive to aphidicolin (Forterre et al., 1984; M. Nakayama and M. Kohiyama, personal communication; R. Schinzel and K.J. Burger, personal communication; H.P. Zabel, J. Winter, H. Fischer, E. Holler, personal communication; D. Prangishvilli and Zillig, 1984), the sequences of ribosomal "A-proteins"

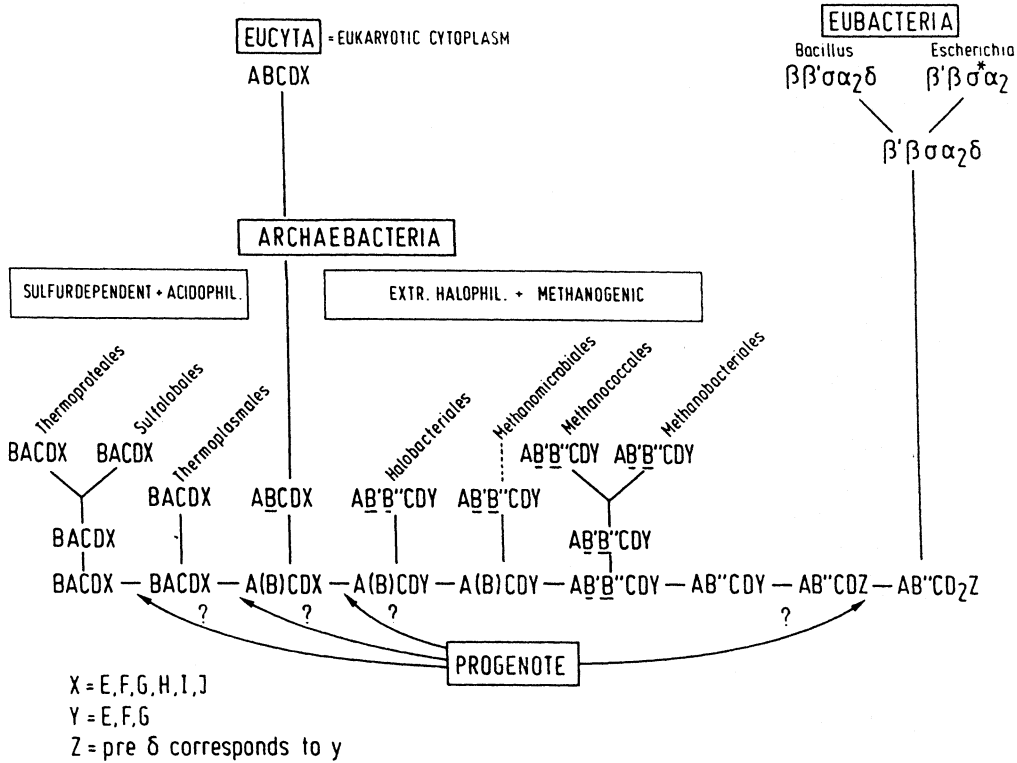

Fig. 6 Flow diagram of component patterns of different DNA dependent RNA polymerase types, corresponding to an unrooted tree. Most probable entering points for conversion to phylogenetic trees indicated by connecting lines from "progenote".

(Matheson and Yaguchi, 1982), the existence of introns in tRNA structural genes (Kaine et al., 1982; Daniels et al., 1984), the absence of formylation of methionyl initiator tRNA, the absence of the 3' terminal CCA sequence in the transcript and certain tRNA modifications (Gupta and Woese, 1980) (Table 2).

In a number of other feature designs, the sulfur dependent Archaebacteria appear closer to eukaryotes than the methanogens and extreme halophiles do (and, vice versa, the methanogens and halophiles closer to the Eubacteria). These comprise the shape of the large ribosomal subunit (Henderson et al., 1984; Lake et al., 1984), the 5S rRNA secondary structure as revealed by Fox et al. (1982), initiator tRNA structure (Kuchino et al., 1982), the complexity of ribosomal protein patterns (Schmid et al., 1982), the extent of rRNA modification (Woese et al., 1984) and cell division mechanisms (no septum formation). RNA polymerase composition appears a particularly valid feature of this type (Table 2).

However, Archaebacteria also have common designs considered typical for Eubacteria: they contain Shine Dalgarno sequences in mRNA and their complements in 16S rRNA, and restriction enzymes. The total sequences of 16S rRNA (Gupta et al., 1983; Leffers and Garrett, 1984; M. Jarsch and A. Böck, total sequence of 16S rRNA gene of Methanococcus vannielii, personal communication) also of the sulfur dependent Sulfolobus (R. Gupta and C. Woese, personal communication) appear more related to that of E. coli than to that of the eukaryote Dictyostelium. Archaebacteria are prokaryotes with all consequences. But this term refers rather to organisation than to phylogenetic status (Table 2).

(9) In summary, Archaebacteria appear as residues of an ancient primitive layer. One of its two surviving branches, comprising the sulfur dependent Archaebacteria, appears close to the ancestral Eucyta, the urkaryotes of Carl Woese, the other, methanogens and extreme halophiles, to the origin of the Eubacteria. An incomplete account of the basis of these views has been given previously (Zillig et al., in the press).

(10) Certain observations seem to contradict this picture, e.g. the exclusive existence of ether lipids in all known Archaebacteria, but of ester lipids in both modern kingdoms

Designs of Features of Archaebacteria

1. Eukaryotic, general	2. Eubacterial, general	3. More eukaryotic in sulfur dependent than in halophilic and methanogenic archaebacteria (and vice versa)
a) ADP ribosylation of EF2 by diphtheria toxin	a) Shine Dalgarno sequences	a) RNA polymerase structure
b) replicating DNA polymerases aphidicolin sensitive	b) occurrence of restriction enzymes	b) shape of large ribosomal subunit
c) ribosomal A protein sequence	c) 16S rRNA total sequences	c) 5S rRNA secondary structure
d) occurrence of introns	(d) prokaryotic organisation)	d) initiator tRNA structure
e) some tRNA modifications		e) complexity of ribosomal protein pattern
f) no formylation of methionyl initiator tRNA		f) cell division mechanisms
g) ACC 3' terminus of tRNA not encoded		

(with one abnormal exception). Another conflicting observation is the complete lack of any sensitivity to ribosome specific antibiotics in sulfur dependent Archaebacteria, whereas methanogens share the sensitivity to a few inhibitors with eukaryotes, that to several others with Eubacteria, and Eubacteria and eukaryotes share responses not exhibited by Archaebacteria (Böck and Kandler, in the press).

Such inconsistencies and the non-congruence of the evolution rates of different genes prohibit the understanding of early divergence of and in the three kingdoms until further comparative studies yield complete agreement of all facts.

References

(1) Achberger, E.C. and Whitely, H.R. (1981). The role of the Delta peptide of the Bacillus subtilis RNA polymerase in promoter selection. J. Biol. Chem. 256, 7424-7432.

(2) Böck, A. and Kandler, O. Antibiotic Sensitivity of Archaebacteria. In "The Bacteria", Vol. 8, in the press.

(3) Böck, A., Jarsch, M., Hummel, H. and Schmid, G. (1984). Evolution of Translation. This volume.

(4) Buhler, J.M. Huet, J., Davies, K.E., Sentenac, A. and Fromageot, P. (1980). Immunological studies of yeast nuclear RNA polymerases at the subunit level. J. Biochem. 225, 9949-9954.

(5) Burgess, R.R. (1976). Purification and Physical Properties of E. coli RNA Polymerase. In "RNA Polymerase". Losick, R., and Chamberlin, M., eds., pp. 69-100. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

(6) Daniels, C.J., Gupta, R. and Doolittle, W.F. (1984). Transcription and excision of a large intron in the $tRNA^{Trp}$ gene of an archaebacterium, Halobacterium volcanii. Submitted to J. Biol. Chem.

(7) DasSarma, S., RajBhandary, U.L. and Khorana, H.G. (1984). Bacterio-opsin mRNA in wild-type and bacterio-opsin deficient Halobacterium halobium strains. Proc. Natl. Acad. Sci. USA 81, 125-129.

(8) Dickel,C.D., Burtis, K.C. and Doi, R.H. (1980). Delta factor increases promoter selectivity by Bacillus subtilis vegetative cell RNA polymerase. Biochem. Biophys. Res. Commun. 95, 1789-1795.

(9) Forterre, P., Elie, C. and Kohiyama, M. (1984). Aphidicolin inhibits growth and DNA synthesis in halophilic archaebacteria. J. Bact. 159, 800-802.

(10) Fox, G.E., Pechmann, K.J., Woese, C.R. (1977). Comparative cataloging of 16S ribosomal RNA: molecular approach to procaryotic systematics. Int. J. System. Bact. 27, 44-57.

(11) Fox, G.E., Stackebrandt, E., Hespell, R.B., Gibson, J., Maniloff, J., Dyer, T.A., Wolfe, R.S., Balch, W.E., Tanner, R.S., Magrum, L.J., Zablen, L.B., Blakemore, R., Gupta, R., Bonen, L., Lewis, B.J., Stahl, D.A., Luehrsen, K.R., Chen, K.N., Woese, C.R. (1980). The phylogeny of prokaryotes. Science 209, 457-463.

(12) Fox, G.E., Luehrsen, K.R. and Woese, C.R. (1982). Archaebacterial 5S ribosomal RNA. Zbl. Bact. Hyg., Orig. C 3, 330-345.

(13) Gierl, A., Zillig, W. and Stetter, K.O. (1982). The role of the components σ and y of the DNA-dependent RNA polymerase of Lactobacillus curvatus in promotor selection. Eur. J. Biochem. 125, 41-47.

(14) Gupta, R. and Woese, C.R. (1980). Unusual modification patterns in the transfer ribonucleic acids of archaebacteria. Current Microbiology 4, 245-249.

(15) Gupta, R., Lanter, J.M. and Woese, C.R. (1983). Sequence of the 16S ribosomal RNA from Halobacterium volcanii, an archaebacterium. Science 221, 656-659.

(16) Gupta, R. (1984). Halobacterium volcanii tRNAs: Identification of 41 tRNAs covering all amino acids, and the sequences of 33 class I tRNAs. J. Biol. Chem., in the press.

(17) Henderson, E., Oakes, M., Clark M.W., Lake, J.A., Matheson, A.T. and Zillig, W. (1984). A new ribosome structure. Science, 225, 510-512.

(18) Herzfeld, F. and Kiper, M. (1976). The reconstitution of Anacystis nidulans DNA-dependent RNA polymerase from its isolated subunits. Eur. J. Biochem. 62, 198.

(19) Hori, T., Itoh, T. and Osawa, S. (1982). The phylogenic structure of the metabacteria. Zbl. Bact. Hyg., I. Abt. Orig. C 3, 18-30.

(20) Huet, J., Sentenac, A. and Fromageot, P. (1982). Spot-immunodetection of conserved determinants in eukaryotic RNA polymerases. J. Biol. Chem. 257, 2613-2618.

(21) Huet, J., Schnabel, R., Sentenac, A. and Zillig, W. (1983). Archaebacteria and eukaryotes possess DNA-dependent RNA polymerases of a common type. The EMBO J. 2, 1291-1294.

(22) Jarsch, M. and Böck, A. (1984). Sequence of the 16S ribosomal RNA gene from Methanococcus vannielii. Submitted to The EMBO J.

(23) Kaine, B.P., Gupta, R. and Woese, C.R. (1983). Putative introns in tRNA genes of prokaryotes. Proc. Natl. Acad. Sci. USA 80, 3309-3312.

(24) Kandler, O. (1982). Cell wall structures and their phylogenetic implications. Zbl. Bakt. Hyg. I. Abt. Orig. C 3, 149-160.

(25) Kessel, M. and Klink, F. (1982). Identification and comparison of eighteen archaebacteria by means of the diphtheria toxin reaction. Zbl. Bakt. Hyg., I. Abt. Orig. C 3, 140-148.

(26) Konheiser, U., Pasti, G., Bollschweiler, C. and Klein, H. (1984). Physical mapping of genes coding for few subunits of methyl CoM reductase component C of Methanococcus voltae. Mol. Gen. Genet., in the press.

(27) Kuchino, Y., Hiara, M., Yabusaki, Y. and Nishimura, S. (1982). Initiator tRNAs from archaebacteria show common unique sequence characteristics. Nature 298, 684-685.

(28) Lake, J.A., Henderson, E., Oakes, M. and Clark, M.W. (1984). Eocytes: A new ribosome structure indicates a kingdom with a close relationship to eukaryotes. Proc. Natl. Acad. Sci. USA 81, 3786-3790.

(29) Langworthy, T.A., Tornabene, T.G. and Holzer, G. (1982). Lipids of archaebacteria. Zbl. Bakt. Hyg., I. Abt. Orig. C 3, 228-244.

(30) Leffers, H. and Garrett R.A. (1984). The nucleotide sequence of the 16S ribosomal RNA gene of the archaebacterium Halococcus morrhuae. The EMBO J. 3, 1613-1619.

(31) Madon, J., Leser, U. and Zillig, W. (1983). DNA-dependent RNA polymerase from the extremely halophilic archaebacterium Halococcus morrhuae. Eur. J. Biochem. 135, 279-283.

(32) Madon, J. and Zillig, W. (1983). A form of the DNA-dependent RNA polymerase of Halobacterium halobium, containing an additional component, is able to transcribe native DNA. Eur. J. Biochem. 133, 471-474 (1983).

(33) Matheson, A.T. and Yaguchi, M. (1982). The evolution of the archaebacterial ribosome. Zbl. Bakt. Hyg., I. Abt. Orig. C 3, 192-199.

(34) Prager, E.M., Wilson, A.C., Lowenstein, J.M. and Sarich, V.M. (1980). Mammoth albumin. Science 209, 287-289.

(35) Prangishvilli, D., Zillig, W., Gierl, A., Biesert, L. and Holz, I. (1982). DNA-dependent RNA polymerases of thermoacidophilic archaebacteria. Eur. J. Biochem. 122, 471-477.

(36) Prangishvilli, D. and Zillig, W. (1984). DNA-Dependent DNA-Polymerases of Thermoacidophilic Archaebacterium Sulfolobus acidocaldarius. Poster presented at the FEMS Symposium "Evolution of Prokaryotes", held in Munich, September 16-18, 1984.

(37) Rexer, B. (1976). Struktur und Aktivität der DNA-abhängigen RNA-Polymerase aus vegetativen und sporulierenden Zellen von B. cereus T. Thesis, Ludwig-Maximilians-Universität München.

(38) Roeder, R.G. (1976). Eukaryotic nuclear RNA polymerases. In "RNA Polymerase". Losick, R. and Chamberlin, M., eds. pp. 285-329, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

(39) Schmid, G., Pecher, Th. and Böck, A. (1982). Properties of the translational apparatus of the archaebacteria. Zbl. Bakt. Hyg., I. Abt. Orig. C 3, 209-217.

(40) Schnabel, R., Zillig, W. and Schnabel, H. (1982). Component E of the DNA-dependent RNA polymerase of the archaebacterium Thermoplasma acidophilum is required for the transcription of native DNA. Eur. J. Biochem. 129, 473-477.

(41) Schnabel, R., Thomm, M., Gerardy-Schahn, R., Zillig, W., Stetter, K.O. and Huet,J. (1983). Structural homology between different archaebacterial DNA-dependent RNA polymerases analyzed by immunological comparison of their components. The EMBO J. 2, 751-755.

(42) Stetter, K.O. and Zillig, W. (1974). Transcription in Lactobacillaceae. DNA-dependent RNA polymerase from Lactobacillus curvatus. Eur. J. Biochem. 48, 527-540.

(43) Stetter, K.O., Winter, J. and Hartlieb, R. (1980). DNA-dependent RNA polymerase of the archaebacterium Methanobacterium thermoautotro phicum. Zbl. Bakt. Hyg., I. Abt. Orig. C 1, 201-218.

(44) Sturm, S., Schönefeld, V., Zillig. W., Janekovic, D., and Stetter, K.O. (1980). Structure and function of the DNA-dependent RNA polymerase of the archaebacterium Thermoplasma acidophilum. Zbl. Bakt. Hyg., I. Abt. Orig. C 1, 12-25.

(45) Thomm, M., Altenbuchner, J. and Stetter, K.O. (1983). Evidence for a plasmid in a methanogenic bacterium. J. Bacteriol. 153, 1060-1062.

(46) Woese, C.R. and Fox, G.E. (1977). The concept of cellular evolution. J. Mol. Evol. 10, 1-6.

(47) Woese, C.R., Magrum, L.J. and Fox, G.E. (1978). Archaebacteria. J. Mol. Evol. 11, 245-252.

(48) Woese, C.R., Gupta, R., Hahn, C.M., Zillig, W. and Tu, J. (1984). The phylogenetic relationships of three

sulfur dependent archaebacteria. System. Appl. Microbiol. 5, 97-105.

(49) Yaguchi, M., Visentin, L.P., Zukur, M., Matheson, A.T., Roy, C. and Strom, A.R. (1982). Amino-terminal sequences of ribosomal proteins from the 30S subunit of archaebacterium Halobacterium cutirubrum. Zbl. Bakt. Hyg., I. Abt. Orig. C 3, 200-208.

(50) Zillig, W., Palm, P. and Heil, A. (1976). Function and Reassembly of Subunits of DNA-Dependent RNA Polymerase. In "RNA Polymerase". R. Losick and M. Chamberlin, eds., pp. 101-125. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

(51) Zillig, W., Stetter, K.O. and Tobien, M. (1978). DNA-dependent RNA polymerase from Halobacterium halobium. Eur. J. Biochem. 91, 193-199.

(52) Zillig, W., Stetter, K.O., Wunderl, S., Schulz, W., Priess, H. and Scholz, I. (1980). The Sulfolobus-"Caldariella" group: Taxonomy on the basis of the structure of DNA-dependent polymerases. Arch. Microbiol. 125, 259-269.

(53) Zillig, W., Stetter, K.O., Schnabel, R., Madon, J. and Gierl, A. (1982) Transcription in archaebacteria. Zbl. Bakt. Hyg., I. Abt. Orig. C 3, 218-227.

(54) Zillig, W., Gierl, A., Schreiber, G., Wunderl, S., Janekovic, D., Stetter, K.O. and Klenk, H.P. (1983a). The archaebacterium Thermofilum pendens represents a novel genus of the thermophilic, anaerobic sulfur respiring Thermoproteales. System. Appl. Microbiol. 4, 79-87.

(55) Zillig, W., Holz, I., Janekovic, D., Schäfer, W. and Reiter, W.D. (1983b). The archaebacterium Thermococcus celer represents a novel genus within the thermophilic branch of the archaebacteria. System. Appl. Microbiol. 4, 88-94.

(56) Zillig, W., Schnabel, R. and Stetter, K.O. (1984). Archaebacteria and the Origin of the Eukaryotic Cytoplasm. In "Current Topics of Microbiology and Immunology", in the press.