Volume 12 Number 8 1984 Nucleic Acids Research

Primary and secondary structure of rat 28 S ribosomal RNA

A.A.Hadjiolov, O.I.Georgiev, V.V.Nosikov* and L.P.Yavachev

Department of Molecular Genetics, Institute of Molecular Biology, Bulgarian Academy of Sciences, Sofia, Bulgaria, and *Institute of Molecular Biology, Academy of Sciences of the USSR, Moscow, USSR

Received 2 February 1984; Revised and Accepted 3 April 1984

ABSTRACT

The primary structure of rat (*Rattus norvegicus*) 28 S rRNA is determined inferred from the sequence of cloned rDNA fragments. The rat 28 S rRNA contains 4802 nucleotides and has an estmated relative molecular mass (M_r, Na-salt) of 1.66 x 10^6. Several regions of high sequence homology with *S.cerevisiae* 25 S rRNA are present. These regions can be folded in characteristic base-paired structures homologous to those proposed for *Saccharomyces* and *E.coli*. The excess of about 1400 nucleotides in the rat 28 S rRNA (as compared to *Saccharomyces* 25 S rRNA) is accounted for mainly by the presence of eight distinct G+C-rich segments of different length inserted within the regions of high sequence homology. The G+C content of the four insertions, containing more than 200 nucleotides, is in the range of 78 to 85 percent. All G+C-rich segments appear to form strongly base-paired structures. The two largest G+C-rich segments (about 760 and 560 nucleotides, respectively) are located near the 5'-end and in the middle of the 28 S rRNA molecule. These two segments can be folded into long base-paired structures, corresponding to the ones observed previously by electron microscopy of partly denatured 28 S rRNA molecules.

INTRODUCTION

Ribosomal RNAs constitute the backbones in the building of the two ribosomal particles. Accordingly, determination of the primary and secondary structures of rRNA constitute an important step in the elucidation of ribosome structure and function. The L-rRNAs of the large ribosomal particles are of particular interest since they are subject to considerable changes during evolution (1,2). Recently, the complete sequence of L-rRNA from *E.coli* and other prokaryotes was established (3,4) and plausible secondary structure models, proposed (4-7). The structure of cytoplasmic L-rRNA of eukaryotes is less understood. The primary structure of *Saccharomyces cerevisiae* (8), *S.carlsbergensis* (9) and *Physarum polycephalum* (10) has been determined and a model for the secondary structure of yeast L-rRNA was deduced (9). The sequence of L-rRNA from higher eukaryotes is still unexplored. Yet, it is known that a marked increase in size is characteristic of these L-rRNA, e.g.

mammalian L-rRNA exceed by about 1500 nucleotides their counterpart in yeast (1). Earlier studies revealed that the difference may be due to the presence of large (G+C)-rich segments in mammalian L-rRNA (11-14). It was shown also that these (G+C)-rich segments resist denaturation treatments and may be visualized by electron microscopy at specific sites along the chain of the L-rRNA (15-18).

In the present work we report the complete sequence of the rat 28 S rRNA inferred from the sequence of the respective rDNA included in a rat rDNA fragment cloned in a λCharon 4A vector (19). A model for the secondary structure of rat 28 S rRNA is also presented and a comparison with the respective structures in yeast is carried out with special emphasis on the location and structure of (G+C)-rich segments.

MATERIALS AND METHODS

The whole 28 S rRNA gene of the rat is included in the λRr IV fragment cloned previously with the use of a Charon 4A vector (19). Treatment with the restriction nucleases EcoR I and BamH I releases four rDNA fragments containing 28 S rDNA sequences. These rDNA fragments, subcloned in pBr 322 by standard methods are designated pRr 19, pRr 13, pRr 20 and pRr 22 [Figure 1] and are used in sequence analysis. As shown earlier (19) the λRr IV rDNA fragment has a distinct organization as compared with the rat λChR - B4 rDNA fragment cloned by Rothblum et al. (20). The 5'- and 3'-termini of the 28 S rRNA gene were identified by S_1-nuclease protection mapping (21) using purified cytoplasmic 28 S rRNA from rat liver. Both termini were homogeneous (data not shown) and their location within the rDNA sequence coincides with previous data for the site of the 5'-end of rat (22) and the 3'-end of mouse (23) 28 S rRNA. The 5'-terminus of the rat 28 S rRNA maps in rDNA fragment pRr 19, while the 3'-terminus is in pRr 22. The sequence of the 28 S rRNA gene included in the four rDNA fragments was determined by the method of Maxam and Gilbert (24) using appropriate rDNA subfragments obtained with different restriction endonucleases. The restriction endonucleases Ava I, Ava II, BamH I, BstN I, Bgl II, Hinf I, Msp I, Pvu II, Taq I and Xma I were products of Biolabs Inc., Beverly, Mass., U.S.A. The restriction endonucleases EcoR I, Mva I and Sau 3A were a kind gift by Dr A.Yanulaitis. The [γ-^{32}P]ATP was prepared by the method of Walseth and Johnson (25) using ^{32}P-orthophosphate from the Radiochemical Centre, Amersham, Bucks, U.K.

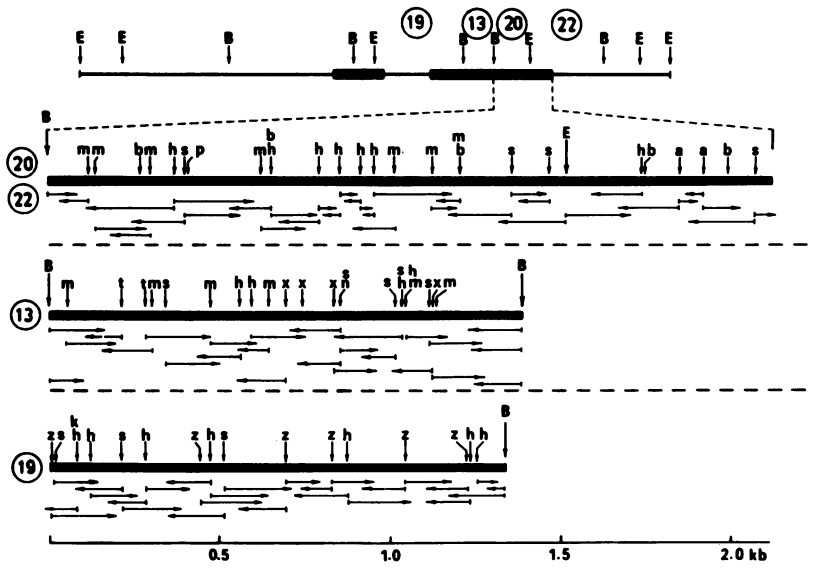

Figure 1. Structure of the rat rDNA repeating unit (above) and expanded maps of the sequenced rDNA fragments containing the 28 S rRNA gene.

The rDNA fragments pRr 19, pRr 13, pRr 20 and pRr 22 (encircled numbers) are obtained by digestion of the λCharon 4A vector, containing λRr IV rDNA (19) with the restriction nucleases EcoR I and BamH I. These rDNA fragments are subcloned in pBr 322 and used for the sequencing of the 28 S rRNA gene. The restriction nuclease sites generating the subfragments used in sequence determination are designated by vertical arrows. Some restriction nuclease sites (e.g. for Msp I) generating shorter rDNA subfragments are not shown. E - EcoR I; B - BamH I; a - Ava I; z - Ava II; b - BstN I; n - Bgl II; h - Hinf I; m - Msp I; k - Mva I; p - Pvu II; s - Sau 3A; t - Taq I; x - Xma I. The horizontal arrows indicate the direction and the size of the sequenced rDNA restriction subfragments obtained with the respective nuclease.

RESULTS AND DISCUSSION

The location of the fragments subcloned in the recombinant plasmids pRr 19, 13, 20 and 22 within the rat rDNA repeating unit and the strategy used in their sequencing are presented in Figure 1. The primary structure of rat 28 S rRNA inferred from the rDNA sequence is given in Figure 2.

The rat 28 S rRNA contains 4802 nucleotides and its G+C content is 67.8%. Knowledge of the rRNA primary structure permits a precise evaluation of its molecular mass. Using the known M_r values for the monosodium salt of the respective nucleotide residues (AMP - 352; GMP - 368; UMP - 329; CMP - 328) the calculated relative molecular mass (M_r) for rat 28 S rRNA is 1.66×10^6.

```
pCCGCGACCUCAGAUCAGACGUGGCGACCCGCUGAAUUUAAGCAUAUUAGUCAGCGGAGGAAAAGAAACUAACCAGGAUUCCCUCAGUAACGGCGAGUGAACAGGGAAGAG
CCCAGCGCCGAAUCCCGCCGCGCGCCGCGGCGCGGGAAAUGUGGCGUACGGAAGACCCACUCCCCGGCGCCGCUCGUGGGGGGCCCAAGUCCUUCUGAUCGAGGCCCAGC
CCGUGGACGGUGUGAGGCCGGUAGCGGCCCCGGCGCGCCGGGCCCGGGUGUUCCCGGAGUCGGGUUGCUUGGGAAUGCAGCCCAAAGCGGGUGGUAACUCCAUCUAAGGC
UAAAUACCGGCACGAGACCGAUAGCCAACAAGUACCGUAAGGGAAAGUUGAAAAGAACUUUGAAGAGAGAGUUCAAGAGGGCGUGAAACCGUUAAGAGGUAAACGGGUGG
GGUCCGCGCAGUCCGCCCGGAGGAUUCAACCCGGCGGCGCGUCCGGCCGGCCGGUGGUCCCGGCGGAUCUUUCCCGCACCCCGUUCCUCCCGACCCCUCCACCCGCGCGU
CUCUCUCCCCCCUCCCCGCGUCCCGCCGUCGCCGUCCCCGCUCCUCCCUCCGGGGGGGUGUCGGCGGGCGCUCCGGCGGCGGGCGCGGGUGUGGUGGGGGCGCGCGGGCG
GGGCCGGGGGUGGGGUCGGGCGGGGGACCGCCCCCCGGUCGGCGACCGGCCGCCGCCGGGCGCACUUCCACCGUGGCGGUGGCCCGCGACCGGCUCCGGGACGGCUGGGA
AGGCCCGGCGGGGAAGGUGGCUCGGGGGGGCGGCGUCACCCGUGGGCGCCGGACCCCGCCCCGAGUGUUACAGCCCCCCGGGCAGCAGCGCUCGCCGAAUCCCGGGGCC
GAGGGAGCCGGAUACCCGUCGCCGCGCUCUCCCCCCGGCCUCUCCCCUCCCGCCCCUCCCCGUGGGGUGGCGGAAAGGGGGGCGGUCGCGGGGGCCGGGCCGCCCCUCCC
ACGGGCGCGACCGCUCUCCCACCCCCCGCGCCUCCGUCGCCCUCUCGGGGUCCGGGGCCCGGGGGGCGGGGCGAACUCUCCCCAGUCCGCCCCGGGCGUCGUCGCGCCGU
CGGGCCCGGGGGGCCGUCCUCACGCGCUCUCCCCCCCCUUCUCGGGGUGGGGGGAGCGAAGCCGAGCGCACGGGGUCGGCGGCGAUGUCGGCUACCCACCCGACCCGUC
UUGAAACACGGACCAAGGAGUCUAACGCGUGCGCGAGUCAGGGGCUCGUCCGAAAGCCGCCGUGGCGCAAUGAAGGUGUGGGCCCCGUUCCCGGGGGCCCCGAGGUGGGA
UCCCGAGGCCUCUCCAGUCCGCCGAGGGCGCACCACCGGCCCGUCUCGCCCGCCGCGCCGGGGAGGUGGAGCACGAGCGUACGCGUUAGGACCCGAAAGAUGGUGAACUA
UGCCUGGGCAGGGCGAAGCCAGAGGAAACUCUGGUGGAGGUCCGUAGCGGUCCUGACGUGCAAAUCGGUCGUCCGACCUGGGUGUAGGGGCGAAAGACUAAUCGAACCAU
CUAGUAGCUGGUUCCUCCCGAAGUUUCCCUCAGGAUAGCUGGCGCUCUCGCAACGCGUUCGCUCGAACCCGCAGUUUUAUCCGGUAAAGCGAAUGAUUAGAGGUCUUGGG
GCCGAAACGAUCUCAACCUAUUCUCAAACUUUAAAUGGGUAAGAAGCCCGGCUCGCUGGCGUGGAACCGCGAGAAUGCGAGUGCCUAGUGGGCCACUUUUGGUAAGCAGA
ACUGGCGCUGCGGGAUGAACCGAACGCCGGGUUAAGGCGCCCGAUGCCGACGCUCAUCAGACCCCAGAAAAGGUUGUUGUUAUUAGACAGCAGGACGGUGGCCAUGGAGG
UCGGAAUCCGCUAAGGGGUGUGUAACAACUCACCUGCCGAAUCAACUAGCCCUGGGGUGGAUGGCGCUGGAGCGUCGGGCCCAUACCCGGCCGUCGCCGGCAGUCGGAAC
GGGACGGGAGCGGCCGCGGGCGCGCGACCCCGGGGGCCGGGCGGCGUCGGCUUCGGCCGGCCGCCGCCCGUCCACCCCCGGGGCUCCCCCCGCGGCGUCGGGCCCGCGGA
GCCUACGCCGCGACGAGUAGGAGGGCCGCUGCGGUCAGCCUUGAAGCCUAGGGCGCGGGCCCGGGUGGAGCCGCCGCAGGUGCAGAUCUUGGUGGUGGUAGCAAAUAUUC
AAACGAGAACUUUGAAGGCCGAAGUGGAGAAGGGUUCCAUGUGAACAGCAGUUGAACAUGGGUCAGUCGGUCCUGAGAGAUGGGCGAGUGCCGUUCCGAAGGGACGGGCG
AUGGCCUCCGUUGCCCUCAGCCGAUCGGAACUGGAGUCGGGGUUCAGAUCCCCGAAUCCGGAGUGGCGGAGAUGAAGUCGGGCGCCCGCGAGGCGUCCAGUGCCGGUAAC
GCACCGAUCCCGGAGAAGCCGGCGGGAGCCCGGGGAGAGUUCUCUUUUCUUUUGUGAAGGGCAGGGCGCCUUGGAAUGGGUUCGCCCCGAGAGAGGGGCCGUGCCUUGGA
AAGCGUCGCGGUUCCGGCGGCGUCCGGUGAGCUCUCGCUGGCCCUUGAAAAUCCGGGGGAGAGGGUGUAAAUCUCGCGCCGGGCCGUACCCAUAUCCGCAGCAGGUCUCC
```

Figure 2. Primary structure of the rat 28 S rRNA as inferred from the sequence of cloned rDNA fragments containing the gene. The BamH I and EcoR I restriction sites delineating the cloned rDNA fragments shown in Figure 1 are boxed. The numbering is in hunderds.

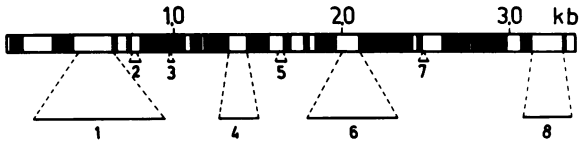

Figure 3. Distribution of conserved and non-conserved segments in *Saccharomyces cerevisiae* and *Rattus norvegicus* L-rRNA.
Regions of more than 80 % sequence homology are designated by black areas aligned along the yeast sequence. The position and the size of identified (G+C)-rich segments (length > 20 nucleotides) in the rat 28 S rRNA are denoted by full horizontal lines and numbered from 1 to 8. The size of the corresponding segments in *Saccharomyces cerevisiae* 25 S rRNA is denoted by the broken lines.

This value is in good agreement with previous estimates (1.65-1.66 x 10^6) obtained by physicochemical analysis of 28 S rRNA from rat and other mammalian species (26,27). The identity of M_r values obtained by the analysis of either rRNA or rDNA makes highly unlikely the removal by splicing of internal segments in 28 S rRNA.

A comparison for sequence homologies between the rat and yeast L-rRNA reveals the existence of several regions of high or moderate homology distributed all along the L-rRNA chain (Figure 3). The regions of high sequence homology are in three major zones delimited by nucleotides (rat) 1405-1945, 3401-3703 and 3888-4307. The regions of moderate homology encompass the 5'-terminal segment (about 430 nucleotides) and a poorly delineated region in the middle of 28 S rRNA (2130-2740).

The regions of high and moderate sequence homology are split by regions of variable length where the homology is lost. In the rat 28 S rRNA all non-homologous regions are characterized by an exceedingly high G+C content (above 75 %) and therefore may be designated as (G+C)-rich segments (GCS). The topology of the eight clearly delineated GCS in the rat 28 S rRNA is also shown in Fig.3. As it can be seen, the length of the identified GCS varies from about 30 in GCS 3 to 760 nucleotides in GCS 1. The GCS shown in Fig.3 replace respective segments in yeast 25 S rRNA although markedly expanded in size. The only exception is provided by GCS 7 (3716-3768) interrupting a fully conserved sequence. Therefore, GCS in L-rRNA cannot be considered strictly as insertions. The great expansion in size of some GCS accounts for practically the total excess of about 1400 nucleotides in rat 28 S rRNA. It is noteworthy that in addition to the expanded GCS, rather long (G+C)-rich stretches, replacing comparable in size stretches of yeast 25 S rRNA, may be

Table 1. Nucleotide composition of the large (G+C)-rich segments in rat 28 S rRNA

Segment of rRNA [position]	N*	Molar ratio				G + C percent
		G	C	A	U	
Total 28 S rRNA	4802	35.7	32.1	16.2	16.0	67.8
GCS 1 [445 - 1204]	760	39.5	42.4	6.7	11.4	81.9
GCS 4 [1946 - 2169]	224	43.3	39.7	9.4	7.6	83.0
GCS 6 [2761 - 3320]	560	40.9	43.7	3.4	12.0	84.6
GCS 8 [4463 - 4687]	225	35.1	43.1	10.7	11.1	78.2

* N - number of nucleotides

also discerned. Two such substituting (G+C)-rich segments are prominent in the rat 28 S rRNA (not shown in Fig.3) encompassing nucleotides 2479-2552 (73 % G+C) and 3818-3884 (85 % G+C).

Because the boundaries of GCS in the rat 28 S rRNA are not always neatly delineated, a more precise analysis of their nucleotide composition is possible only for the four large GCS. The data (Table 1) show that the G+C content of these GCS is in the range of 78 to 85 %, figures markedly higher than the G+C content of total 28 S rRNA. Further, although numerous rather long (10-20 nucleotides) GC-stretches are common to all large GCS, single nucleotide (G or C) stretches longer than 9 nucleotides are not found.

The existence of numerous regions of sequence homology between rat and yeast L-rRNA prompted us to attempt the construction of a model for the secondary structure of rat 28 S rRNA. For the conserved regions in the rat 28 S rRNA the comparative approach to model construction (28,29) was followed with *Saccharomyces* as reference (9) and with the models for *E.coli* (4-7) as a background. We adopted the delineation of structural domains and the presentation given by Veldman *et al.* (9) for the secondary structure of *Saccharomyces* 25 S rRNA since it appears to be most closely related to the rat 28 S rRNA. Where possible, the numbering of concensus helices in *E.coli* 23 S rRNA given by Maly and Brimacombe (7) is also included. Possible secondary structure of non-conserved segments in rat 28 S rRNA is based on a computer-assisted analysis of possible base-pairing interactions within the respective rRNA segments. In the case of GCS 1 and GCS 6 we probed also the

possibility for their folding into double-stranded structures analogous to the ones observed by electron microscopy of partly denatured L-rRNA molecules of different vertebrates (15-17), including the rat (17,18). We used the extended (instead of the skeleton) presentation in order to outline the actual interactions of individual nucleotides. The numbering used throughout corresponds to that for the rat 28 S rRNA given in Fig.2.

In Figure 4 A to H we present a model for the secondary structure of rat 28 S rRNA encompassing domains I to VII (9). The following aspects of the

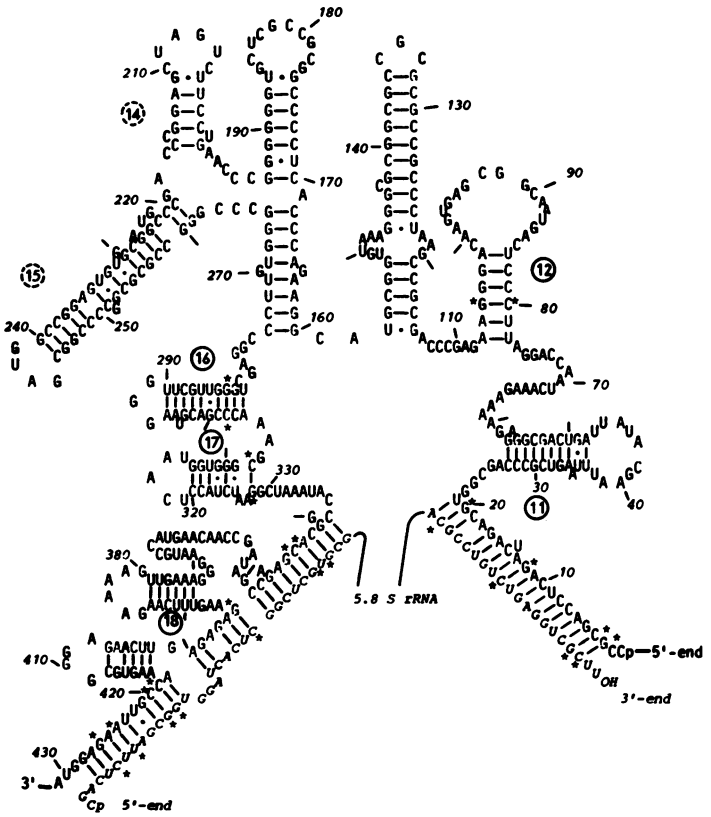

Figure 4 (general legend). Model for the secondary structure of rat 28 S rRNA.

The secondary structure domains I to VII correspond to those proposed for *Saccharomyces* (9). The numbering (italics) is that for rat 28 S rRNA (see Fig.2). Compensating base changes are denoted by *. The encircled numbers denote possible correspondance to concensus helices in *E.coli* (7).
Figure 4-A. Domain I. The segments of rat 5.8 S rRNA possibly base-paired to 28 S rRNA sequences are presented in italics.

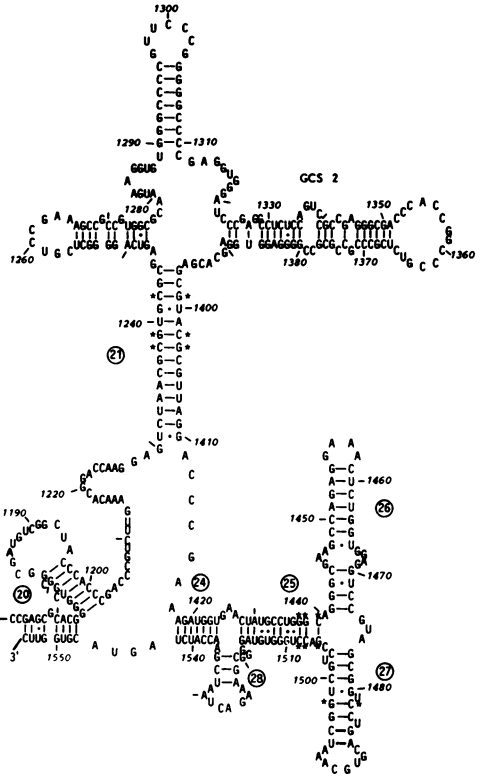

Figure 4-B. Domain II (conserved part) of rat 28 S rRNA

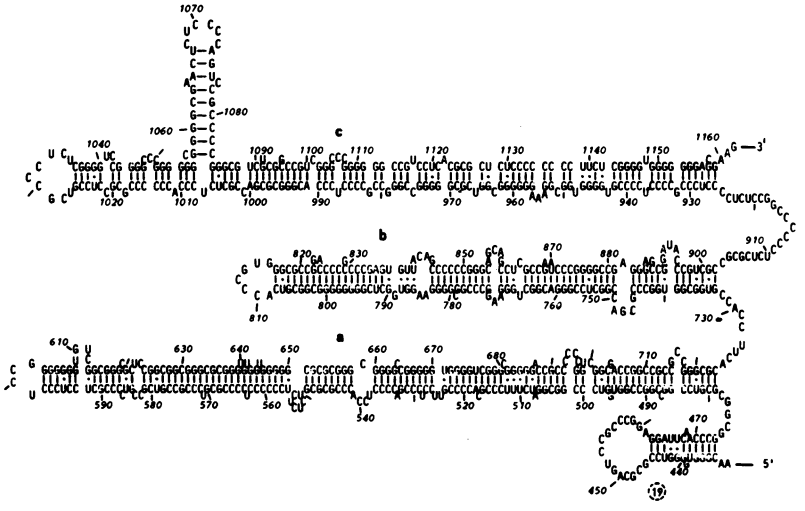

Figure 4-C. GCS 1 located at the junction between domain I and II of rat 28 S rRNA

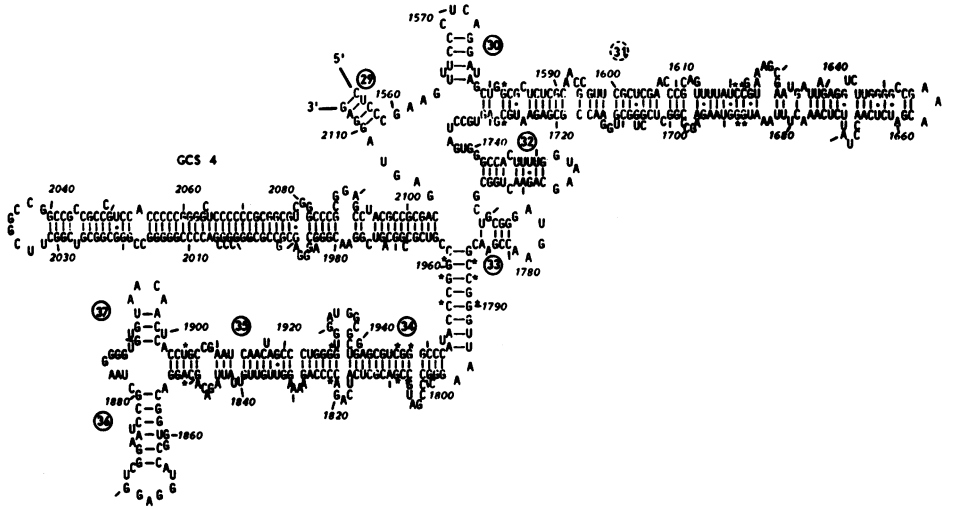

Figure 4-D. Domain III of rat 28 S rRNA

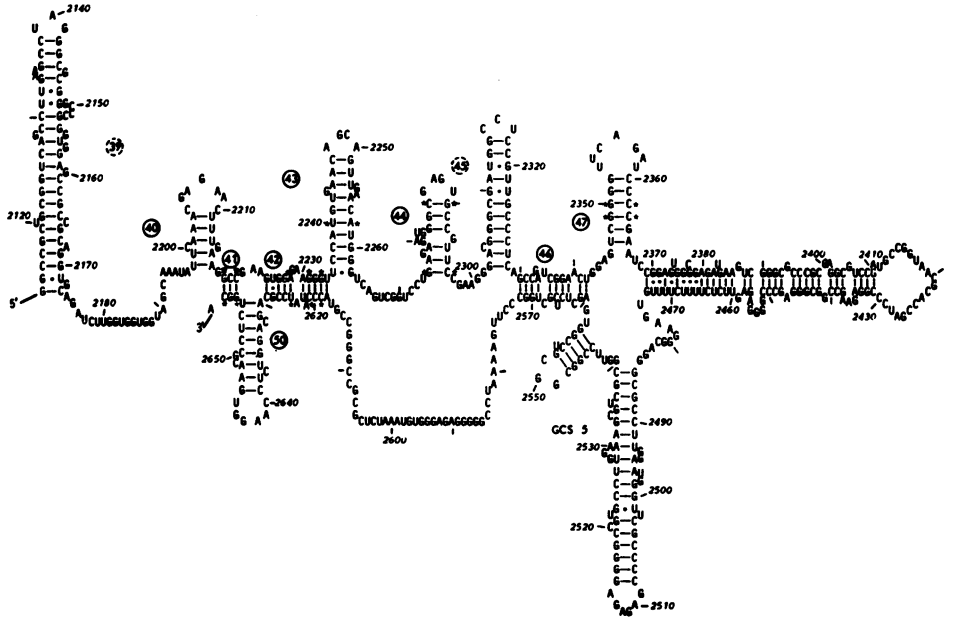

Figure 4-E. Domain IV of rat 28 S rRNA

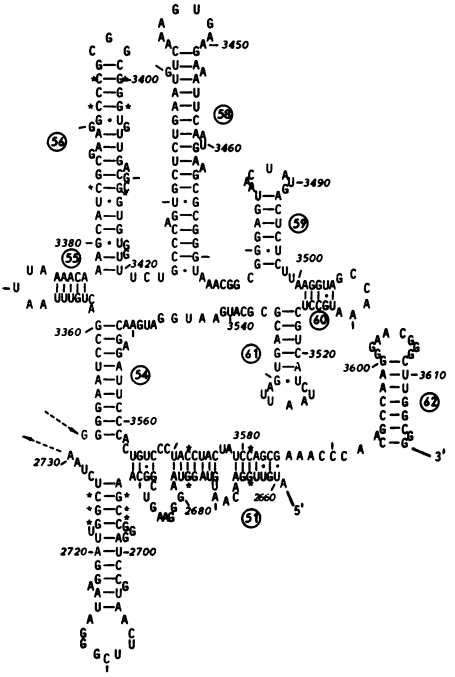

Figure 4-F. Domain V (conserved part) of rat 28 S rRNA

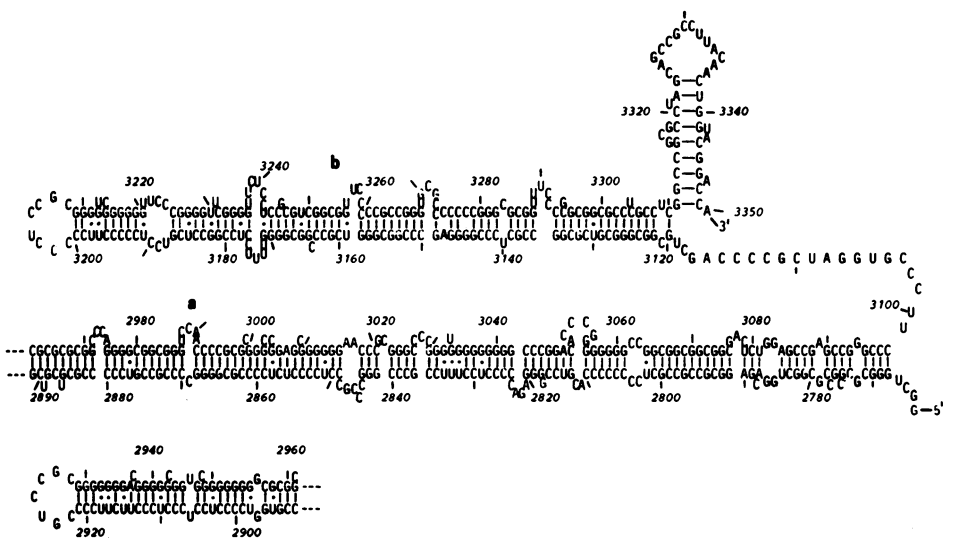

Figure 4-G. Secondary structure of GCS 6 in domain V of rat 28 S rRNA

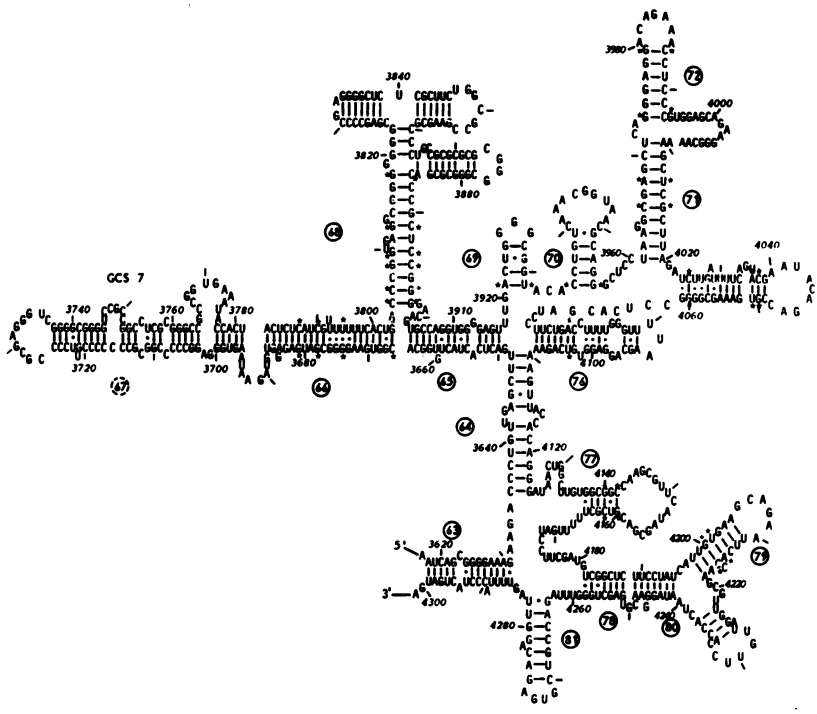

Figure 4-H. Domain VI of rat 28 S rRNA

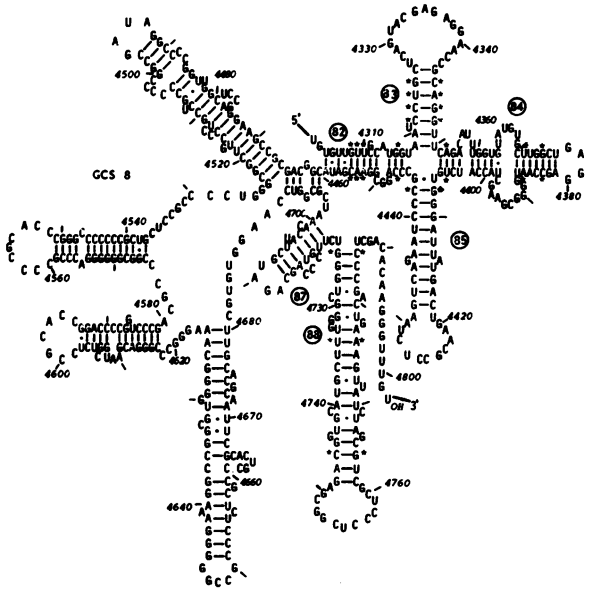

Figure 4-I. Domain VII of rat 28 S rRNA

proposed model are noteworthy.

Domain I (1-431) closely resembles the structures proposed for this region in the yeast (9) and the mouse (30) L-rRNA. The possibility for base-pairing with the 5'- and 3'-termini of 5.8 S rRNA is supported by the presence of numerous compensatory base changes (cbc) in both helices. In particular, a compensated A ⟶ G substitution at position 395 confirms the middle helix at the 5'-terminus of 5.8 S rRNA. The *E.coli* concensus helices 11, 12, 16 and 17 are reconfirmed by cbc, while helix 18 is strongly conserved. The region in the middle of domain I (115 to 276) is markedly enriched in rather long, non-conserved (G+C)-rich stretches easily folded into strongly base-paired structures.

Domain II (432-1554) contains a conserved part (Fig.4-B), corresponding to the homologous domain in yeast and GCS 1 (Fig.4-C). The conserved part includes concensus helices 21, 25 and 27 confirmed by several cbc and the conserved helices 24, 26 and 28. The putative counterpart of concensus helix 20 appears to allow stronger interactions than in *Saccharomyces*, but it is not confirmed by cbc. The large insertions present in this part of domain II of *Saccharomyces* (9) are not conserved and their structure in the rat (1245-1396) appears to be remodeled, in particular for the part corresponding to GCS 2 (1332-1388). Strictly, the large GCS 1 is located at the junction between domains I and II. It begins by the counterpart of concensus helix 19, confirmed by cbc. The GCS 1 proper contains about 760 nucleotides (see Table 1). Aside from the nucleotides forming the junctions to domain I via helix 19 and to the conserved part of domain II, the remaining 680 nucleotides in GCS 1 can be folded into three long helices (a, b and c in Fig.4-C) Base-pairing in GCS 1 helices is enhanced by the asymmetry in the distribution of G and C along the polynucleotide chain (Table 2). The size of the base-paired part of GCS 1 (680 nucleotides) matches almost perfectly the size of this structure evaluated to 670 nucleotides by electron microscopy of partly denatured rat 28 S rRNA (17). A shorter middle base-paired loop is also detected upon electron microscopy of this rRNA (18).

Domain III (1555-2113) includes (Fig.4-D) the strongly conserved concensus helices 29, 30, 32, 33, 36 and 37. Helices 34 and 35 are partly conserved and supported by several cbc. Helix 31 is present, but is markedly longer than in yeast. The distal conserved part is confirmed by two cbc. The proximal part contains GCS 3 (1581-1612) possibly base-paired to newly acquired nucleotide stretches. One cbc is present also at the origin of this helix. GCS 4 constitutes a strongly base-paired loop totally replacing helix

Table 2. Distribution of G and C in the two strands of the folded GCS 1 and GCS 6 of rat 28 S rRNA

Helix*	Strand	Number of G and C in strand**			Ratio G / C	Number of base-pairs in helix	
		Total	G	C		G - C	G · U
GCS 1a	u	110	72	38	1.90	83	12
	l	95	27	68	0.40		
GCS 1b	u	67	30	37	0.81	51	10
	l	63	42	21	2.00		
GCS 1c	u	84	47	37	1.27	71	8
	l	86	37	49	0.76		
GCS 6a	u	153	97	56	1.73	113	15
	l	127	42	85	0.49		
GCS 6b	u	84	42	42	1.00	59	10
	l	67	31	36	0.86		

* The designation of helices in GCS is the same as in Fig.4-C and 4-G
** All G and C are counted except in terminal loops or lateral helices
u and l denote upper and lower strand as shown in Fig.4-C and 4-G

38 and its yeast counterpart. In fact, this helix includes only part of GCS 4 (1963-2104, see Table 1), while the remaining part is in domain IV.

Domain IV (2114-2659) is moderately conserved (Fig.4-E). It contains three long non-conserved stretches including part of GCS 4 (2114-2169) and the whole GCS 5 (2479-2553). Only partial comparison with *Saccharomyces* is possible (see ref.9). Domain IV contains the strongly conserved helices 40 and 50, as well as helices 41, 42, 43, 44 and 47 supported by cbc. The counterparts of concensus helices 39 and 45 are totally remodeled (helix 39 includes part of GCS 4). GCS 5 and the preceeding sequence may form two long helical structures, but their relation to helices 48 and 49 is uncertain.

Domain V (2660-3616) contains a strongly conserved part (Fig.4-F) and GCS 6 (Fig.4-G). The strongly conserved helices 54, 55, 58, 59, 60, 61 and 62 are present, but not confirmed by cbc. Several cbc confirm concensus helices 51 and 56 as well as a helix (51a) proposed for *Saccharomyces* (9). and encompassing nucleotides 2692-2726 in the rat. The large insertion next to this helix in yeast is totally replaced by GCS 6 characterized by its very high G+C content (see Table 1). The majority of the nucleotides in GCS 6 can be folded into two double-stranded loops of unequal length (a and b in Fig.4-G) including 540 nucleotides. In this case too base-pairing is enhanced by the asymmetry in G and C distribution in helix a (see Table 2). Electron microscopy of partly denatured rat 28 S rRNA also reveals the presence of two

long loops unequal in length and encompassing about 560 nucleotides (17,18). A transition helix, analogous to helix 51a can be also formed at the junction of GCS 6 with the conserved part of domain V.

Domain VI (3617-4301) is strongly conserved for considerable parts of its sequence (Fig.4-H). Concensus helices 63, 64, 78, 80 and 81 are strongly conserved and devoid of cbc. Helices 65, 66, 68, 69, 70, 71, 72, 76, 77 and 79 are confirmed by several cbc. The helix encompassing nucleotides 4023 to 4061 appears to be better supported by cbc than the two concensus helices (74 and 75) proposed for *E.coli* and yeast (7,9), but further evaluation seems appropriate. Two non-conserved (G+C)-rich structures are present in domain VI. One corresponds to GCS 7 and replaces concensus helix 67 forming a large expansion, while the other (3818-3884) substitutes for a similar in length, but non-conserved, sequence in yeast (9).

Domain VII (4302-4802) is only moderately conserved (Fig.4-I). Concensus helices 82, 83, 84, 85 and 88 are confirmed by several cbc, although helix 88 seems to be markedly longer than its counterpart in yeast (9). Helix 87 is present, but in spite of several base substitutions no clear-cut cbc can be identified and its stability is altered. The large GCS 8 encompassing 225 nucleotides (see Table 1) substitutes for an "insertion" present at this site in yeast (9). The GCS 8 corresponds to the (G+C)-rich fragment identified previously at the 3'-terminus of mammalian 28 S rRNA by chemical-enzymatic methods (31). Folding into relatively strongly base-paired loops is possible, but remains to be ascertained by cbc.

The analysis of the primary and secondary structure of rat 28 S rRNA reveals the existence of a highly conserved backbone structure characteristic for all prokaryotic and eukaryotic L-rRNA. However, the appearance of large (G+C)-rich segments, folded in strongly base-paired structures, appears to be favoured in the evolution of higher eukaryotes. It should be added that the tendency for an enrichment in G+C does not seem to be restricted to the formation of (G+C)-rich segments. Thus, out of 48 unambiguously identified cbc in the comparison of yeast and rat L-rRNA, 67 % result in the substitution of A-U (or U-A) with G-C (or C-G) base-pairs. In summary, the enhanced stability of helical structures in vertebrate L-rRNA may be attributed to: (a) formation of expanded GCS at specific sites; (b) substitution of relatively long sequences by (G+C)-rich sequences of similar size, and (c) the enrichment in G-C pairs of the conserved helices. The biological role of the higher stability of vertebrate L-rRNA remains to be elucidated.

ACKNOWLEDGMENTS

The authors are indebted to Drs M.Kermekchiev and L.Antonov for their expert assistance in some experiments.

REFERENCES

1. Cox,R.A.(1977) Progress Biophys.Mol.Biol., 32, 193-231.
2. Gerbi,S.A., Gourse,R.L. and Clark,C.G.(1982) in The Cell Nucleus (Busch, H. and Rothblum,L.,eds) Vol.X, pp.351-386, Academic press, New York.
3. Brosius,J., Dull,T.J. and Noller,H.F.(1980) Proc.Nat.Acad.Sci.USA, 77, 201-207.
4. Branlant,C., Krol,A., Machatt,M.A., Ebel,J.P., Edwards,K. and Kossel,H. (1981) Nucleic Acids Res., 9, 4303-4324.
5. Glotz,C., Zwieb,C., Brimacombe,R., Edwards,K. and Kossel,H.(1981) Nucleic Acids Res., 9, 3287-3306.
6. Noller,H.F., Kop,J., Wheaton,V., Brosius,J., Gutell,R.R., Kopylov,A.M., Dohme,F., Herr,W., Stahl,D.A., Gupta,R. and Woese,C.R.(1981) Nucleic Acids Res., 9, 6167-6189.
7. Maly,P. and Brimacombe,R.(1983) Nucleic Acids Res., 11, 7263-7286.
8. Georgiev,O.I., Nikolaev,N., Hadjiolov,A.A., Skryabin,K.G., Zakharyev,V.M. and Bayev,A.A.(1981) Nucleic Acids Res., 9, 6953-6958.
9. Veldman,G.M., Klootwijk,J., Regt,V.C., Planta,R.J., Branlant,C., Krol,A. and Ebel,J.P.(1981) Nucleic Acids Res., 9, 6935-6952.
10. Otsuka,T., Noniyama,H., Yoshida,H., Kukita,T., Kuhara,S. and Sakaki,Y. (1983) Proc.Nat.Acad.Sci.USA, 80, 3163-3167.
11. Hadjiolov,A.A., Venkov,P.V. , Dolapchiev,L.B. and Genchev,D.D.(1967) Biochim.Biophys.Acta, 142, 111-127.
12. Hadjiolov,A.A. and Milchev,G.I.(1967) Compt.Rend.Acad.Sci.Bulgarie, 20 , 1333-1336.
13. Delihas,N.(1967) Biochemistry, 6, 3356-3362.
14. Cox,R.A., Huvos,P. and Godwin,E.A.(1973) Israel J.Chem., 11, 407-422.
15. Wellauer,P.K. and Dawid,I.B.(1973) Proc.Nat.Acad.Sci.USA, 70, 2827-2831.
16. Wellauer,P.K. and Dawid,I.B.(1975) Brookhaven Symp.Biol., 26, 214-223.
17. Schibler,U., Wyler,T. and Hagenbuchle,O.(1975) J.Mol.Biol., 94, 503-517.
18. Dabeva,M.D., Dudov,K.P., Hadjiolov,A.A., Emanuilov,I. and Todorov,B.N. (1976) Biochem.J., 160, 495-503.
19. Braga,E.A., Yussifov,T.N. and Nosikov,V.V.(1982) Gene, 20, 145-156.
20. Rothblum,L.I., Parker,D.L., Cassidy,B.(1982) Gene, 17, 75-77.
21. Berk,A.J. and Sharp,P.A.(1977) Cell, 12, 721-732.
22. Subrahmanyam,C.S., Cassidy,B., Busch,H. and Rothblum,L.I.(1982) Nucleic Acids Res., 10, 3667-3680.
23. Kominami,R., Mishima,Y., Urano,Y., Sakai,M. and Muramatsu,M. (1982) Nucleic Acids Res., 10, 1963-1979.
24. Maxam,A.M. and Gilbert,W. (1980) Methods Enzymol., 65, 499-580.
25. Walseth,T.F. and Johnson,R.A.(1979) Biochim.Biophys.Acta, 562, 11-31.
26. Peterman,M.L. and Pavlovec,A.(1963) J.Biol.Chem., 238, 3717-3724.
27. Cammarano,P., Pons,S., Londei,P. (1975) Acta Biol.Med.Germanica, 34, 1123-1135.
28. Fox,G. and Woese,C.R. (1975) Nature, 256, 505-507.
29. Brimacombe,R., Maly,P. and Zwieb,C.(1983) Progress Nucleic Acids Res.Mol. Biol., 28, 1-48.
30. Michot,B., Bachellerie,J.P. and Raynal,F.(1982) Nucleic Acids Res., 10, 5273-5283.
31. Hadjiolov,A.A., Cox,R.A. and Huvos,P.(1975) Biochem.J., 147, 625-628.

APPENDIX

After this paper was submitted for publication we recieved information on the work of Chan et al. [Nucleic Acids Res.,11 (1983) 7819-7831] reporting the complete sequence (4718 nucleotides) of rat 28 S rRNA. Comparison with our sequence shows that the two sequences are in good general agreement with each other showing differences at 90 sites, most of which (78) are single base substitutions, insertions or deletions. These may be due to differences between the two rat rDNA clones and/or to technical errors. In our rDNA clone, the presence of BstN I sites at positions (1st nucleotide of site) *3348*, *4445* and *4690*, and the absence of a Taq I site at position *2335*, is supported by the restriction analysis using these enzymes. A segment of 72 nucleotides (pos.*4548-4619*), located between two Ava I sites, is not present in the sequence published by Chan et al.. It is noteworthy that the presence of this segment yields a rat 28 S rRNA ∝-sarcin fragment of 466 nucleotides, more closely corresponding to the size (about 488 nucleotides) for this fragment derived earlier from gel electrophoresis analyses [Endo and Wool, J.Biol. Chem., 257 (1982) 9054-9060].