1374–1382 Nucleic Acids Research, 1994, Vol. 22, No. 8

# Probing the structure of mouse Ehrlich ascites cell 5.8S, 18S and 28S ribosomal RNA *in situ*

Lovisa Holmberg, Yvette Melander and Odd Nygård*
Department of Zoological Cell Biology, Arrhenius Laboratories E5, Stockholm University, S-106 91 Stockholm, Sweden

Received February 7, 1994; Revised and Accepted March 23, 1994

## ABSTRACT

**The secondary structure of mouse Ehrlich ascites 18S, 5.8S and 28S ribosomal RNA *in situ* was investigated by chemical modification using dimethyl sulphate and 1-cyclohexyl-3-(morpholinoethyl) carbodiimide metho-*p*-toluene sulphonate. These reagents specifically modify unpaired bases in the RNA. The reactive bases were localized by primer extension followed by gel electrophoresis. The three rRNA species were equally accessible for modification *i.e.* approximately 10% of the nucleotides were reactive. The experimental data support the theoretical secondary structure models proposed for 18S and 5.8/28S rRNA as almost all modified bases were located in putative single-strand regions of the rRNAs or in helical regions that could be expected to undergo dynamic breathing. However, deviations from the suggested models were found in both 18S and 28S rRNA. In 18S rRNA some putative helices in the 5'-domain were extensively modified by the single-strand specific reagents as was one of the suggested helices in domain III of 28S rRNA. Of the four eukaryote specific expansion segments present in mouse Ehrlich ascites cell 28S rRNA, segments I and III were only partly available for modification while segments II and IV showed average to high modification.**

## INTRODUCTION

The eukaryotic 18S and 28S ribosomal RNAs are considerably larger than their prokaryotic homologs, 16S and 23S rRNA. Despite the differences in length, the eukaryotic and prokaryotic rRNAs contain several regions with substantial sequence homologies. This has allowed construction of phylogenetic secondary structure models for both the 16S and the 23S-like rRNAs using comparative sequence analysis (1, 2). In the latest versions of the secondary structure models of eukaryotic 18S and 28S rRNAs, part of the additional eukaryote specific sequences are found in large expansion segments that have not been arranged into defined secondary structures (1, 2).

The predicted secondary structure models for prokaryotic 16S and 23S rRNA have gained support from a variety of experimental tests (for a review see 3) but little experimental work has been done on the structure of the eukaryotic rRNAs. Furthermore, the studies have so far focused on the secondary structures of the deproteinized rRNAs (4–8). The structural information from these studies shows substantial deviation from the predicted secondary structure models. However, in prokaryotes it has been observed that structural information collected *in situ* is more consistent with the predicted secondary structure models (9). Little is known about the structure of the eukaryotic rRNAs *in situ*. Hogan *et al.* (10, 11) studied the distribution of kethoxal reactive guanosine residues in yeast 40S and 60S subunits whereas in higher eukaryotes, the only detailed studies on the secondary structure of rRNA *in situ* have been performed on 5S rRNA and 5.8S rRNA as well as on domain V of *Xenopus laevis* 28S rRNA (7, 8, 12).

In this report we have studied the structure of mouse Ehrlich ascites cell 18S, 28S and 5.8S rRNA within derived 40S and 60S subunits, using the two single-strand specific reagents dimethyl sulphate (DMS) and 1-cyclohexyl-3-(morpholinoethyl) carbodiimide metho-*p*-toluene sulphonate (CMCT).

## MATERIALS AND METHODS

### Chemicals

Dimethyl sulphate and CMCT were from Aldrich Chemie (Germany). Deoxy and dideoxy nucleotides were from Boehringer Mannheim (Germany). [$\gamma$-$^{32}$P]ATP and T4 polynucleotide kinase were from Amersham International (UK). Superscript reverse transcriptase was from Life Technologies, Inc..

### cDNA primers

cDNA primers were synthesiszed as described by Caruthers *et al.* (13). The specific sequences used for primer annealing were for (i) 28S rRNA, $C^{155}$-$A^{169}$, $G^{268}$-$C^{282}$, $G^{403}$-$G^{417}$, $U^{503}$-$U^{517}$, $U^{607}$-$C^{621}$, $G^{745}$-$C^{759}$, $A^{855}$-$U^{869}$, $G^{1004}$-$G^{1018}$, $U^{1130}$-$C^{1144}$, $G^{1324}$-$G^{1338}$, $U^{1507}$-$C^{1521}$, $U^{1661}$-$A^{1675}$, $G^{1835}$-$U^{1849}$, $C^{1887}$-$C^{1891}$, $C^{2025}$-$A^{2039}$, $G^{2148}$-$U^{2162}$, $G^{2295}$-$G^{2309}$, $G^{2505}$-$U^{2519}$, $G^{2603}$-$U^{2617}$, $G^{2664}$-$C^{2678}$, $U^{2821}$-$U^{2835}$, $A^{2919}$-$C^{2933}$, $U^{3035}$-$U^{3049}$, $G^{3096}$-$U^{3110}$, $A^{3193}$-$U^{3207}$, $U^{3268}$-$G^{3282}$, $G^{3360}$-$A^{3374}$, $U^{3493}$-$C^{3507}$, $U^{3508}$-$A^{3522}$, $G^{3610}$-$G^{3624}$, $U^{3693}$-$U^{3707}$, $G^{3876}$-$A^{3890}$, $G^{4022}$-$A^{4036}$, $C^{4159}$-$G^{4173}$,

*To whom correspondence should be addressed

$A^{4275}$-$U^{4289}$, $G^{4312}$-$U^{4326}$, $A^{4379}$-$G^{4393}$ and $A^{4677}$-$A^{4691}$ (ii) 5.8S rRNA $G^{143}$-$U^{157}$ (iii) 18S rRNA, $G^{108}$-$G^{122}$, $U^{220}$-$A^{234}$, $U^{302}$-$U^{316}$, $G^{479}$-$C^{493}$, $U^{660}$-$A^{674}$, $A^{811}$-$U^{825}$, $U^{956}$-$U^{970}$, $C^{1080}$-$G^{1094}$, $G^{1257}$-$G^{1271}$, $U^{1405}$-$C^{1419}$, $C^{1598}$-$G^{1612}$ and $U^{1831}$-$U^{1845}$.

### Preparation of ribosomal subunits

Ribosomes were prepared from mouse Ehrlich ascites cells using the method described by Sundkvist and Staehelin (14).The pelleted ribosomes, mainly monosomes, were suspended in a buffer containing 0.5 M KCl, 20 mM Tris/HCl, pH 7.6, 3 mM $MgCl_2$ and 11 mM $\beta$-mercaptoethanol and the ribosomal subunits separated by gradient centrifugation as described by Nygård and Nika (15). The isolated subunits were pelleted by centrifugation and suspended in 0.25 M sucrose, 70 mM KCl, 30 mM Hepes/KOH, pH 7.6, 2 mM $Mg(CH_3COO)_2$ and 5 mM $\beta$-mercaptoethanol to a concentration of 6 $\mu$M. The subunits were frozen and stored at −80°C until used.

### Modification of rRNA

Chemical modification was performed as previously described (12). The reagents DMS and CMCT were added to the samples containing derived subunits at a final concentration of 20 or 90 $\mu$M for DMS and 20 or 100 mM for CMCT. The samples, final volume 400 $\mu$l, were incubated, for 15 min in the case of CMC-T and for 5 min in the case of DMS, at 37°C to allow statistical modification of the rRNA (16). Control samples were incubated in the absence of DMS or CMCT but were otherwise treated exactly as the modified samples. Dimethyl sulphate modifies single-strand adenines and cytosines whereas CMCT modifies unpaired uridines and guanines (16). In addition, CMCT, at pH 7.6, reacted with cytosines.

The rRNA was extracted from the ribosomes with phenol according to Brawerman *et al.* (17), dissolved in $H_2O$ and the RNA-concentration adjusted to 1 pmol/$\mu$l. The material was stored in small aliquotes at −80°C.

### Identification of modification sites

The cDNA primers were end-labelled using [$\gamma$-$^{32}$P]ATP and purified as previously described (12). For primer extension (18), 1 pmol of labelled probe was annealed to 1 pmol control or modified rRNA in 50 mM Hepes/KOH, pH 7.6, 122 mM KCl. After incubation for 3 min at 90°C, the samples were cooled to 45°C, reverse transcriptase, 50 units, together with the four dNTPs, final concentrations 0.5 mM, were added and the samples incubated for 15 min at 45°C. Sequencing of the rRNA was as previously described (12). The sequencing and primer extension products were analyzed by electrophoresis on gels containing 8% (mass by volume) acrylamide and 0.4% (mass by volume) bisacrylamide (12).

## RESULTS AND DISCUSSION

We have analyzed the accessibility of rRNA for chemical modification in derived 40S and 60S subunits isolated from mouse Ehrlich ascites cells. These subunits were considered functional based on the following criteria; i) the derived subunits bind initiation factors and Met-$tRNA_f$ (19) ii) the subunits efficiently associate into 80S ribosomes and the re-associated ribosomes bind elongation factor 2 in the presence of GTP and GTP analogues and hydrolyse GTP in an eEF-2 dependent manner (20) iii) the 80S reassociated ribosomes translate globin mRNA in the presence of elongation and initiation factors (21).

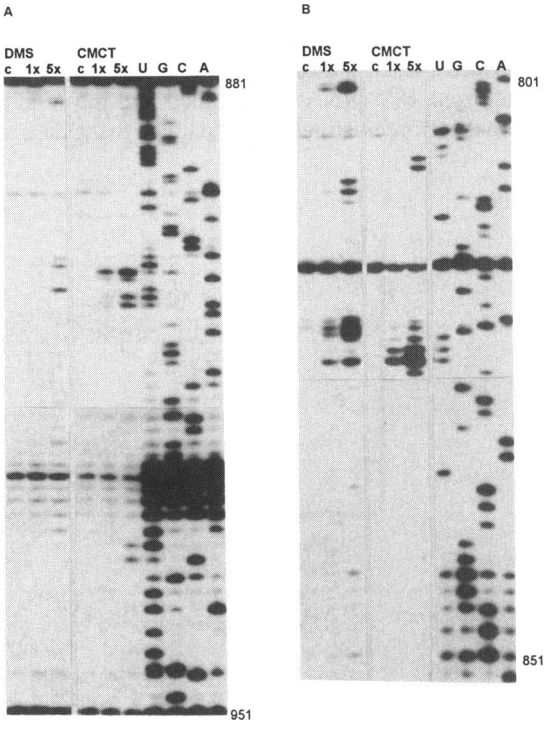



**Figure 1.** Autoradiographs showing the chemical accessibility of 18S and 28S rRNA isolated from derived 40S and 60S subunits, respectively. The chemical accessibility was analyzed using the single strand specific reagents DMS and CMC-T. The concentrations used were 20 and 90 $\mu$M for DMS and 20 and 100 mM for CMCT corresponding to 1× and 5×, respectively. Control samples (c) were incubated in the absence of modifying reagent. **A.** Nucleotides 881−951 in 18S rRNA. **B.** Nucleotides 801−850 in expansion segment I of 28S rRNA. The lanes U, G, C and A were used for determining the sequence.

The chemical reagents used in this study, CMCT and DMS, specifically modify single-strand nucleotides and are frequently used in analysing the secondary structure of RNA molecules (16, 18, 22). The position of the modified nucleotides in the rRNA was localized by primer extension followed by gel-electrophoresis and autoradiography (typical gels are shown in fig. 1). The accessible sites were denoted as highly reactive, if they were modified at the lowest concentration of the reagent used, or moderately reactive, if they were only modified at the highest concentration. Seventeen nucleotides were denoted as hypersensitive *i.e.* they were extensively modified by the reagents at the lowest concentrations used.

During the experiments we observed that some of the bases in the rRNA functioned as natural stops for the reverse transcriptase (Figs. 2 and 3, table 1). The accessibility of these bases for modification could therefore not be evaluated. Many of the natural stops coincided with the position of 2′-O-methylated nucleotides (23) although several known 2′-O-methylated bases were efficiently copied by the enzyme. Natural stops also occurred at positions containing other modified nucleotides such as pseudouridine and 3-methyluridine.

### Structure of 18S ribosomal RNA in isolated 40S subunits

Mouse 18S rRNA consists of 1869 nucleotides (24) and is divided into three separate domains (Fig. 2). A general analysis of the

secondary structure of the entire mammalian 18S rRNA and a more detailed study on the structure of its 5'-domain have recently been published (4, 5). However, in these reports the structure of deproteinized 18S rRNA was studied and the modification patterns are therefore not directly comparable to the data reported here. As already noted in prokaryotes (9), the rRNA is as a rule less accessible for modification in the intact ribosomal particle than in the stripped RNA. Thus, the number of modification sites reported for the naked 18S rRNA is much higher than that seen in the ribosomal subunit but the majority of the modification sites

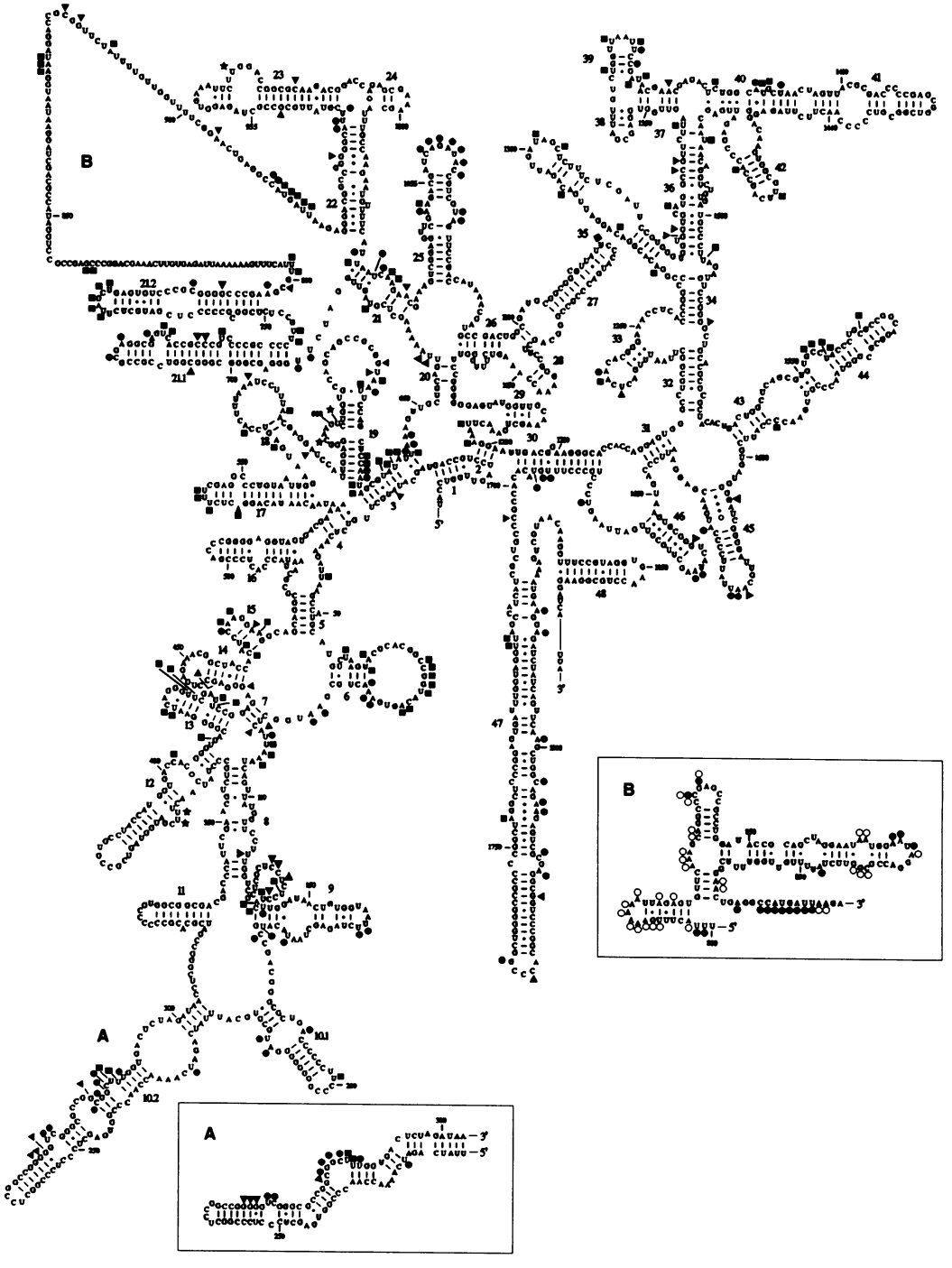
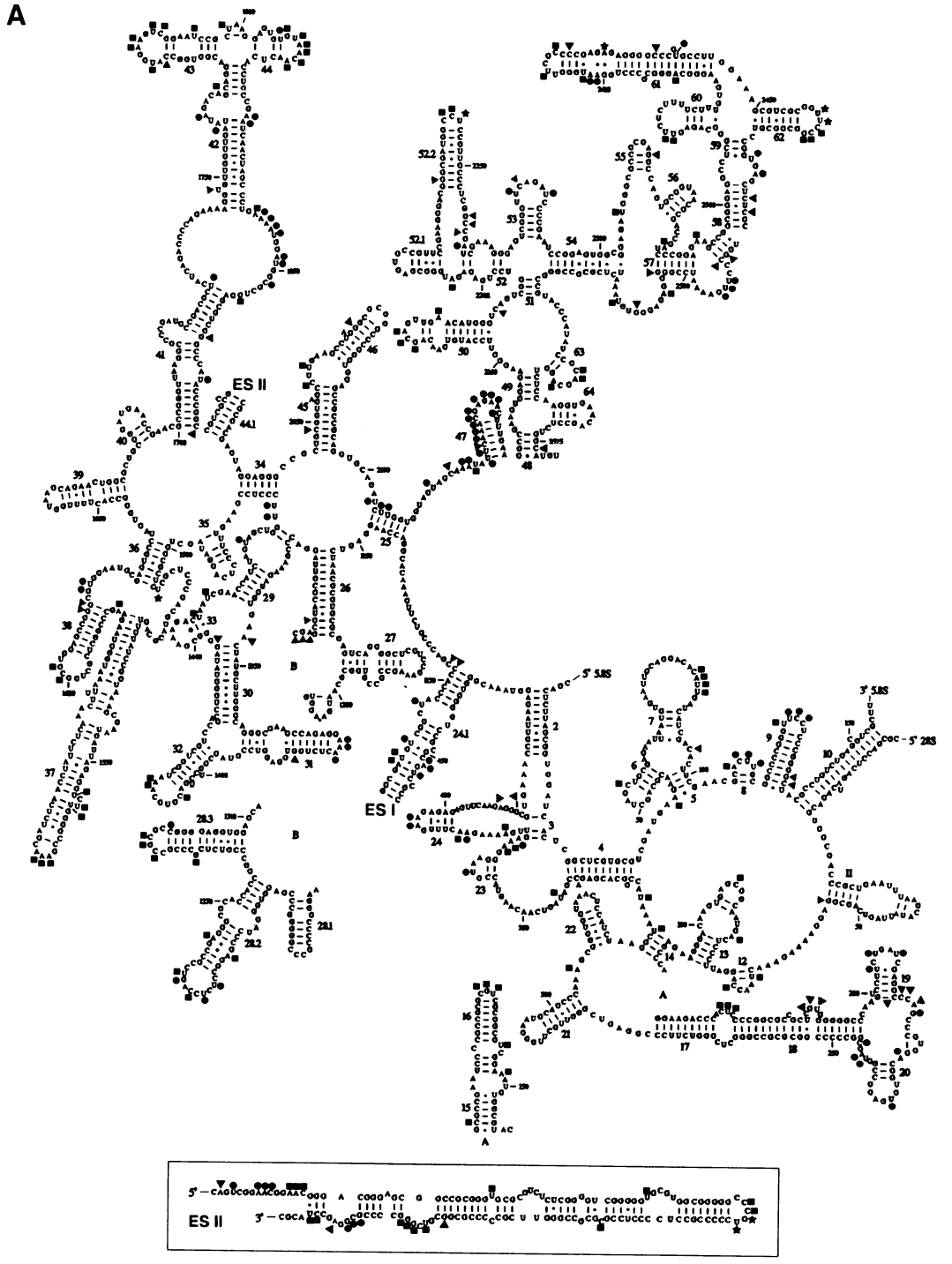



**Figure 2.** Secondary structure model of mouse 18S rRNA showing the distribution of bases accessible for chemical modification in derived 40S subunits. The structure model is from Gutell *et al.* (1) and the helix-numbering is according to Neefs *et al.* (51). The extent of modification of the bases is indicated as; moderately reactive (●), highly reactive (■) and hyper-sensitive (★). Bases that could not be analyzed due to natural stops are marked by (▲). Insert A. Alternative secondary structure model for helix 10.2. Insert B. Tentative secondary structure model for the expansion segment. The model was constructed based on the chemical accessibility of 18S rRNA within the derived 40S (52). Bases accessible to modification by DMS or CMCT are indicated as (●). The modifications found in naked 18S rRNA from rabbit (4) are indicated as (○).

found in the RNA *in situ* are also found in the deproteinized RNA. However, we found a few sites that were apparently only accessible in the intact 40S subunit. Similar observations made in *E.coli* are suggested to reflect conformational changes in the rRNA caused by assembly of the subunit (9).

*The 5'-domain.* The first 658 nucleotides of 18S rRNA constitute the 5'-domain (Fig. 2). The domain contained 86 nucleotides that were available for modification by DMS or CMCT and was thus the most extensively modified domain in 18S rRNA. Sixtytwo of these nucleotides were located in sequences that, according to the theoretical secondary structure model, were single-strand sequences whereas the remaining 24 modified nucleotides were engaged in putative helices (Fig. 2). However, 17 of the latter sites involved bases that were located in structures generally referred to as dynamic unstable structures (8) *i.e.* at a helix terminus, next to a non-canonic base-pair or a bulged nucleotide or involved in a non-canonic base pair. The remaining seven nucleotides were found in putative helices (Fig. 2). Thus, major differences between the obtained single-strand specific modification pattern and the theoretical secondary structure model were found. Extensive modification of helical bases were



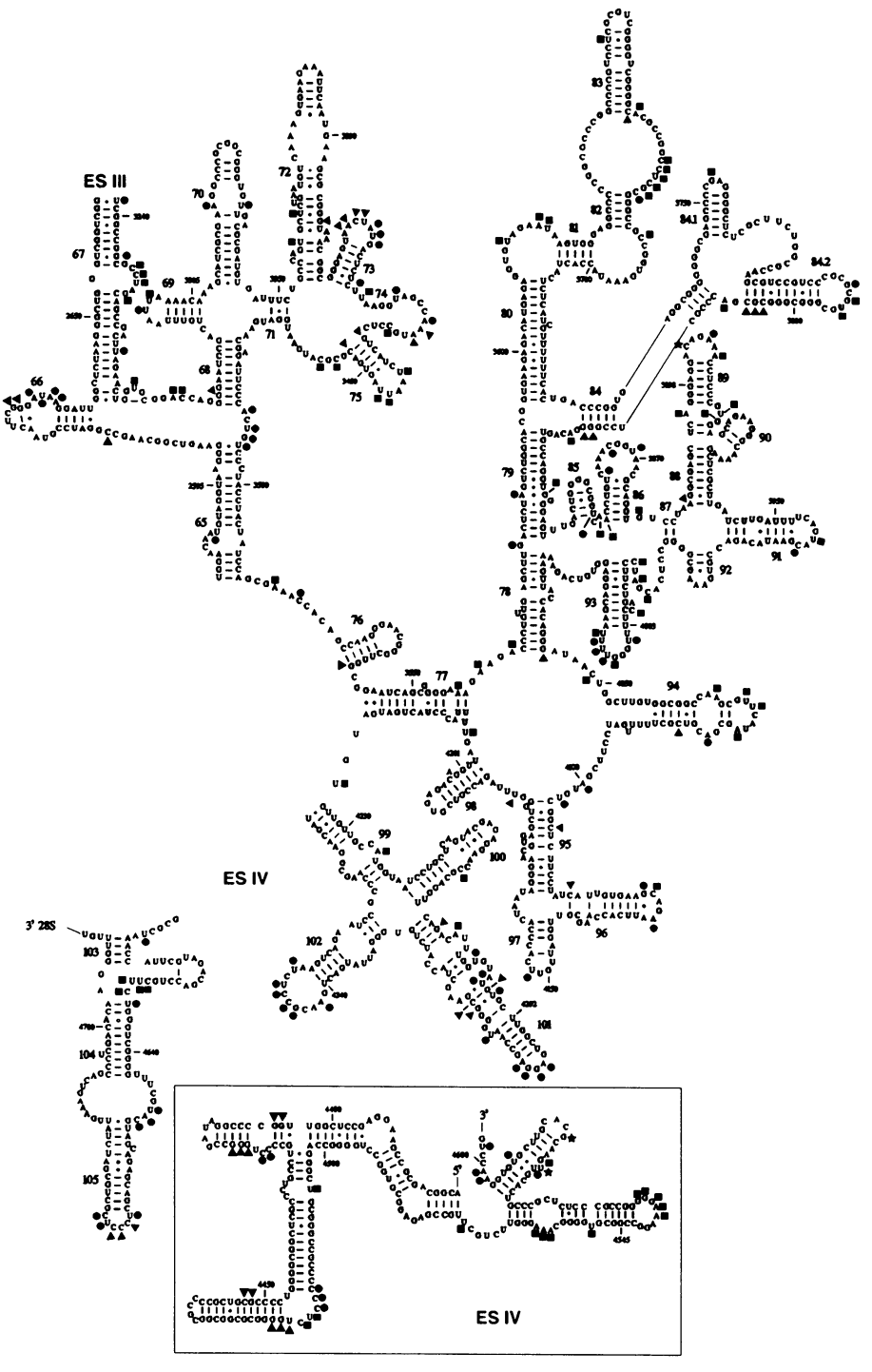



**Figure 3. A.** Secondary structure model for the 5'-half of mouse 5.8S/28S rRNA showing location of bases accessible for chemical modification in 60S subunits. The structure model is from Gutell *et al.* (1). The position of the eukaryote specific expansion segment is indicated by ES. For details see legend to figure 2. Insert. Tentative secondary structure models for expansion segment II. **B.** Secondary structure model for the 3'-half of mouse 28S rRNA showing location of bases accessible for chemical modification in 60S subunits. See figure 3A for details. Insert. Tentative secondary structure model for expansion segment IV.

observed in the adjacent helices 19 and 3 as well as in helix 10.2. In the first region, 14 out of the 20 modified nucleotides were involved in putative Watson–Crick base pairs. A similar modification has also been reported in deproteinized 18S rRNA, although modification seems to be more pronounced *in situ* (4, 5). These observations suggest that this helical region, if present, must be partly destabilized to allow modification by the single-strand specific reagents. Interestingly, the homologous region in prokaryotic 16S rRNA is totally protected against modification in the assembled subunit and only marginally hit by single-strand specific chemical reagents in its deproteinized form (9). In helix 10.2 the modified helical bases were part of a cluster containing seven reactive bases. This part of the sequence also contained an additional uracil that was not found in mouse 18S rRNA (24).

**Table 1.** Modified bases in the expansion segments of mouse Ehrlich ascites cell 28S rRNA

| Expansion segment | Position | Reactivity | Position | Reactivity | Position | Reactivity |
|---|---|---|---|---|---|---|
| I | $U^{467}$ | ++ | $U^{468}$ | + | $A^{470}$ | + |
| (nucleotides | $A^{471}$ | + | $U^{486}$ | + | $C^{504}$ | ns |
| 460–1102) | $C^{505}$ | ns | $U^{517}$ | + | $U^{533}$ | + |
| | $A^{546}$ | + | $U^{566}$ | + | $U^{567}$ | + |
| | $C^{568}$ | ++ | $C^{571}$ | + | $G^{582}$ | ns |
| | $C^{583}$ | + | $G^{687}$ | ns | $G^{688}$ | ns |
| | $A^{692}$ | ++ | $U^{694}$ | ++ | $U^{702}$ | ++ |
| | $U^{708}$ | + | $G^{709}$ | + | $C^{734}$ | ++ |
| | $G^{735}$ | + | $G^{745}$ | ++ | $U^{747}$ | ++ |
| | $U^{779}$ | ns | $C^{780}$ | ns | $A^{781}$ | ++ |
| | $G^{782}$ | ++ | $A^{792}$ | ++ | $A^{793}$ | ++ |
| | $C^{794}$ | ns | $C^{795}$ | ns | $A^{796}$ | ++ |
| | $A^{801}$ | ++ | $C^{802}$ | ++ | $U^{811}$ | ++ |
| | $U^{812}$ | ++ | $A^{813}$ | ++ | $C^{814}$ | ++ |
| | $G^{823}$ | ns | $C^{827}$ | ++ | $A^{828}$ | ++ |
| | $C^{829}$ | ++ | $G^{830}$ | ++ | $U^{831}$ | ++ |
| | $U^{832}$ | ++ | $U^{833}$ | + | $C^{901}$ | ns |
| | $C^{902}$ | ns | $C^{903}$ | ns | $G^{908}$ | ns |
| | $G^{909}$ | ns | $G^{910}$ | ns | $U^{913}$ | + |
| | $G^{915}$ | ns | $G^{916}$ | ns | $G^{917}$ | ns |
| | $G^{921}$ | ns | $G^{922}$ | ns | $G^{923}$ | ns |
| | $G^{932}$ | + | $U^{956}$ | + | $C^{957}$ | + |
| | $C^{958}$ | ns | $U^{971}$ | + | $C^{972}$ | + |
| | $C^{977}$ | ns | $U^{978}$ | ++ | $U^{980}$ | ++ |
| | $C^{981}$ | + | $G^{985}$ | ns | $C^{986}$ | ++ |
| | $U^{1046}$ | ++ | $U^{1071}$ | ++ | $U^{1073}$ | ++ |
| | $C^{1074}$ | ++ | $C^{1075}$ | ++ | $C^{1076}$ | ++ |
| | $C^{1084}$ | ++ | $C^{1085}$ | ++ | $C^{1091}$ | ns |
| | $G^{1097}$ | ns | $U^{1098}$ | ++ | $G^{1100}$ | ++ |
| | $C^{1102}$ | ns | | | | |
| III | | | | | | |
| (nucleotides | $A^{2675}$ | ++ | $A^{2676}$ | ++ | $U^{2684}$ | + |
| 2664–3237) | $C^{2710}$ | + | $C^{2728}$ | + | $G^{2734}$ | ns |
| | $A^{2737}$ | + | $G^{2738}$ | ++ | $A^{2739}$ | + |
| | $U^{2751}$ | ++ | $U^{2752}$ | ++ | $C^{2756}$ | + |
| | $C^{2757}$ | + | $C^{2758}$ | ++ | $C^{2788}$ | + |
| | $A^{2895}$ | ns | $U^{2907}$ | + | $U^{2908}$ | ++ |
| | $C^{2924}$ | ns | $A^{2926}$ | + | $C^{2934}$ | + |
| | $G^{2948}$ | ns | $A^{2957}$ | + | $U^{2959}$ | ns |
| | $C^{2968}$ | + | $G^{2981}$ | ns | $A^{2982}$ | +++ |
| | $C^{2983}$ | +++ | $A^{2989}$ | + | $G^{2994}$ | ++ |
| | $U^{3019}$ | + | $C^{3020}$ | + | $A^{3021}$ | + |
| | $A^{3053}$ | ++ | $U^{3064}$ | + | $C^{3070}$ | ns |
| | $G^{3075}$ | ns | $C^{3077}$ | ns | $C^{3145}$ | ns |
| | $C^{3157}$ | ns | $U^{3159}$ | + | $G^{3160}$ | + |
| | $U^{3161}$ | + | $C^{3163}$ | + | $G^{3170}$ | ns |
| | $U^{3178}$ | ++ | $U^{3184}$ | ++ | $U^{3185}$ | ++ |
| | $C^{3210}$ | ns | $U^{3218}$ | ++ | $U^{3219}$ | ++ |
| | C3221 | + | G3222 | + | G3235 | ++ |

The bases are denoted as moderately (+) or highly (++) reactive. Natural stops (ns).

Thus, in ascites cells the sequence between nucleotides 284–291 was identical to that found in rabbit, HeLa cells and *Xenopus laevis* (5, 25, 26) and allowed an alternative arrangement of the secondary structure for helix 10.2 (Fig. 2). The modification data reported here were in good agreement with this alternative model.

Based on experiments performed with deproteinized 18S rRNA, the apical loop of helix 6 has been suggested to participate in tertiary interactions with the terminal loop of helix 12 (5). Our *in situ* data show that the apical loop of helix 12 is completely protected from modification and contains a sequence of six nucleotides, CCGUGG, that are complementary to the six non-exposed nucleotides, GCACGG, in the loop of helix 6 (Fig. 2). Thus, the data reported here do not exclude a possible interaction between these two loops *in situ*. However, it should be pointed out that Hogan *et al.*, (10) observed extensive modification *in situ* of the yeast rRNA structure homologous to the apical loop of helix 12. Furthermore, in prokaryotes the sequence corresponding to the loop of helix 6 does not interact with RNA or protein (27).

The universally conserved 16 memberred loop of helix 19 (Fig. 2) showed limited accessibility for modification both *in situ* and in naked 18S rRNA (4, 5). The modification pattern was similar to that observed in prokaryotic 16S rRNA (9). The sequences $G^{603}GC^{605}$ and $C^{623}CG^{625}$ located in the bulge of the stem and in the apical loop of helix 19, respectively, have been suggested as partners in the formation of a pseudoknot (28).However, in the 40S particle the base $C^{605}$ was hypersensitive to single-strand RNA specific modification, suggesting that the pseudoknot, if present, must undergo dynamic changes that allow modification by the reagents.

*Central domain.* Nucleotides 659 to 1203 constitute the central domain of 18S rRNA (Fig. 2). The domain accommodates a 127 nucleotides long expansion segment that is unstructured in the model in fig. 2. A total of 42 modifications were detected within the modelled regions. All sites were found in putative single-strand sequences or in dynamic unstable structures. Thus, the modification pattern of the central domain was consistent with the theoretical secondary structure model (Fig. 2). Most of the modification sites found in the central domain of 18S rRNA *in situ* were also found in deproteinized 18S rRNA [Fig. 2 and reference (4)]. However, the region containing helices 23 and 24 and the inter helical sequence connecting helices 25 and 26 were almost exclusively modified in the deproteinized 18S rRNA. In *E.coli* the helices homologous to helices 23 and 25 are juxtaposed as seen by intra-RNA cross-linking (29). This part of the 16S rRNA contains the binding sites for proteins S11 and S15 as deduced from cross-linking and footprinting experiments (29, 30). The *E.coli* protein S11 is homologous to protein S14 from rat (31), suggesting that S14 could be involved in protecting the central domain of 18S rRNA against chemical modification within the 40S subunit.

The central domain contains a nucleotide sequence that has not been modelled in the secondary structure model in fig. 2. However, a model has been suggested for the identical sequence found in rabbit 18S rRNA (4). Using this model, we observed that the majority of the 15 accessible bases found in ascites 18S rRNA were located in helical regions. Therefore we constructed an alternative model based on the *in situ* modification pattern (Fig. 2). This model was found to be in reasonable agreement with the modification pattern found in deproteinized RNA (4) although some bases in putative stable canonic base pairs were available for modification. However, the discrepancy is approximately similar to that observed between the average modification patterns obtained in naked RNA and the phylogenetic secondary structure models.

*The 3'-domain.* The 3'-domain of Ehrlich ascites 18S rRNA is composed of nucleotides 1204 to 1869 (Fig. 2). The domain contained 52 nucleotides that were accessible for modification. Thus, this domain was the least available for modification of the domains in 18S rRNA. Most of the modified nucleotides were located in putative single-strand sequences but 13 bases were found in dynamic unstable structures of the RNA. Three modified bases were found in the putative stable helices 40, 42 and 44. In helices 40 and 44 the modified helical bases were part of larger clusters of modified bases, indicating that these bases were part of generally exposed sequences.

The sites accessible for modification in the 3'-domain of 18S rRNA *in situ* were also found to be exposed in deproteinized 18S rRNA with the exception that the apical part of hairpin 47 was more modified in the ribosome than in naked 18S rRNA (4). The region containing helix 41 and its apical loop was totally protected against modification in the 40S subunit (Fig. 2) but extensively modified in naked rRNA. In prokaryotes the homologous helix is cross-linked to proteins S9 and S10 and the region also contains footprinting sites for protein S9 (29, 30). The two *E.coli* proteins S9 and S10 are homologous to rat ribosomal proteins S16 and S20, respectively (32, 33). Thus, the protection of the region seen in 18S rRNA *in situ* could originate from an interaction of these two proteins with the 18S rRNA.

## Structure of 5.8S and 28S ribosomal RNA within 60S subunits

Mouse 28S rRNA consists of 4712 bases (34). The phylogenetic secondary structure model of 28S rRNA (1, 2) divides the rRNA into six separate domains (Fig. 3) in which approximately 66% of the total nucleotides are included. The remaining bases are found in four large expansion segments. Due to the limited experimental data available on the structure of eukaryotic 28S rRNA a comparison with existing data can only be made for the 5.8S rRNA part of domain I and for domain V.

*Domain I.* Eukaryotic 5.8S rRNA is considered to be the phylogenetic result of expanding the 23S rRNA and separating the 5'-region from the main part of the rRNA (35, 36). Thus, in eukaryotes 5.8S rRNA and the 5'-end of the 28S rRNA, nucleotides 1 to 430, form a hybrid structure that is homologous to domain I in prokaryotic 23S rRNA. In 5.8S rRNA 14 modifications were detected. These bases were positioned in single-strand sequences of the rRNA or in dynamic unstable structures.

The modification pattern generated by CMCT and DMS, was in agreement with previous observations although we find considerably fewer exposed nucleotides in mouse 60S subunits than previously reported for the rat 60S particle (7). The discrepancy between the two results may originate from differences in the ionic compositions used during subunit preparation and modification. This assumption is substantiated by a recent report in which the modification pattern of 5.8S rRNA is shown to vary with the magnesium concentration (37).

The 28S rRNA part of domain I contained 35 bases that were accessible for modification by DMS and CMCT (Fig. 3a). Thirty-one of the modified bases were found in predicted single-strand sequences whereas the remaining modified bases were located to dynamic unstable regions. Hence, the experimental data was in agreement with the suggested secondary structure model (2). The limited chemical reactivity of the bases in domain I of 28S rRNA as well as the generated modification pattern was similar to that reported for domain I of prokaryotic 23S rRNA *in situ* (38, 39).

*Domain II.* Nucleotides 1143 to 2110 constitute domain II (Fig. 3a). The domain comprises a long eukaryote specific sequence (expansion segment II), nucleotides 1887–2025, that is not included in the phylogenetic secondary structure model. A total of 70 modified bases were detected within the modelled regions of the domain. Fifty-nine of these sites were located in single-strand sequences of the rRNA whereas nine sites were found in dynamic unstable regions. The remaining two modified nucleotides were located internally in helix 25.

The helix-loop arrangement containing helices 43 and 44 (Fig. 3a) was one of the most accessible regions found in 28S rRNA. This region contained a total of 13 highly reactive and one moderately reactive base. An additional five reactive bases were found in the internal loop of the adjacent helix 42. The homologous region in prokaryotes is also very accessible to modification (40) although the ribosomal protein L11 and the pentameric protein complex L10.$(L12)_4$ bind to helices 43 and 44 and to helix 42, respectively (41). However, binding of L10.$(L12)_4$ to the rRNA also increases the exposure of several bases in the internal loop of helix 42. Interestingly several of these sites were homologous to the nucleotides that were reactive in the ascites 60S subunit. Furthermore, the available data for the yeast 28S type of rRNA confirm that the apical loop of helix

37 and the region homologous to helices 43 and 44 (Fig. 3a) are accessible for modification in the intact subunit (11).

*Domain III.* Domain III is composed of nucleotides 2125 to 2580 (Fig. 3a). During sequencing we noted that nucleotide $C^{2418}$ found in mouse 28S rRNA (34) was replaced by $U^{2418}$ in mouse Ehrlich ascites cells. The domain contained 55 DMS and CMC-T reactive nucleotides including four hyper sensitive bases. Forty-six bases were members of putative single-strand elements and two sites involved bases engaged in non-canonic base pairs. The remaining modification sites involved Watson–Crick base pairs located in helices 47 and 61. In the latter helix both nucleotides in a suggested base-pair were available for modification by the single-strand specific reagents. The modified bases in helix 47 were part of a larger cluster in which 15 out of 17 bases were accessible to CMCT and DMS modification. All modified bases showed limited reactivity suggesting that the hairpin may be flexible and only exist in a subpopulation of the 60S ribosomal subunits.

Although domain III was the most exposed domain in 28S rRNA, the region centered around helix 51 was completely protected from modification. The homologous region in 23S rRNA is also protected from modification in the assembled subunit due to a direct interaction with ribosomal protein L23 (42).

*Domain IV.* Domain IV, nucleotides 2581 to 3511, contains some of the most conserved sequences in the 28S rRNA (43). In addition, the domain contains a long eukaryote specific sequence (expansion segment III), nucleotides 2664–3237, that has not been included in the secondary structure model (Fig. 3b). Our analysis showed that the modelled regions contained 40 bases that were accessible for DMS and CMCT modification (Fig. 3b). Thirtynine of the bases were positioned in single-strand regions or in dynamic unstable structures. The remaining base was located in a putative stable base-pair in helix 67.

The observed modification pattern was similar to that seen in prokaryotic 50S subunits (39) with the exception that hairpins 72 to 75 were more available for modification in 60S subunits than in the 50S particles. In the latter particles the homologous region is largely protected from modification by ribosomal protein L2 (44). Apparently, no similar protection is exerted by any of the eukaryotic ribosomal proteins.

*Domain V.* Domain five of 28S rRNA, nucleotides 3544 to 4222, is generally considered to contain the peptidyl transferase centre of the ribosome (3). In this domain we observed that nucleotides $C^{3834}$ and $C^{3842}$, found in mouse 28S rRNA, were replaced by $G^{3834}$ and $G^{3842}$ in Ehrlich ascites cells. A total of 62 DMS and CMCT reactive bases were detected in domain V (Fig. 3b). All accessible sites were located in putative single-strand sequences or in dynamic unstable structures. Thus, the chemical modification pattern of domain V was in agreement with the theoretical secondary structure model in fig. 3b.

The chemical accessibility of domain V in *Xenopus laevis* 28S rRNA has recently been investigated using the single-strand specific reagents DMS, CMCT and kethoxal (8). The obtained modification pattern is similar to that reported here for mouse 28S rRNA. Reactivity differences were, however, seen in the apical loops of helices 92 and 93. The former sequence was protected from modification in mouse 60S subunits but accessible in *X.laevis*, while the loop of hairpin 93 was reactive in mouse but protected in Xenopus (8).

*Domain VI.* Domain VI, nucleotides 4223 to 4712, contains an eukaryote specific sequence (expansion segment IV), nucleotides 4379 to 4603, that is not included in the secondary structure model. The modelled part of the domain showed a total of 31 modification sites located in suggested single-strand or dynamic unstable sequences (Fig. 3b). Consequently, the experimental data support the theoretical secondary structure model for domain VI.

The domain contains the so called $\alpha$-sarcin/ricin loop at the apex of helix 100 (45). This part of the domain is involved in protein synthesis elongation and is assumed to bind the two elongation factors (46, 47). Despite the fact that the loop is accessible for modification by the N-glycosidase ricin, the region only contained one site that was available for chemical modification in 60S subunits from mouse Ehrlich ascites cells (Fig. 3b) and in yeast (11). These results are in clear contrast to the strong modifications seen in the $\alpha$-sarcin/ricin region in prokaryotic 50S particles (48).

In prokaryotes chemical cross-linking and footprinting studies have shown that ribosomal protein L3 interacts with helix 99 and the hinge between helices 103 and 104 (48, 49). Protein L3 has recently been shown to be homologous to mammalian L3 (50). Apparently, mammalian L3 did not protect the homologous sequences in mouse 28S rRNA as this part of the rRNA was accessible for modification in the 60S subunits.

*The expansion segments.* Approximately one third of the bases in 28S rRNA are found in non conserved eukaryote specific expansion segments. The first expansion segment, found between domains I and II, contains helix 24.1 and a large sequence, nucleotides 460 to 1102, which has not been included in the secondary structure model. The putative helix 24.1 contained seven bases that were available for modification. Six of the bases were distributed on both sides of a short four base pair helical segment containing two non-canonic base pairs. The extensive modification of the helix indicates that the secondary structure of this part of the 28S rRNA differs from that suggested in the secondary structure model in fig. 3a.

Expansion segment I was partly difficult to analyze due to long runs of guanines and cytosines. These sequences also resulted in a relatively high number of natural stops for the reverse transcriptase (table 1). During sequencing we observed that the sequence derived from the mouse Ehrlich ascites 28S rRNA differed from that found in mouse 28S rRNA (34). The short sequence between nucleotides $A^{813}$ and $U^{833}$ showed a series of deletions, insertions and base substitutions giving the following sequence: $A^{813}$CAGCUCCGGGCGCACGUUU$^{833}$. This sequence is one nucleotide shorter than the corresponding sequence in mouse 28S rRNA. We have however, chosen not to adjust the numbering of the nucleotides in fig. 3 and table 1 as we cannot exclude the possibility of a compensatory insertion in the long runs of alternating C and G:s found in the expansion segment.

Altogether 60 modified nucleotides were detected in the expansion segment. The accessible nucleotides were not evenly distributed as long runs of the rRNA such as the sequence between nucleotides $G^{584}$ and $C^{691}$ were totally protected from modification. The inaccessibility of the nucleotides and the high content of guanines and cytosines, 97 out of 108 bases, in this sequence suggest that it forms a stable helix in the 60S subunit.

The accessible regions of the expansion segment consisted of larger clusters of reactive bases. Three of the clusters contained continuous rows of four or more very accessible bases (Table 1), suggesting that these bases formed loops in the 28S rRNA.

Domain II in 28S rRNA contains a 140 nucleotides long expansion segment (expansion segment II). In this segment we observed that nucleotide $C^{2024}$, found in mouse 28S rRNA (34), was substituted for by a guanosine in mouse Ehrlich ascites cells. The expansion segment contained 23 bases that were accessible for DMS and CMCT modification *i.e.* 17% of the bases. Thus, expansion segment II was extremely accessible for modification. Nineteen of the reactive nucleotides occurred in groups consisting of two or more modified bases. One of the clusters contained two highly reactive and two hypersensitive nucleotides. The modification data are summarized in the tentative secondary structure model in fig. 3a (insert). As seen, the tentative model is in good agreement with the experimental data reported here.

Expansion segment III, nucleotides 2664 to 3237, located in domain IV showed little exposure to chemical modification and long stretches of the rRNA were completely protected from modification by the reagents used (Table 1). Of the 41 accessible nucleotides, 26 were found in clusters of two or more reactive nucleotides. Despite the overall low reactivity of the segment, two hypersensitive nucleotides were detected.

Domain VI contains expansion segment IV composed of 225 nucleotides. This expansion segment contained 26 bases, including two hypersensitive nucleotides, that were accessible to modification by CMCT and DMS. Two very modified clusters with three and five reactive bases, respectively were observed in the segment. The modification pattern is shown in the probationary secondary structure model in fig. 3b (insert). This model is in good agreement with our experimental data.

## ACKNOWLEDGEMENTS

We are grateful to Dr R.R.Gutell for making the latest versions of the theoretical secondary structure models of mouse ribosomal RNA available to us prior to publication. We thank Dr B.-M.Sjöberg for construction of secondary structure models and Dr L.Nilsson for discussion and careful reading of the manuscript. We are indebted to Mrs B.Lundberg for technical assistance. This work was supported by the Swedish Natural Science Research Council grant B-BU 8463-307.

## REFERENCES

1. Gutell, R.R., Larsen, N. and Woese, C.R. (1993) In *Ribosomal RNA: Structure, Evolution, Gene Expression and Function in Protein Synthesis* (Zimmermann, R.A. and Dahlberg, A.E eds) CRC Press, Baca Raton, FL., in press.
2. Gutell, R.R., Schnare, M.N. and Gray, M.W. (1992) *Nucl. Acids Res. 20* (Supplement), 2095–2109.
3. Noller, H.F. (1991) *Annu. Rev. Biochem.* **60**, 191–227.
4. Rairkar, A., Rubino, H.M. and Lockard, R.E. (1988) *Biochemistry* **27**, 582–592.
5. Mandiyan, V. and Boublik, M. (1990) *Nucl. Acids Res.* **18**, 7055–7062.
6. Andersen, J., Delihas, N., Hanas, J.S. and Wu, C.-W. (1984) *Biochemistry* **23**, 5752–5759.
7. Liu, W., Lo, A.C. and Nazar, R.N. (1983) *J. Mol. Biol.* **171**, 217–224.
8. Stebbins-Boaz, B. and Gerbi, S.A. (1991) *J. Mol. Biol.* **217**, 93–112.
9. Moazed, D., Stern, S. and Noller, H.F. (1986) *J. Mol. Biol.* **187**, 399–416.
10. Hogan, J.J., Gutell, R.R. and Noller, H.F. (1984) *Biochemistry* **23**, 3322–3330.
11. Hogan, J.J., Gutell, R.R., and Noller, H.F. (1984) *Biochemistry* **23**, 3330–3335.
12. Holmberg, L., Melander, Y. and Nygård, O. (1992) *J. Biol. Chem.* **267**, 21906–21910.
13. Caruthers, M.H., Beaton, G., Wu, J.V. and Wiesler, W. (1992) Methods Enzymology **211**, 3–20.
14. Sundkvist, I.C. and Staehelin, T. (1975) *J. Mol. Biol.* **99**, 401–418.
15. Nygård, O. and Nika, H. (1982) *EMBO J.* **1**, 357–362.
16. Ehresmann, C., Baudin, F., Mongel, M., Romby, P., Ebel, J.-P. and Ehresmann, B. (1987) *Nucl. Acids Res.* **15**, 9109–9128.
17. Brawerman, G., Mendecki, J. and Lee, S.Y. (1972) *Biochemistry* **11**, 637–641.
18. Christiansen, J., Egebjerg, J., Larsen, N. and Garrett, R.A. (1990) In *Ribosomes and Protein Synthesis. A Practical Approach* (Spedding, G. ed) pp 229–252, IRL Press, Oxford, U.K.
19. Westermann, P., Nygård, O. and Bielka, H. (1981) *Nucl. Acids Res.* **9** 2387–2396.
20. Nygård, O. and Nilsson, L. (1989) *Eur. J. Biochem.* **179**, 603–608.
21. Nygård, O. and Westermann, P. (1982) *Biochim. Biophys. Acta* **697**, 263–269.
22. Stern, S., Moazed, D. and Noller, H.F. (1988) *Methods Enz.* **164**, 481–489.
23. Maden, B.E.H. (1990) *Prog. Nucl. Acids Res. Mol. Biol.* **39**, 241–303.
24. Raynal, F., Michot, B. and Bachellerie, J.-P. (1984) *FEBS Lett.* **167**, 263–268.
25. Connaughton, J.F., Rairkar, A. and Kumar, A. (1984) *Nucl. Acids Res.* **12**, 4731–4745.
26. Salim, B. and Maden B.E.H. (1981) *Nature* **291**, 205–208.
27. Woese, C.R., Winker, S. and Gutell, R.R. (1990) *Proc. Nat. Acad. Sci. USA* **87**, 8467–8471.
28. Woese, C.R. and Gutell, R.R. (1989) *Proc. Nat. Acad. Sci. USA* **86**, 3119–3122.
29. Brimacombe, R., Atmadja, J., Stiege, W. and Schuler, D. (1988) *J. Mol. Biol.* **199**, 115–136.
30. Stern, S., Powers, T., Changchien, L.M. and Noller, H.F. (1989) *Science* **244**, 783–790.
31. Paz, V., Chan, Y.L., Glück, A. and Wool, I.G. (1989) *Nucl. Acids Res.* **17**, 9484.
32. Chan, Y.L., Paz, V., Olvera, J. and Wool, I.G. (1990) *FEBS Lett.* **263**, 85–88.
33. Chan, Y.L. and Wool, I.G. (1990) *Biochim. Biophys. Acta* **1049**, 93–95.
34. Hassouna, N., Michot, B. and Bachellerie, J.-P. (1984) *Nucl. Acids Res.* **12**, 3563–3583
35. Nazar, R.N. (1980) *FEBS Lett.* **119**, 212–214.
36. Ware, V.C., Tague, B.W., Clark; C.G., Gourse, R.L., Brand, R.C. and Gerbi, S.A. (1983) *Nucl. Acids Res.* **11**, 7795–7817.
37. Lo, A.C., Liu, W., Culham, D.E. and Nazar, R.N. (1987) *Biochem. Cell Biol.* **65**, 536–542.
38. Egebjerg, J., Leffers, H., Christensen, A., Andersen, H. and Garett, R.A. (1987) *J. Mol. Biol.* **196**, 125–136.
39. Egebjerg, J., Larsen, N. and Garrett, R.A. (1990) In *The structure, function and evolution of ribosomes* (Hill, W.E. ed.) Washington, D.C., American society for microbiology. pp 168–178.
40. Egebjerg, J., Douthwaite, S. and Garrett, R.A. (1989) *EMBO J.* **8**, 607–611.
41. Egebjerg, J., Douthwaite, S.R., Liljas, A. and Garrett, R.A. (1990) *J. Mol. Biol.* **213**, 275–288.
42. Leffers, H., Kjems, J., Ostergaard, L., Larsen, N. and Garrett, R.A. (1987) *J. Mol. Biol.* **195**, 43–61.
43. Noller, H.F. (1984) *Annu. Rev. Biochem.* **53**, 119–162.
44. Beauclerk, A.A.D. and Cundliffe, E. (1988) *EMBO J.* **7**, 3589–3594.
45. Endo, Y. and Wool, I.G. (1982) *J. Biol. Chem.* **257**, 9054–9060.
46. Hausner, T.P., Atmadja, J. and Nierhaus, K. (1987) *Biochimie* **69**, 911–923.
47. Moazed, D., Robertson, J.M. and Noller, H.F. (1988) *Nature* **334**, 362–364.
48. Leffers, H., Egebjerg, J., Andersen, A., Christensen, T. and Garrett, R.A. (1988) *J. Mol. Biol.* **204**, 507–522.
49. Osswald, M., Greuer, B. and Brimacombe, R. (1990) *Nucl. Acids Res.* **18**, 6755–6760.
50. Kuwano,Y and Wool, I.G. (1992) *Biochem. Biophys. Res. Commun.* **187**, 58–64.
51. Neefs, J.-M. and De Wachter, R. (1990) *Nucl. Acids Res.* **18**, 5695–5704.
52. Devereux, J., Haeberli, P., and Smithies, O. (1984) *Nucl. Acids Res.* **12**, 387–395.