Chemical Probing of Adenine Residues within the Secondary Structure of Rabbit 18S Ribosomal RNA†

Asha Rairkar, Heidi M. Rubino, and Raymond E. Lockard*
Department of Biochemistry, Cornell University Medical College, 1300 York Avenue, New York, New York 10021
Received April 21, 1987; Revised Manuscript Received September 29, 1987

ABSTRACT: The location of unpaired adenine residues within the secondary structure of rabbit 18S ribosomal RNA was determined by chemical probing. Naked 18S rRNA was first prepared by digestion of purified 40S subunits with matrix-bound proteinase K in sodium dodecyl sulfate, thereby omitting the use of nucleic acid denaturants. Adenines within naked 18S rRNA were chemically probed by using either diethyl pyrocarbonate or dimethyl sulfate, which specifically react with unpaired nucleotides [Peattie, D. A., & Gilbert, W. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 4679-4682]. Adenine modification sites were identified by polyacrylamide sequencing gel electrophoresis either upon aniline-induced strand scission of ^{32}P-end-labeled intact and fragmented rRNA or by primer extension using sequence-specific DNA oligomers with reverse transcriptase. The data indicate good agreement between the general pattern of adenine reactivity and the location of unpaired regions in 18S rRNA determined by comparative sequence analysis [Chan, Y.-L., Gutell, R., Noller, H. F., & Wool, I. G. (1984) *J. Biol. Chem. 259*, 224-230]. The overall reactivity of adenine residues toward single-strand-specific chemical probes was, also, similar for both rabbit and *Escherichia coli* small rRNA. The number of strongly reactive adenines appearing within phylogenetically determined helical segments, however, was greater in rabbit 18S rRNA than for *E. coli* 16S rRNA. Some of these adenines were found clustered in specific helices. Such differences suggest a greater irregularity of many of the helical elements within mammalian 18S rRNA, as compared with prokaryotic 16S rRNA. These helical irregularities could be important for protein association and also may represent biologically relevant flexible regions of the molecule.

In both prokaryotes and eukaryotes, the small ribosomal subunit is the focal point where specific protein factors, initiator Met-tRNA and mRNA, organize into a functional initiation complex. Translational initiation in eukaryotes, however, is a much more intricate process involving at least nine factors which must interact with the 40S subunit in a discreet way to ensure proper positioning of the Met-tRNA with the AUG initiator codon (Jagus et al., 1981; Kozak, 1983). The chemical composition of the eukaryotic 40S subunit is also different from its prokaryotic counterpart. It has been determined that 16S rRNA comprises ~60% of the mass of the 30S subunit (Wittman, 1982), while chemical determination of rat liver 40S subunits has revealed that 18S rRNA contributes 45% of its mass (Wool, 1980). This difference is largely due to the association of at least 10 additional ribosomal proteins with 18S rRNA, giving a total of ~31 proteins for the 40S subunit compared with only 21 in the *Escherichia coli* 30S particle (Wool, 1980). Such a variation is likely a reflection of a more complex eukaryotic cellular environment and underscores the need for studying the structure of mammalian small ribosomal RNA as it relates to eukaryotic protein biosynthesis.

A tentative secondary structure model has recently been proposed for rat 18S rRNA derived by comparative sequence analysis (Chan et al., 1984). The overall structural organization of the 18S rRNA is quite similar to eubacterial 16S rRNA, having four major demarcated domains. The increased size of the rat small rRNA (1871 nucleotides) compared with *E. coli* 16S rRNA (1542 nucleotides) is attributed largely to two specific insertion elements, as well as to the lengthening of helices within other regions of the molecule (Chan et al., 1984). Secondary structures for the insertion elements were not presented in these studies; however, a structure prediction for the major insert, 680-917, of rat 18S rRNA has recently been proposed (Choi, 1985) which is compatible with two T_1 ribonuclease cleavage sites on the 40S subunit. Partial structure data are available for yeast 18S rRNA from probing 40S subunits with the chemical probe kethoxal which reacts with unpaired guanine residues (Hogan et al., 1984). Of the 48 kethoxal reactive sites detected, 36 were found to be within phylogenetically determined single-stranded regions. Of the 12 discrepancies, some could be explained by their possible weak nearest-neighbor stacking interactions, while others might actually represent regions of the molecule which undergo possible conformational rearrangements (Hogan et al., 1984). Using the single-strand-specific chemical probe dimethyl sulfate in conjunction with DNA primer extension, Bachellerie and co-workers (Lempereur et al., 1985) have probed the structure of the first 320 nucleotides of the 5′-structural domain of yeast 18S rRNA. Substantial reactivity was observed in this region with naked 18S rRNA but was greatly reduced in the 40S subunit. Data are also available for the 3′-minor domain of yeast 18S rRNA employing both chemical and enzymatic structure probes (Douthwaite et al., 1983). Direct experimental determination of five secondary structural elements in *Xenopus laevis* 18S rRNA was recently performed upon analysis of base-paired fragments generated by partial ribonuclease digestion of the rRNA (Atmadja et al., 1984). These data were used to refine some phylogenetically derived helical regions and additionally identified a possible eukaryotic "switch" structure.

It has previously been determined that unpaired nucleotides in prokaryotic 16S rRNA show a much higher degree of sequence conservation than do base-paired nucleotides (Gutell et al., 1985). Furthermore, a disproportionate number of these

† This research was supported by National Institutes of Health Grant GM 35467.

0006-2960/88/0427-0582$01.50/0

universally conserved nucleotides appear to be adenine residues. There is also growing evidence suggesting the specific involvement of unpaired adenines in ribosomal protein association (Moazed et al., 1986a; Christiansen & Garrett, 1985; Gregory & Zimmermann, 1986; Stern et al., 1986). To eventually understand the role of unpaired nucleotides in mammalian small rRNA, the location of unpaired adenine residues within the secondary structure of rabbit 18S ribosomal RNA was determined by chemical probing. A method is described for isolating naked 18S rRNA from rabbit 40S subunits which omits the use of nucleic acid denaturants. Adenines within naked 18S rRNA were chemically probed by using either diethyl pyrocarbonate (DEPC)[1] or dimethyl sulfate (Me_2SO_4), which specifically react with unpaired nucleotides (Peattie & Gilbert, 1980). The adenine reactivity data were found to be in generally good agreement with the phylogenetically derived structure model for mammalian 18S rRNA (Chan et al., 1984). However, there were also some interesting differences with the proposed structure. The possible relevance of these differences to both protein binding and conformational flexibility is considered and discussed.

Materials and Methods

Chemicals and Enzymes. Ultrapure sucrose (RNase free) was purchased from Schwarz/Mann. Polyacrylamide gel electrophoresis reagents were obtained from Baker Chemicals. Ultrapure urea was supplied by Bio-Rad. Reagent-grade formamide and diethyl pyrocarbonate (DEPC) were purchased from Sigma Chemical Co. Dimethyl sulfate (Me_2SO_4) was obtained from Aldrich Chemical Co. Centricon-30 microconcentrators were supplied by Amicon Corp. The silver-staining kit was from NEN Research Inc. Deoxyoligomers I [5′-d(CATCGTTTATGGTCGGAACTAC)3′] and III [5′-d(CGCTCGGGCCTGCTTTAA)3′] were synthesized by Collaborative Research Inc. Deoxyoligomer II [5′-d-(GTCCAAGAATTTCACCTCTAG)3′] was synthesized in the laboratory of Dr. P. Model, Rockefeller University, New York, NY. Deoxynucleotides, dideoxynucleotides, and T_4 polynucleotide kinase were supplied by Boehringer-Mannheim. T_4 RNA ligase was from P-L Biochemicals Inc. RNase T_1 and RNase U_2 were prepared by Sankyo and obtained through Calbiochem. Matrix-bound proteinase K was obtained from E. Merck. *E. coli* RNase H was a kind gift of Dr. Robert Crouch (NIH). Avian myeloblastosis virus (AMV) reverse transcriptase was purchased from Life Sciences Inc. [γ-^{32}P]ATP (4500 Ci/mmol) and [^{32}P]Cp (1905 Ci/mmol) were purchased from ICN Biochemical Inc.

Preparation of Naked 18S rRNA. Isolation of reticulocytes from anemic New Zealand white rabbits and preparation of polyribosomes have been previously described (Lockard & RajBhandary, 1976). Polyribosomes were dissociated with GTP and puromycin (Blobel & Sabatini, 1971), and the resulting 40S ribosomal subunits were isolated by centrifugation through linear 15–30% w/w sucrose gradients containing 500 mM KCl, 5 mM β-mercaptoethanol, and 50 mM triethanolamine hydrochloride (pH 7.5) buffer. Naked 18S rRNA was prepared from purified 40S ribosomal subunits by mild deproteinization with matrix-bound proteinase K (300 μg/mL) at 37 °C in 25 mM triethanolamine hydrochloride, pH 7.5, 1% sodium dodecyl sulfate (SDS), and 5 mM β-mercaptoethanol, followed by fractionation on linear 15–30% w/w sucrose gradients containing 100 mM NaCl, 0.5% SDS, 5 mM β-mercaptoethanol, and 50 mM triethanolamine hydrochloride (pH 7.5). Fractions containing naked 18S rRNA were pooled and concentrated by diafiltration through Centricon-30 microconcentrators. The detailed procedure is outlined in Figure 1. The integrity and homogeneity of the isolated naked 18S rRNA were determined by both 5′- and 3′-^{32}P-end-labeling (Lockard et al., 1982). High-sensitivity amino acid analyses of the naked 18S rRNA were performed by using the precolumn phenyl isothiocyanate derivatization procedure (Bidlingmeyer et al., 1984).

Chemical Probing of Naked 18S rRNA. Chemical modification procedures were similar to those described before (Peattie & Gilbert, 1980). Prior to carbethoxylation with DEPC or methylation with Me_2SO_4, 20 μg of naked 18S rRNA was preincubated in 200 μL of structure buffer (200 mM NaCl, 20 mM $MgCl_2$, and 50 mM sodium cacodylate, pH 7.0) for 10–15 min at 37 °C and allowed to slow cool to room temperature. For carbethoxylation, 4 μL (2% v/v final concentration) or 5 μL (2.5% v/v final concentration) of DEPC was added to 20 μg of 18S rRNA in 200 μL of structure buffer. The reaction was incubated for 15 min at 37 °C with frequent gentle mixing, terminated by the addition of 50 μL of DEPC stop buffer (1.5 M sodium acetate) plus 3 volumes of 95% ethanol, and rapidly frozen to −70 °C. The modified 18S rRNA was precipitated, pelleted, and dissolved in 0.3 M sodium acetate and reprecipitated with ethanol. Pellets were repeatedly washed with 80% ethanol and finally dissolved in water. For methylation, 0.25 μL (0.12% v/v final concentration), 0.5 μL (0.25% v/v final concentration), or 1 μL (0.5% v/v final concentration) of Me_2SO_4 was added to 20 μg of 18S rRNA in 200 μL of structure buffer. The mixture was incubated for 15 min at 37 °C. The reaction was terminated by the addition of 100 μL of cold Me_2SO_4 stop buffer (1.0 M Tris-HCl, pH 7.5, 1 M β-mercaptoethanol, 1.5 M sodium acetate, and 0.1 M EDTA) plus 3 volumes of 95% ethanol and rapidly frozen to −70 °C. The modified RNA was pelleted and dissolved in 0.3 M sodium acetate and reprecipitated with ethanol. The methylated 18S rRNA was further reduced with alkaline sodium borohydride to stabilize the modifications (Peattie & Gilbert, 1980) and then dissolved in water. Control (unmodified) RNA samples, for both the methylation and carbethoxylation reactions, were treated identically except for the omission of Me_2SO_4 and DEPC, respectively.

Identification of Modification Sites within 18S rRNA. Four different approaches were used to identify methylated and carbethoxylated adenine residues within the entire 18S rRNA.

(1) ^{32}P-End-Labeling of Intact 18S rRNA. Intact carbethoxylated 18S rRNA was ^{32}P-end-labeled at its 5′-terminus with [γ-^{32}P]ATP and T_4 polynucleotide kinase or at its 3′-terminus with [^{32}P]Cp and T_4 RNA ligase (Lockard et al., 1982). After purification and recovery of the ^{32}P-end-labeled 18S rRNA (Lockard et al., 1978), the labeled rRNA was treated with 1 M aniline, pH 4.5, at 45 °C for 20 min in the dark to induce strand scission. The resulting fragments were electrophoretically fractionated on long polyacrylamide sequencing gels along with enzymatic sequencing reactions (Lockard et al., 1982). The gels were autoradiographed using Kodak XAR-5 film and the sites of modification within the 5′ domain (nucleotides 1–400) and the 3′ domain (nucleotides 1500–1858) identified.

(2) Site-Specific Cleavage with RNase T_1. To map the chemically reactive adenine residues within internal regions

[1] Abbreviations: DEPC, diethyl pyrocarbonate; DMS, dimethyl sulfate (Me_2SO_4); RNase, ribonuclease; EDTA, ethylenediaminetetraacetic acid; Tris, tris(hydroxymethyl)aminomethane; SDS, sodium dodecyl sulfate; TEA, triethanolamine.

of the molecule, naked 18S rRNA was carbethoxylated with DEPC and site specifically cleaved with RNase T_1 (Connaughton et al., 1984a). The resultant fragments were labeled at their 5′-termini with [γ-^{32}P]ATP and T_4 polynucleotide kinase and fractionated on a preparative 3.5% polyacrylamide gel (80 cm long × 20 cm wide × 1.5 mm thick) containing 7 M urea as detailed previously (Connaughton et al., 1984a). Upon purification and recovery, the 5′-^{32}P-end-labeled fragments 1, 2, and 3 were treated with acidic aniline as described above to induce strand scission at the modified nucleotides. The reaction products were electrophoretically fractionated on polyacrylamide gels and the sites of modification identified after autoradiography.

(*3*) *Site-Specific Cleavage with RNase H.* Carbethoxylated 18S rRNA was hybridized with a *Pst*I DNA linker and the RNA/DNA hybrid digested with *E. coli* RNase H as previously detailed (Connaughton et al., 1984a). The resulting fragments were ^{32}P-end-labeled at their 3′-termini and fractionated on a preparative 3.5% polyacrylamide gel (80 cm × 20 cm × 1.5 mm) containing 7 M urea as described for the T_1 RNase derived fragments. Purified 3′-^{32}P-end-labeled fragments A and B (Connaughton et al., 1984a) were subjected to aniline-induced strand scission and the sites of modification identified.

(*4*) *Primer Extension with Reverse Transcriptase.* The three synthetic deoxyoligomers (I, II, and III) listed in Chemicals and Enzymes were 5′-^{32}P-end-labeled using [γ-^{32}P]ATP and T_4 polynucleotide kinase and purified on preparative 20% polyacrylamide gels run in 7 M urea (Inoue & Cech, 1985); 0.03 pmol of a specific 5′-^{32}P-end-labeled DNA primer was heat denatured (90 °C, 3 min) and added to 0.3 pmol of Me_2SO_4-modified 18S rRNA dissolved in water and equilibrated at 65 °C. After 10 min at 65 °C, the reaction mixture was brought to a final concentration of 50 mM Tris-HCl, pH 8.3, 6 mM $MgCl_2$, 40 mM KCl, and 20 mM dithiothreitol by adding 10× reverse transcriptase buffer (500 mM Tris-HCl, pH 8.3, 400 mM KCl, 60 mM $MgCl_2$, and 200 mM dithiothreitol) and further incubated at 65 °C for 7 min before slow cooling to room temperature. To the rRNA/DNA hybrid mixture, 10× extension buffer (4 mM each of dATP, dGTP, dCTP, and dTTP in 50 mM Tris-HCl, pH 8.3) was added to give a final concentration of 400 μM for each dNTP in a 20-μL reaction volume. Six units of AMV reverse transcriptase were then added to the mixture, and elongation of the primer was performed at 37 °C for 30 min. One microliter of 10× chase buffer (2 mM each of dATP, dGTP, dCTP, and dTTP in 50 mM Tris-HCl, pH 8.3) was then added and the incubation continued for an additional 15 min. The reaction was terminated by addition of 100 μL of 0.3 M sodium acetate and 400 μL of 95% ethanol, chilled at −80 °C for 2 h, and the transcripts were pelleted by centrifugation. Control (unmodified) rRNA was employed as a template for the dideoxy sequencing reactions (Qu et al., 1983). Transcription products for all primer extension reactions were vacuum dried and resuspended in 10 μL of loading solution (80% deionized formamide, 10 mM NaOH, 1 mM EDTA, 0.1% xylene cyanol, and 0.1% bromophenol blue). Upon heat denaturation at 90 °C for 2 min followed by quick chilling to 4 °C, the samples were fractionated by electrophoresis on both 10% and 20% polyacrylamide sequencing gels (Lockard et al., 1978). Sites of modification on all gels were determined by microdensitometric analysis of autoradiograms using either an UltroScan XL (LKB Instruments) or a DU-8B scanning spectrophotometer (Beckman Instruments). Peak areas for each band were calculated and these values corrected by subtracting areas from corresponding control lane bands.

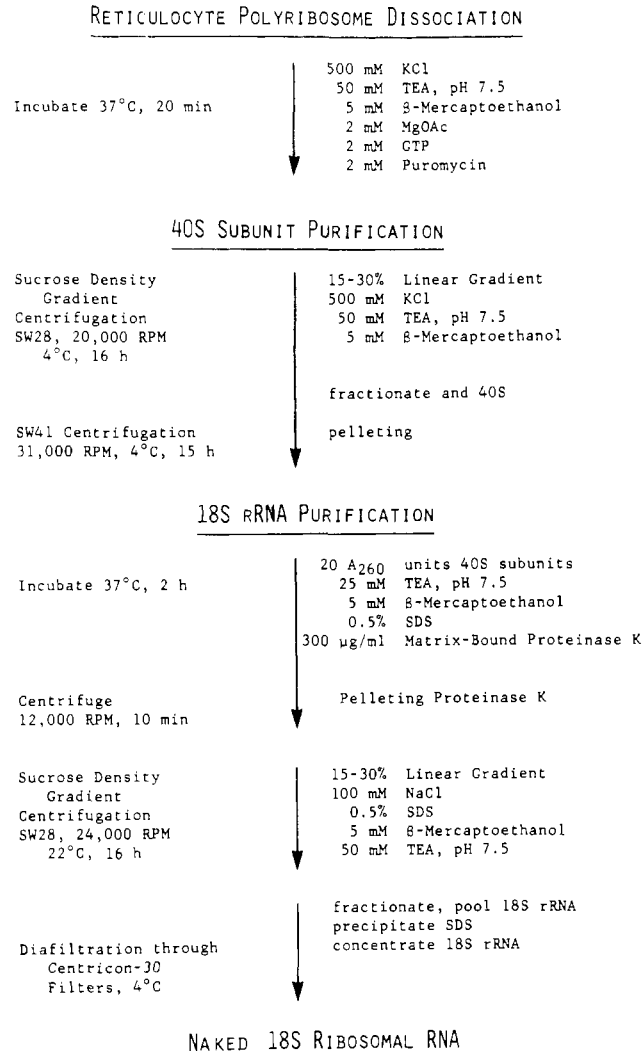


FIGURE 1: Preparation of naked 18S rRNA from rabbit reticulocytes.

RESULTS

Preparation of Naked 18S rRNA. Figure 1 outlines a method for preparing naked 18S rRNA from rabbit reticulocytes. Purified 40S subunits were first digested with matrix-bound proteinase K in 1% sodium dodecyl sulfate (SDS), followed by fractionation of the RNA on SDS–sucrose density gradients and concentration using Centricon-30 microconcentrators. Care was taken to avoid possible disruption of the rRNA structure by omitting organic solvents and denaturants such as urea, phenol, EDTA, ethanol, etc. The resultant 18S rRNA was found to be intact by polyacrylamide gel electrophoretic analysis of both 5′- and 3′-^{32}P-end-labeled rRNA (Connaughton et al., 1984b). Figure 2 shows an SDS–polyacrylamide gel electrophoretic analysis of the naked 18S and of 18S rRNA prepared by phenol extraction. The absence of ribosomal protein bands in both lanes suggests that the naked 18S rRNA is comparable in purity to the phenol-extracted preparation. In order to obtain a quantitative estimate of RNA purity, amino acid analysis was performed by using a precolumn phenyl isothiocyanate (PITC) derivatization procedure (Bidlingmeyer et al., 1984). RNA samples were hydrolyzed in 6 M hydrochloric acid at 110 °C for 24 h and derivatized in a Pico-Tag workstation (Waters Associates). The quantity of RNA used for each analysis was sufficient enough to detect a protein impurity present in less than 1% by weight. Analyses indicated that both the naked and the

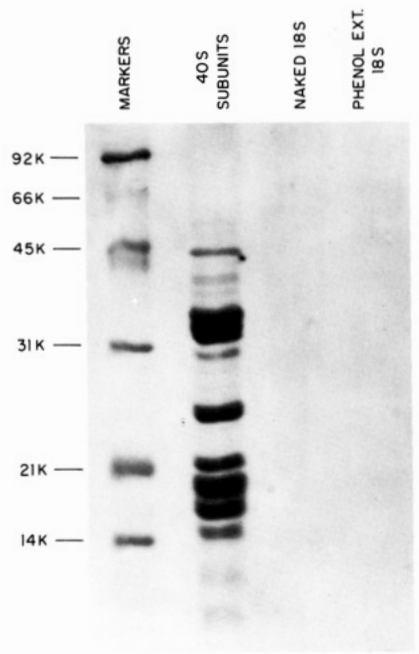


FIGURE 2: Protein analysis of naked 18S ribosomal RNA by SDS–polyacrylamide gel electrophoresis (15% PAGE; silver staining). Lanes contained 0.5 A_{260} unit of rRNA.

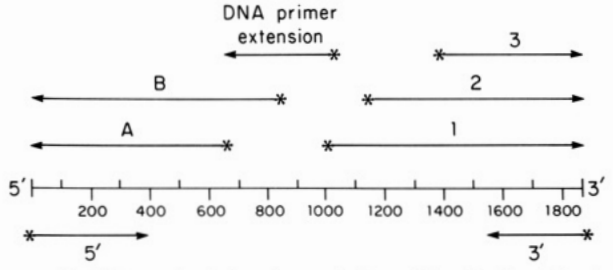


FIGURE 3: Scheme depicting the methods used for the identification of chemically reactive bases in rabbit 18S rRNA. Asterisks designate ^{32}P-end-label. (1, 2, and 3) T_1 RNase digestion fragments. (A and B) RNase H digestion fragments. (DNA primer extension) Utilization of [5′-^{32}P]DNA oligomers I, II, and III (see Materials and Methods).

phenol-extracted 18S rRNA preparations were not less than 99.7% RNA by weight.

Chemical Probing of Adenine Residues with Diethyl Pyrocarbonate (DEPC). Several studies indicate that diethyl pyrocarbonate (DEPC) can be used successfully for detecting single-stranded adenines in a variety of RNA molecules (Peattie & Gilbert, 1980; Van Stolk & Noller, 1984; Christiansen & Garrett, 1985). DEPC reacts at both the N-6 and N-7 positions of adenine residues (Vincze et al., 1973) and can be used as an indirect measure of single strandedness. Reactivity is generally indicative of an absence of base stacking. Lack of reactivity, however, is more difficult to interpret since it could result not only from base pairing and tertiary interactions but also from the presence of stabilizing divalent cations (Peattie & Gilbert, 1980). We have employed three different approaches for the identification of DEPC modification sites within naked 18S rRNA as depicted schematically in Figure 3. Either intact or fragmented rRNA was end-labeled with ^{32}P, and modified adenines were identified, upon aniline-induced strand scission, as bands on an autoradiogram of a polyacrylamide sequencing gel. Initial studies were directed toward determination of modification sites within the 5′ and 3′ domains of 18S rRNA. Intact-modified 18S rRNA was ^{32}P-end-labeled either at its 5′-terminus with [γ-^{32}P]ATP and T_4 polynucleotide kinase or at its 3′ end with [^{32}P]Cp and T_4 RNA ligase (Lockard et al., 1982).

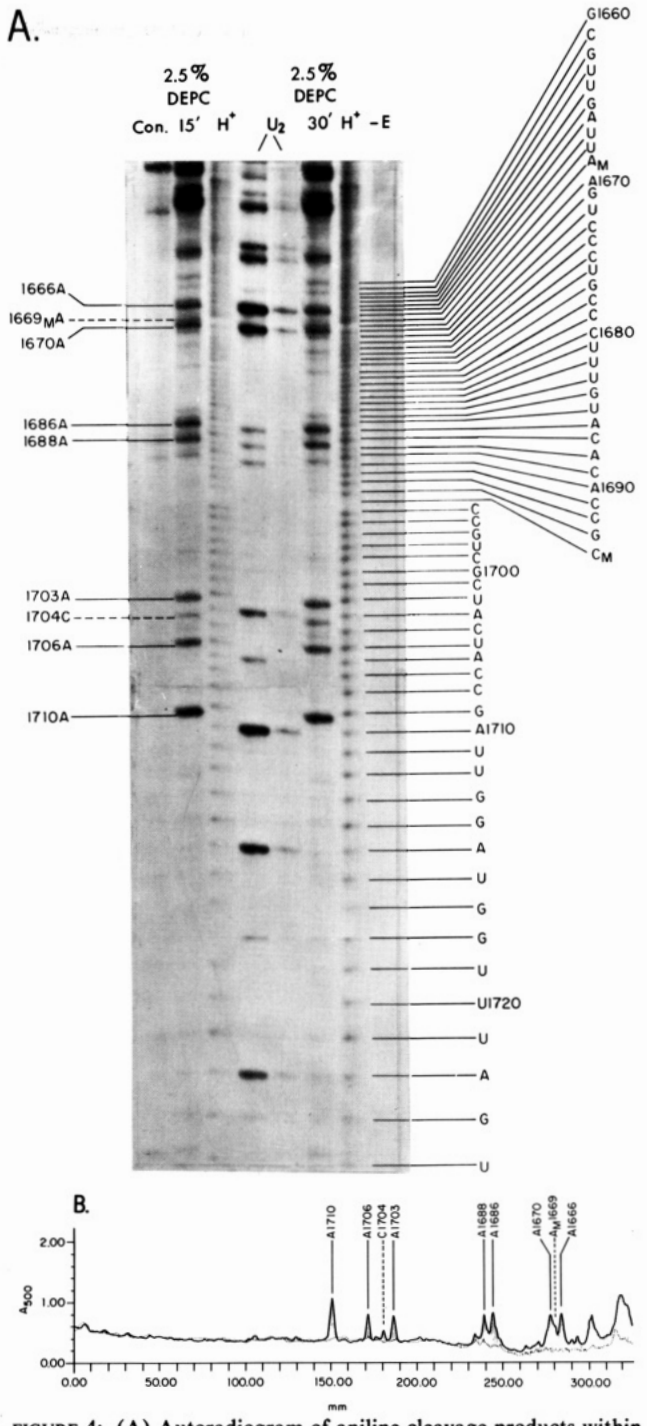


FIGURE 4: (A) Autoradiogram of aniline cleavage products within the region 1660–1724 derived from DEPC-modified 3′-^{32}P-end-labeled 18S rRNA electrophoresed on a 10% polyacrylamide gel in 95% formamide. Gel dimensions were 0.4 mm thick × 33 cm wide × 100 cm long. Strong and weak modifications are indicated by (—) and (---), respectively. From left to right: (Con.) aniline-treated control; (2.5% DEPC-15′) 2.5% DEPC v/v, 15 min, 37 °C; (H^+) partial acid hydrolysis; (U_2) U_2 RNase, 0.25 × 10^{-4} unit/μg of RNA and 0.25 × 10^{-5} unit/μg of RNA; (2.5% DEPC-30′) 2.5% DEPC v/v, 30 min, 37 °C; (−E) minus enzyme. (B) Densitometric scan of lane (2.5% DEPC-15′) for above autoradiogram.

The ^{32}P-end-labeled RNAs were purified and then treated with aniline to induce strand scission. Resulting cleavage products were electrophoretically fractionated on polyacrylamide sequencing gels run in 95% formamide (Lockard et al., 1982). Figure 4A shows a representative gel analysis of DEPC modification sites in the region 1660–1724 of the 3′ domain. Visual identification of these sites was usually straightforward.

Microdensitometric scanning of the autoradiogram, however, was necessary for a more quantitative estimate of reactivity. Figure 4B shows a densitometric scan of the autoradiogram appearing in Figure 4A. Care was taken not to overexpose gels so as to ensure a linear relationship between radioactivity and absorbance. The shaded areas within individual peaks indicate the difference in absorbance after subtraction of control lane bands. Seven strongly reactive adenines were identified in this region of the RNA. Am_{1669}, which is only partially resolved on this gel, was found to be weakly reactive in several other experiments. A_{1690}, A_{1715}, and A_{1722} are unreactive. It is noteworthy that we have identified occasional modification of cytosine residues with DEPC. In Figure 4A,B, C_{1704} is a reproduciby weak modification site. Carbethoxylation of cytosine has been described previously (Vincze et al., 1973) and has been observed as a minor modification reaction in 16S rRNA (Van Stolk & Noller, 1984).

To map DEPC reactive sites within internal regions of the molecule, modified 18S rRNA was site specifically cleaved with RNase H or T_1 ribonuclease. Figure 3 depicts schematically the overlapping fragments generated by using either enzyme. We previously determined that site-directed cleavage of 18S rRNA is possible using a *Pst*I DNA linker in conjunction with *E. coli* RNase H (Connaughton et al., 1984a). As shown in Figure 3, two unique fragments, A and B, are generated which overlap with the 5′ domain. These fragments were labeled at their 3′-termini with ^{32}P and purified from 3.5% preparative polyacrylamide gels, and their modification sites were determined upon aniline-induced strand scission (Connaughton et al., 1984a). Specific cleavage with T_1 ribonuclease generated the three overlapping fragments 1, 2, and 3, which were ^{32}P-end-labeled at their 5′-termini. These labeled fragments were also purified on preparative 3.5% polyacrylamide gels and modification sites determined upon aniline cleavage (Connaughton et al., 1984a).

Figure 5A shows a representative gel analysis of DEPC modification sites in the region 1158–1219 of T_1 fragment 2. This area is of interest since it encompasses a putative pseudoknot as well as the region demarcating the 3′-major domain (see Discussion). Visual identification of reactive sites was again straightforward. The microdensitometric scan shown in Figure 5B indicated eight strongly reactive and four weakly reactive adenine residues. A_{1166}, A_{1169}, A_{1177}, A_{1178}, A_{1179}, A_{1200}, and A_{1210} were all unreactive. A_{1186} was found to be very weakly reactive. Clearly, the reactivity of some nucleotides is very marginal and often difficult to evaluate. We have observed occasional reactivity of guanine residues as typified by the weak reactivity of G_{1194}. Carbethoxylation of guanine has been described previously (Vincze et al., 1973) and was also observed as a minor modification reaction in 16S rRNA (Van Stolk & Noller, 1984; Douthwaite et al., 1983).

Chemical Probing of Adenine Residues with Dimethyl Sulfate (Me_2SO_4). Sites of modification within the remaining portion of the molecule were determined by using the new method of DNA primer extension with reverse transcriptase as depicted in Figure 3 (Lempereur et al., 1985; Inoue & Cech, 1985; Moazed et al., 1986a). A specific 5′-^{32}P-end-labeled synthetic DNA oligomer is first annealed to Me_2SO_4-modified 18S rRNA and then extended by using AMV reverse transcriptase in the presence of four deoxynucleoside triphosphates. With Me_2SO_4-modified RNA, the reverse transcriptase will stop one nucleotide before a methylated base provided the modification impairs Watson–Crick base pairing between the newly synthesized cDNA and the template RNA. m^1A, m^3C, m^3U, and also m^1G can be detected as Me_2SO_4-induced

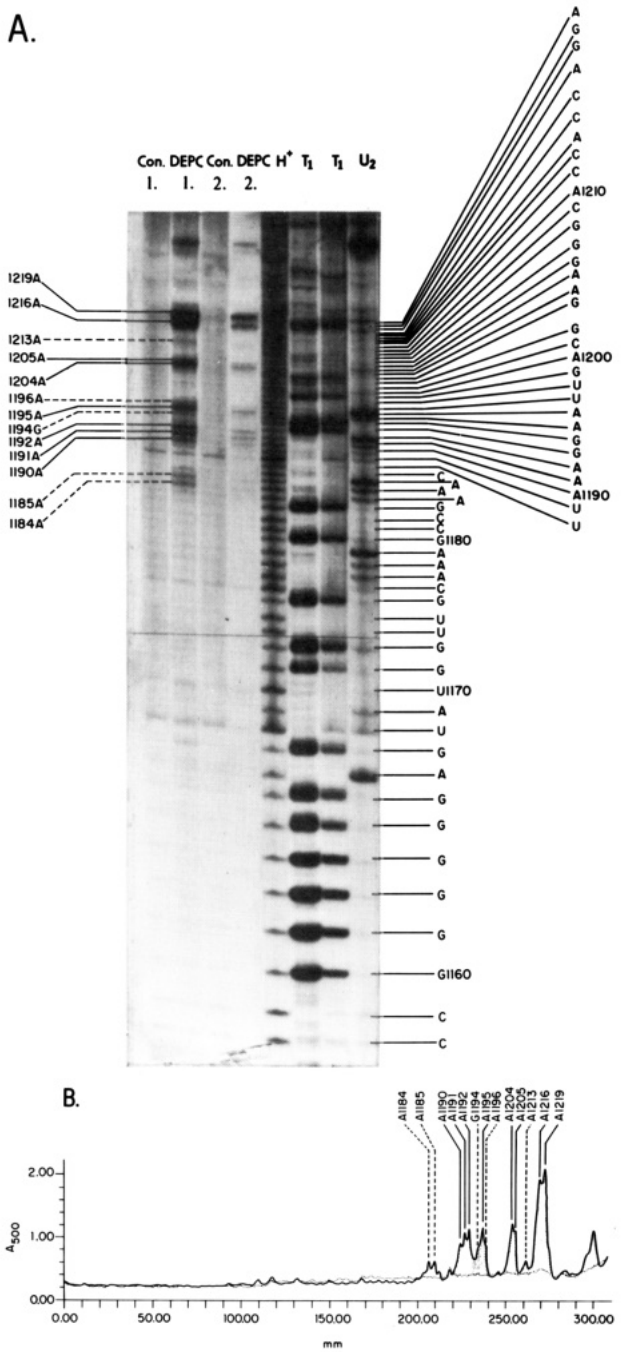


FIGURE 5: (A) Autoradiogram of aniline cleavage products within the region 1158–1219 of DEPC-modified RNase T_1 fragment 2 electrophoresed on a 10% polyacrylamide gel in 95% formamide. Gel dimensions were as in Figure 4A. Strong and weak modifications are indicated by (—) and (---), respectively. From left to right: (Con. 1) aniline-treated control; (DEPC 1) 2% DEPC v/v, 15 min, 37 °C, aniline reaction 20 min, 45 °C; (Con. 2) aniline-treated control; (DEPC 2) 2% DEPC v/v, 15 min, 37 °C, aniline reaction 10 min, 45 °C; (H^+) partial acid hydrolysis; (T_1) T_1 RNase, 1.25×10^{-3} unit/μg of RNA and 1.25×10^{-4} unit/μg of RNA; (U_2) U_2 RNase, 0.25×10^{-4} unit/μg of RNA. (B) Densitometric scan of lane (DEPC 1) for above autoradiogram.

elongation blocks. Such stops give rise to bands on polyacrylamide sequencing gels corresponding to the extension of the [5′-^{32}P]DNA primer to the nucleotide preceding the modified position. Identification of the modified base is facilitated by running dideoxy sequencing reactions on the same gel.

Three sequence-specific DNA primers were utilized (I, II, and III; see Materials and Methods) which allowed determination of methylated sites within the region overlapping

Table I: Reactivity of Adenine Residues toward both DEPC and DMS within Region 650–825 of Rabbit 18S rRNA[a]

position	reactivity	
	DEPC (N-7)	DMS (N-1)
A_{654}	++	–
Am_{658}	+	+
A_{659}	++	++
A_{660}	++	++
A_{661}	++	++
A_{662}	++	++
A_{669}	++	++
A_{675}	++	+
A_{684}	–	–
A_{693}	–	–
A_{711}	–	–
A_{716}	–	–
A_{720}	–	–
A_{759}	–	–
A_{767}	+	++
A_{772}	++	–
A_{790}	+	++
A_{791}	++	–
A_{798}	+	++
A_{804}	++	–
A_{805}	++	++
A_{806}	++	++
A_{807}	++	++
A_{808}	++	++
A_{809}	++	–
A_{812}	++	+
A_{814}	++	+
A_{821}	+	+
A_{822}	+	+
A_{823}	+	++

[a] Relative reactivities were determined by microdensitometric scanning of autoradiograms and represent a consensus from several independent experiments. Unreactive (–); weakly reactive (+); strongly reactive (++).

RNase T_1 fragment 1 and RNase H fragment B as depicted in Figure 3. Figure 6A shows a representative polyacrylamide sequencing gel analysis of Me_2SO_4 modification sites within the region 880–942 determined by extension of [^{32}P]DNA primer II, which is complementary to the sequence 949–969. This region is of interest since it corresponds to the prokaryotic ribosomal protein S_{15} binding domain (see Discussion). Since visual identification of Me_2SO_4 reactive sites was often not straightforward, microdensitometric scanning of both modified and control lanes was helpful. Figure 6B shows a densitometric scan of the 0.25% DMS (Me_2SO_4) lane in Figure 6A. The shaded areas within individual peaks indicate differences in absorbance after subtraction of corresponding bands in unmodified lanes. The peak areas of each band in the modified lane were calculated and these values then corrected by subtracting control band areas. A modification was considered strong if its calculated and corrected area was not less than 60% of the band with the greatest corrected intensity on that gel. Several overlapping gels were scanned to ensure consistency and reproducibility of strong modification sites. In Figure 6B, A_{934}, A_{922}, A_{916}, A_{915}, and A_{899} were determined to be strongly reactive. Interestingly, a cluster of three guanines (G_{932}, G_{931}, and G_{930}) were also found to be strongly modified. Although m^1G modification was previously reported to be a minor methylation reaction (Lempereur et al., 1985), clearly some guanine residues within the tertiary structure of 18S rRNA are accessible to Me_2SO_4 at this position. Such additional information was often helpful in conjunction with adenine reactivity data in discerning unpaired regions of the molecule.

Table I compares adenine reactivity toward both DEPC and Me_2SO_4 within the region 650–825. An examination of all adenines within this region indicated good agreement between DEPC and Me_2SO_4 reactivity for 25 of the 30 residues. There were five adenines (A_{654}, A_{772}, A_{791}, A_{804}, and A_{809}) which were strongly reactive with DEPC but unreactive toward Me_2SO_4. The specific accessibility of the N-7 and not the N-1 position toward chemical probes may be related to how these residues are positioned within the three-dimensional structure of the rRNA (see Discussion). Since the nucleotide sequence of rabbit 18S rRNA is 95% homologous to rat 18S rRNA (Connaughton et al., 1984a; Chan et al., 1984), the rabbit rRNA was folded in accordance with the phylogenetically derived secondary structure model for rat 18S rRNA (Chan et al., 1984). Figure 7 summarizes both the DEPC and Me_2SO_4 reactivity data and compares it with the phylogenetically derived secondary structure model for rabbit 18S rRNA. Also included in this figure are corrections of the originally published nucleotide sequence for rabbit 18S rRNA (Connaughton et al., 1984a).

Table II: Reactivity of Adenine Residues toward Single-Strand-Specific Chemical Probes within Rabbit and *E. coli* Small Ribosomal RNA

	reactivity			
	strong	weak	unreactive	total
rabbit	190 (45%)	81 (19%)	149 (36%)	420
E. coli[a]	171 (44%)[b]	72 (19%)	145 (37%)	388

[a] Data obtained by the primer extension method (Moazed et al., 1986). [b] Includes both strong and moderately reactive adenines (Moazed et al., 1986).

Table III: Distribution of Strongly Reactive Adenine Residues within the Phylogenetically Derived Secondary Structure Models for Rabbit and *E. coli* Small Ribosomal RNA

	single-stranded	base-paired	total
rabbit	135	55 (19)[b]	190
E. coli[a]	158	13 (2)[b]	171

[a] Moazed et al. (1986). [b] Strongly reactive adenines after exclusion of those with weak predicted nearest-neighbor stacking interactions (Tinoco et al., 1973).

Discussion

From Figure 7, it is evident that there is good agreement between the general pattern of adenine reactivity and unpaired regions of the rRNA as determined by comparative sequence analysis. In fact, the overall reactivity of adenine toward single-strand-specific chemical probes was similar for both rabbit and *E. coli* as indicated in Table II. The distribution of these reactive sites within the phylogenetically derived secondary structure models, however, was somewhat different for these two rRNAs, as summarized in Table III and revealed in more detail below.

5′ Domain 1–642. The helices (A_{22}–C_{30})·(G_{634}–U_{642}) and (U_{31}–A_{38})·(U_{552}–A_{560}) organize residues 1–642 into a 5′ domain. The corresponding domain in *E. coli* 16S rRNA has been shown to interact with both ribosomal proteins S_4 and S_{20} in the early assembly of the 30S subunit (Mizushima & Nomura, 1970). Previous analysis of 16S rRNA with DEPC (Van Stolk & Noller, 1984) indicated that adenine residues within the region 109–279 of the *E. coli* 5′ domain were relatively unreactive, suggesting the presence of a highly stable structure possibly involved in the early events of ribosome assembly. We also have observed low reactivity toward DEPC in the corresponding region 100–350 of rabbit 18S rRNA.

A structure prediction for the small eukaryotic insertion element 560–584 compatible with the DEPC reactivity data is proposed in Figure 7. The proximity of this hairpin structure

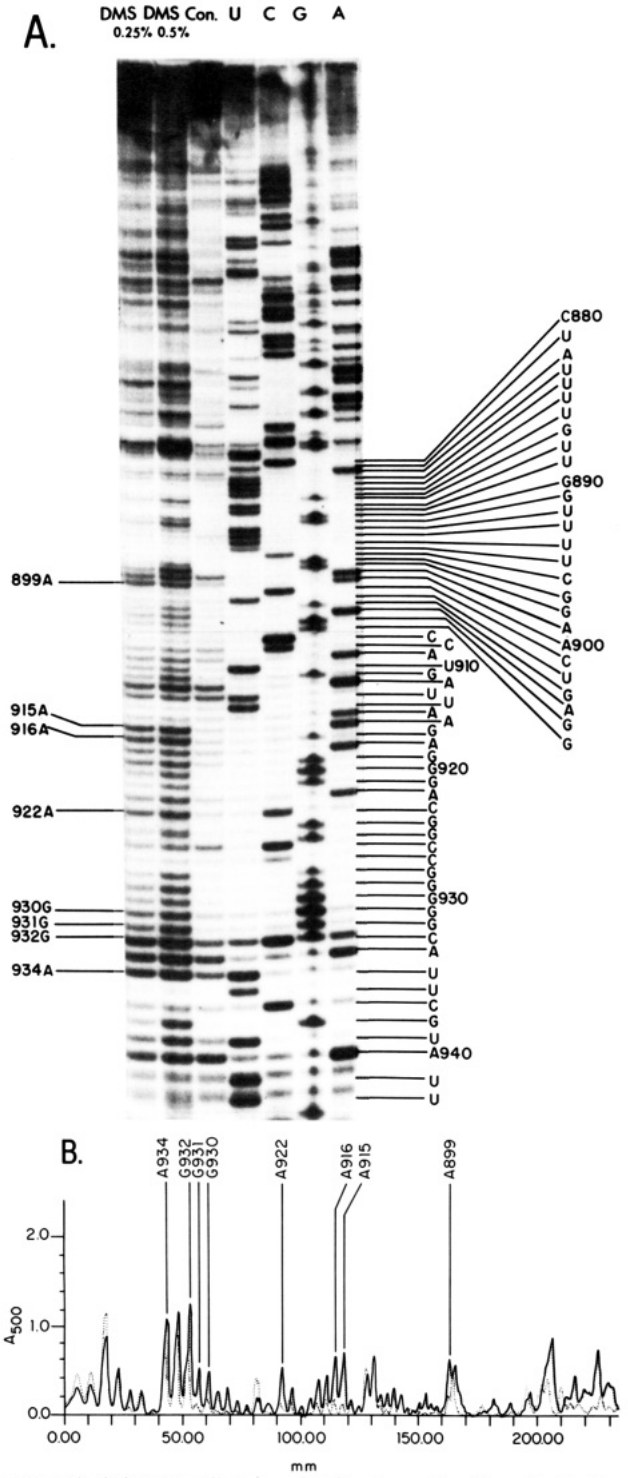


FIGURE 6: (A) Autoradiogram of 10% polyacrylamide gel indicating Me_2SO_4 (DMS) modification sites within the primer II extended region 880–942 of 18S rRNA. Strong modifications are indicated by (—). From left to right: (DMS 0.25%) and (DMS 0.50%) 0.25% and 0.50% dimethyl sulfate (DMS) v/v (15 min, 37 °C); (Con.) control (unmodified rRNA); (U, C, G, and A) dideoxy sequencing reactions. (B) Densitometric scan of lane (DMS 0.25%) for above autoradiogram.

to the protein S_4 binding domain of *E. coli* (Stern et al., 1986) is interesting and may be an important augmentation for the assembly of the eukaryotic 40S subunit.

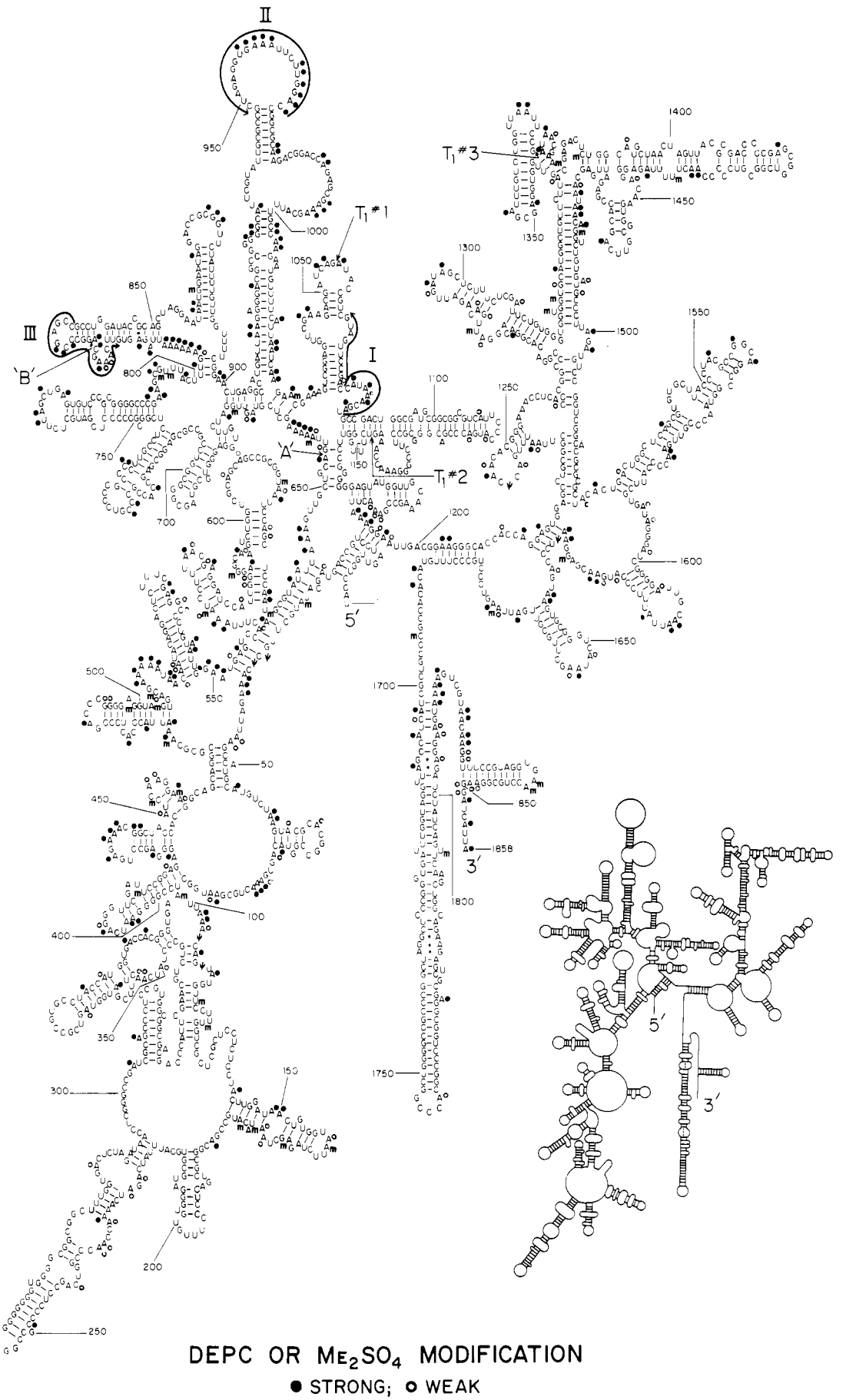
The putative interaction of the region U_{12}–U_{15} with A_{1192}–A_{1195} in rabbit 18S rRNA is an unusual type of pairing between two loops and has been previously termed a "pseudoknot" (Maly & Brimacombe, 1983; Pleij et al., 1985; Moazed et al., 1986a). The base pairing of these two loops is not entirely supported by the DEPC reactivity data shown in Figure 5. A_{1192} and A_{1195} were observed to be strongly modified and G_{1194} weakly reactive. However, similar lability of this region has been noted in both naked 16S rRNA and 30S subunits (Moazed et al., 1986a) as well as for yeast 40S subunits (Hogan et al., 1984).

As shown in Figure 7, 65 strongly reactive adenine residues were observed within the rabbit 5′ domain. Even though 45 of these modifications were found to reside in single-stranded regions of the phylogenetically derived secondary structure model, 20 discrepancies remained. The reactivities for 11 of these differences, however, could be rationalized by their possible weak nearest-neighbor stacking interactions (Tinoco et al., 1973; Freier et al., 1986) as summarized in Table IV. Three of these reactive adenines were found to be adjacent to a bulged nucleotide, while eight reside at the ends of hairpin loops. The opening and closing of A·U base pairs both internally and also at the ends of helical segments have been previously documented (Pörschke, 1974; Mandal et al., 1979) and may account for this observed reactivity.

Central Domain 650–1189. The helices $(C_{650}$–$G_{655})\cdot(C_{1158}$–$G_{1163})$ and $(G_{663}$–$C_{666})\cdot(G_{1025}$–$C_{1028})$ organize the central domain into discreet segments. This domain in *E. coli* 16S rRNA has been shown to interact with ribosomal proteins S_6, S_8, S_{15}, and S_{18} (Gregory et al., 1984). The prokaryotic S_8 binding domain, however, is absent in eukaryotic small ribosomal RNA (Gutell et al., 1985) and appears to have been replaced by the large insertion element 667–909 in rabbit 18S rRNA. The secondary structure shown for this element in Figure 7 is similar to that recently proposed by Choi (1985), which was found compatible with two cleavage sites generated by partial digestion of rat liver 40S subunits with T_1 ribonuclease. One of the base-paired fragments recently obtained by partial digestion of *Xenopus laevis* 18S rRNA (Atmadja et al., 1984) was a duplex corresponding to helix (746–762)·(773–789). This information in conjugation with our chemical modification data lends support for the proposed structure in Figure 7. Helix $(U_{859}$–$G_{869})\cdot(C_{880}$–$G_{890})$, however, contains three strongly reactive adenine residues (A_{860}, A_{861}, and A_{868}). Two of these, A_{860} and A_{861}, are adjacent to G·U wobble base pairs which might account for their reactivity toward dimethyl sulfate. The additional modification of A_{868}, however, suggests that this helix may only be marginally stable in the naked rRNA.

Helix $(U_{910}$–$A_{934})\cdot(U_{1000}$–$A_{1023})$, which is analogous to the prokaryotic S_{15} binding domain, also appears partially accessible to chemical structure probes. As summarized in Table IV, reactivities of A_{916}, A_{922}, A_{934}, and A_{1023} might be rationalized by poor nearest-neighbor stacking. A_{915}, A_{1019}, and A_{1020}, however, appear to be stably stacked within the two-dimensional representation of this phylogenetically proven helix (Chan et al., 1984). Additional modification of C_{1002} and C_{1003}, as well as their base-paired counterparts G_{931} and G_{932} observed in Figure 6, supports the notion that this helix is partially unstable in the naked 18S rRNA. The region 1085–1180 was found to be very resistant to chemical modification toward DEPC, suggesting that it is engaged in tertiary interactions not visible by this method. Interestingly, this region has recently been implicated in a possible eukaryotic "switch" structure in *Xenopus laevis* 18S rRNA (Atmadja et al., 1984).

3′-Major Domain 1203–1683. The 3′-terminal region is organized by the helix $(G_{1203}$–$A_{1210})\cdot(U_{1676}$–$U_{1683})$ into a major domain $(G_{1203}$–$U_{1683})$ and a minor domain $(G_{1684}$–$A_{1858})$. In 16S rRNA, the sequence 1050–1070 (1327–1341 in rabbit), which is normally base paired with 1190–1210 (1477–1490



DEPC OR Me_2SO_4 MODIFICATION

● STRONG; o WEAK

FIGURE 7: Secondary structure model for rabbit 18S rRNA. (I, II, and III) Primer binding regions; ("A" and "B") RNase H + *Pst*I directed cleavage sites; (1, 2, and 3) RNase T_1 cleavage sites; (Ψ, Am, Um, Cm, and Gm) naturally occurring modified bases. (●) Strongly reactive toward DEPC or Me_2SO_4; (o) weakly reactive toward DEPC. The inset figure is a schematic diagram of the secondary structure of rabbit 18S rRNA. Structure model includes corrections of the originally published sequence (Connaughton et al., 1984a).

Table IV: Sites of Strongly Reactive Adenine Residues Observed To Be Base Paired within the Phylogenetically Derived Secondary Structure Model for Rabbit 18S rRNA[a]

base pair	favorably stacked	noncanonical	adjacent to bulged nucleotide	adjacent to internal or hairpin loop	adjacent to G·U pair	base pair	favorably stacked	noncanonical	adjacent to bulged nucleotide	adjacent to internal or hairpin loop	adjacent to G·U pair
A_{38}·U_{552}				*		A_{934}·U_{1000}				*	
A_{77}·U_{63}	*					A_{976}·U_{942}	*				
A_{107}·U_{344}	*					A_{1019}·U_{914}	*				
A_{228}·U_{276}				*		A_{1020}·U_{913}	*				
A_{361}·U_{383}				*		A_{1023}·U_{910}				*	
A_{379}·U_{365}	*					A_{11922}·U_{15}				*	
A_{404}·U_{413}	*					A_{1195}·U_{12}				*	
A_{405}·U_{412}				*		A_{1204}·U_{1682}					*
A_{429}·U_{444}			*			A_{1205}·U_{1681}	*				
Am_{458}·U_{451}			*			A_{1219}·U_{1634}					*
A_{476}·U_{504}				*		A_{1278}·U_{1311}			*		
A_{479}·U_{501}	*					A_{1328}·U_{1490}					*
A_{535}·U_{528}	*					A_{1349}·U_{1354}				*	
A_{545}·U_{520}			*			A_{1417}·U_{1422}				*	
A_{554}·Ψ_{36}	*					A_{1434}·U_{1404}				*	
A_{589}·U_{627}				*		A_{1435}·U_{1403}	*				
A_{630}·U_{586}	*					A_{1442}·U_{1396}	*				
A_{631}·U_{585}				*		A_{1478}·U_{1338}			*		
A_{633}·U_{31}				*		A_{1480}·U_{1336}			*		
A_{638}·U_{26}	*					A_{1481}·U_{1335}	*				
A_{814}·G_{848}		*				A_{1538}·U_{1577}					*
A_{826}·U_{840}				*		A_{1579}·U_{1536}				*	
A_{860}·U_{889}					*	A_{1632}·U_{1221}				*	
A_{861}·U_{888}					*	A_{1642}·U_{1664}				*	
A_{868}·U_{881}	*					A_{1703}·U_{1810}	*				
A_{915}·U_{1018}	*					A_{1710}·G_{1803}		*			
A_{916}·U_{1017}			*			A_{1811}·U_{1702}				*	
A_{922}·U_{1011}					*	(total) 55	19	2	8	19	7

[a] Asterisk indicates neighborhood of reactive base-paired adenine residue.

in rabbit), was also found in association with the sequence 385–400 (431–446 in rabbit) (Glotz & Brimacombe, 1980). It was proposed that these two regions may possibly undergo a conformational rearrangement during the functioning of the 30S subunit (Brimacombe et al., 1983). Furthermore, a mutation at position 1192 in 16S rRNA, resulting in conversion of a G·C base pair to a G·U base pair, generates resistance to the aminoglycoside antibiotic spectinomycin (Sigmund et al., 1984). Such data support the view that this region plays an important role during prokaryotic protein synthesis, perhaps via an intramolecular switch mechanism (Sigmund et al., 1984). We have found the corresponding helix in rabbit 18S rRNA, (G_{1320}–U_{1339})·(A_{1477}–U_{1498}), partially accessible toward DEPC modification. As indicated in Figure 7, A_{1328}, A_{1478}, A_{1480}, and A_{1481} are strongly modified. A_{1328} is stacked adjacent to a U·G base pair, while A_{1478} and A_{1480} are next to two bulged nucleotides, which could account for their reactivity. Weaker modifications were also observed with DEPC at residues G_{1326} and G_{1494}, lending additional support for the notion that this is a metastable helix in naked 18S rRNA. Interestingly, two strongly reactive kethoxal modification sites corresponding to G_{1491} and G_{1494} of rabbit have been detected in yeast 40S subunits, suggesting that a portion of this helix is normally unpaired and accessible toward chemical probes. As seen in Figure 7, the putative complement to this region in the 5′ domain, 431–446, is also readily accessible toward DEPC. Further experiments, of course, are essential in determining whether this region plays a role in eukaryotic protein synthesis.

We have observed reproducible reactivity of residues A_{1204} and A_{1205} within the helix (G_{1203}–A_{1210})·(U_{1676}–U_{1683}) demarcating the 3′ domain, as seen in Figure 5. The accessibility of these adenines suggests a partial unpairing of this helix in the naked 18S rRNA. Loosening of the corresponding stem in naked 16S rRNA (Moazed et al., 1986a) as well as in "inactive" 30S subunits (Moazed et al., 1986b) has been noted previously.

3′-Minor Domain 1684–1858. Results on DEPC modification of the 3′-minor (functional) domain are essentially in agreement with the phylogenetic model, as well as with kethoxal modification data on yeast 40S subunits (Hogan et al., 1984) and yeast 18S rRNA (Douthwaite et al., 1983). We have also observed similar modification of these adenine residues within rabbit reticulocyte 40S subunits,[2] suggesting that the structure in this region of the naked 18S rRNA mirrors the "native" state. Modification of A_{1703} and A_{1811} residing at the end of the large hairpin loop suggests that pairing between U_{1702}–A_{1811} and A_{1703}–U_{1810} is disrupted. The absence of base pairing between nucleotides at positions 1701 and 1812 has recently been correlated with paromomycin resistance in *Tetrahymena* (Spangler & Blackburn, 1985), suggesting that subtle structural changes within this highly conserved region have a profound effect on translation. Indeed, we have observed subtle changes in reactivity for many residues within the regions 1686–1704 and 1810–1826 of the putative "decoding site" (Noller et al., 1986) when 40S subunits, 80S monosomes, and polyribosomes were compared.[2]

Implications of Adenine Reactivity Data. Table III compares the distribution of strongly reactive adenine residues within single-stranded and base-paired regions of the phylogenetically derived models for *E. coli* and rabbit small ribosomal RNA. It is clear that a relatively greater number of adenines were observed to be reactive within helical segments of rabbit 18S rRNA (55) compared with *E. coli* 16S rRNA (13). If these numbers are corrected by excluding adenines with weak predicted nearest-neighbor stacking interactions as summarized in Table IV, 19 strongly reactive sites remain

[2] H. Rubino, A. Rairkar, and R. E. Lockard, unpublished experiments.

within helices for rabbit compared with 2 for *E. coli*. While any discrepancy between chemical probing data and the phylogenetic model could be interpreted as evidence against the model, such differences could also be due to (i) additional specificities for the chemical probes, (ii) experimental disruption of the naked 18S rRNA secondary structures, or (iii) actual marginal stability of many helical segments.

Since DEPC modification is an indirect measure of single strandedness by preferentially reacting at the N-6 and N-7 positions (Peattie & Gilbert, 1980; Vincze et al., 1973), it is possible that some of the observed adenine reactivities are unrelated to the absence of base pairing. As indicated in Table I for the region 650–825, a good correlation was found between the reactivity of adenine residues toward both Me_2SO_4 and DEPC. However, there were 5 adenines out of 30 which were strongly reactive with DEPC, but unreactive toward Me_2SO_4. The specific accessibility of the N-7 and not the N-1 position of purines has been observed previously (Moazed et al., 1986b). Such reactivity may be related to how these residues are positioned within the tertiary conformation of the molecule and also may be a reflection of a localized change within the RNA A(A′) helix. Such conformational sensitivity for DEPC has been suggested previously by both the reactivity of purines within B-DNA (Herr et al., 1982) and their hyperreactivity within Z-DNA (Herr, 1985). It is noteworthy that there are 78 adenine residues which appear to be single stranded in the two-dimensional phylogenetic model but were found to be unreactive toward single-strand-specific chemical probes. It is possible that these nucleotides are actually engaged in complex tertiary interactions in the naked 18S rRNA.

Recent studies comparing the chemical reactivity of naked 16S rRNA with that in the 30S subunit indicate a general "tuning" of the structure upon ribosomal protein association to bring it into accord with the phylogenetically predicted model (Moazed et al., 1986; Stern et al., 1986). It is conceivable that removing ~31 ribosomal proteins from rabbit 18S rRNA destabilizes base pairing within some helical segments. Although we cannot completely rule out disruption of weak structural elements due strictly to our experimental procedures, the clustering of some reactive adenines within specific helical segments suggests that these regions may actually be marginally stable in the absence of protein or be flexible elements within the rRNA structure.

From studies with *E. coli* 30S subunits, there appears to be extensive protection of conserved, unpaired adenine residues, suggesting that they may play an important role in the assembly process (Moazed et al., 1986a). The direct involvement of unpaired adenines as well as other residues within bulged nucleotides, bulged loops, and internal loops has already been implicated in ribosomal protein S_4 binding (Stern et al., 1986), S_8 binding (Gregory & Zimmerman, 1986), and L_{18} protein binding to *E. coli* 5S RNA (Christiansen & Garrett, 1985). The apparent requirement for such helical irregularities may be a reflection of a general difficulty for many proteins to "read" RNA A(A′) helices (Seeman et al., 1976; Bubienko et al., 1983; Kean et al., 1985; Stern et al., 1986). Localized helical deformations may, therefore, be essential for specific protein recognition. It has been proposed that such irregularities may be subtle and can arise from sequence-specific base stacking irregularities (Bubienko et al., 1983). We tend to favor the interpretation that the observed differences in adenine reactivity may reflect both obvious and subtle irregularities of many of the helical segments in mammalian 18S rRNA, as compared with prokaryotic 16S rRNA. This notion is consistent with the greater number of ribosomal proteins which must eventually associate with 18S rRNA, as well as the possible unique roles 18S rRNA might additionally play during eukaryotic protein biosynthesis.

ACKNOWLEDGMENTS

We thank Dr. Robert Crouch (NIH) for the kind gift of *E. coli* RNase H and Vijay Khadse and James Geddis for technical assistance. We thank Dr. Kathy Scotto and José Walewski for assistance in densitometry and Dr. Andrew Seddon for performing amino acid analyses. We also thank Lucille Morgan and Terri McKernan for preparing the manuscript.

REFERENCES

Atmadja, J., Brimacombe, R., & Maden, B. E. H. (1984) *Nucleic Acids Res. 12*, 2649–2667.

Bidlingmeyer, B. A., Cohen, S. A., & Tarvin, T. L. (1984) *J. Chromatogr. 336*, 93–104.

Blobel, G., & Sabatini, D. (1971) *Proc. Natl. Acad. Sci. U.S.A. 68*, 390–394.

Brimacombe, R., Maly, P., & Zwieb, C. (1983) *Prog. Nucleic Acid Res. Mol. Biol. 28*, 1–48.

Bubienko, E., Cruz, P., Thomason, J. F., & Borer, P. N. (1983) *Prog. Nucleic Acid Res. Mol. Biol. 30*, 41–90.

Chan, Y.-L., Gutell, R., Noller, H. F., & Wool, I. G. (1984) *J. Biol. Chem. 259*, 224–230.

Choi, Y. C. (1985) *J. Biol. Chem. 260*, 12769–12772.

Christiansen, J., & Garrett, R. A. (1985) in *Structure, Function, and Genetics of Ribosomes* (Hardesty, B., & Kramer, G., Eds.) pp 253–269, Springer-Verlag, New York.

Connaughton, J. F., Rairkar, A., Lockard, R. E., & Kumar, A. (1984a) *Nucleic Acids Res. 12*, 4731–4745.

Connaughton, J. F., Kumar, A., & Lockard, R. E. (1984b) in *Eukaryotic Gene Expression* (Kumar, A., Ed.) pp 203–221, Plenum Press, New York.

Douthwaite, S., Christensen, A., & Garrett, R. A. (1983) *J. Mol. Biol. 169*, 249–279.

Freier, S. M., Kierze, K. R., Jaeger, J. A., Sugimoto, N., Caruthers, M. H., Neilson, T., & Turner, D. H. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 9373–9377.

Glotz, C., & Brimacombe, R. (1980) *Nucleic Acids Res. 8*, 2377–2395.

Gregory, R. J., & Zimmermann, R. A. (1986) *Nucleic Acids Res. 14*, 5761–5776.

Gregory, R. J., Zeller, M. L., Thurlow, D. L., Gourse, R. L., Stark, M. J. R., Dahlberg, A. E., & Zimmermann, R. A. (1984) *J. Mol. Biol. 178*, 287–302.

Gutell, R. R., Weiser, B., Woese, C. R., & Noller, H. F. (1985) *Prog. Nucleic Acid Res. Mol. Biol. 32*, 155–216.

Herr, W. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 8009–8013.

Herr, W., Corbin, V., & Gilbert, W. (1982) *Nucleic Acids Res. 10*, 6931–6944.

Hogan, J. J., Gutell, R. R., & Noller, H. F. (1984) *Biochemistry 23*, 3322–3330.

Inoue, T., & Cech, T. R. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 648–652.

Jagus, R., Anderson, W. F., & Safer, B. (1981) *Prog. Nucleic Acid Res. Mol. Biol. 25*, 127–185.

Kean, J. M., White, S. A., & Draper, D. E. (1985) *Biochemistry 24*, 5062–5070.

Kozak, M. (1983) *Microbiol. Rev. 47*, 1–45.

Lane, D. J., Pace, B., Olsen, G. J., Stahl, D. A., Sogin, M., & Pace, N. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 6955–6959.

Lempereur, L., Nicoloso, M., Rielhl, N., Ehresmann, B., & Bachellerie, J. P. (1985) *Nucleic Acids Res. 13*, 8339–8357.

Lockard, R. E., & RajBhandary, U. L. (1976) *Cell (Cambridge, Mass.) 9*, 747-760.
Lockard, R. E., Alzner-Deweerd, B., Heckman, J. E., MacGee, J., Tabor, M. W., & RajBhandary, U. L. (1978) *Nucleic Acids Res. 5*, 37-56.
Lockard, R. E., Connaughton, J. F., & Kumar, A. (1982) *Nucleic Acids Res. 10*, 3445-3457.
Maly, P., & Brimacombe, R. (1983) *Nucleic Acids Res. 11*, 7263-7286.
Mandal, C., Kallenbach, N. R., & Englander, S. W. (1979) *J. Mol. Biol. 135*, 391-411.
Mizushima, S., & Nomura, M. (1970) *Nature (London) 226*, 1214-1217.
Moazed, D., Stern, S., & Noller, H. F. (1986a) *J. Mol. Biol. 187*, 399-416.
Moazed, D., Van Stolk, B. J., Douthwaite, S., & Noller, H. F. (1986b) *J. Mol. Biol. 191*, 483-493.
Peattie, D. A., & Gilbert, W. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 4679-4682.
Pleij, C. W. A., Rietveld, K., & Bosch, L. (1985) *Nucleic Acids Res. 13*, 1717-1731.
Pörschke, D. (1974) *Biophys. Chem. 2*, 97-101.
Qu, L.-H., Michot, B., & Bachellerie, J. P. (1983) *Nucleic Acids Res. 11*, 5903-5920.
Seeman, N. C., Rosenberg, J. M., & Rich, A. (1976) *Proc. Natl. Acad. Sci. U.S.A. 73*, 804-808.
Sigmund, C. D., Ettayebi, M., & Morgan, E. A. (1984) *Nucleic Acids Res. 12*, 4653-4663.
Spangler, E. A., & Blackburn, E. H. (1985) *J. Biol. Chem. 260*, 6334-6340.
Stern, S., Wilson, R. C., & Noller, H. F. (1986) *J. Mol. Biol. 192*, 101-110.
Tinoco, I., Borer, P. N., Dengler, B., Levine, M. D., Uhlenbeck, O. C., Crothers, D. M., & Gralla, J. (1973) *Nature (London), New Biol. 246*, 40-41.
Van Stolk, B. J., & Noller, H. F. (1984) *J. Mol. Biol. 180*, 151-177.
Vincze, A., Henderson, R. E. L., McDonald, J. J., & Leonard, N. J. (1973) *J. Am. Chem. Soc. 95 (8)*, 2677-2682.
Wittman, H. G. (1982) *Annu. Rev. Biochem. 51*, 35-65.
Wool, I. G. (1980) in *Ribosomes, Structure, Function, and Genetics* (Chambliss, G., Ed.) pp 797-824, University Park Press, Baltimore, MD.

Primary Structure Homologies between Two Zinc Metallopeptidases, the Neutral Endopeptidase 24.11 ("Enkephalinase") and Thermolysin, through Clustering Analysis†

T. Benchetrit,‡ V. Bissery,§ J. P. Mornon,§ A. Devault,∥ P. Crine,∥ and B. P. Roques*,‡

Département de Chimie Organique, U 266 INSERM, UA 498 CNRS, UER des Sciences Pharmaceutiques et Biologiques, 4 avenue de l'Observatoire, 75006 Paris, France, Laboratoire de Minéralogie et Cristallographie, Université Paris VI, 4 Place Jussieu, 75005 Paris, France, and Département de Biochimie, Université de Montréal, Montréal, Québec, Canada H3C 3J7

Received April 28, 1987; Revised Manuscript Received August 20, 1987

ABSTRACT: Analogies in the sequences of two related zinc metallopeptidases, the bacterial thermolysin (316 amino acids) and the recently cloned neutral endopeptidase 24.11 ("enkephalinase", 749 amino acids), have been demonstrated by a hydrophobic cluster analysis method derived from the Lim theory. Two sequence alignments are proposed for the entire primary structure of thermolysin and the C-terminal part of endopeptidase 24.11. Except for an arginine residue, all the amino acids involved in the active site of thermolysin have been retrieved in both models of endopeptidase 24.11 within conserved clustered structures. The first model is characterized by a deletion of the Ca^{2+}-binding coil present in thermolysin and the second by replacement of this coil by two α-helices. In both models an Arg residue can be located in the active site of the neutral endopeptidase.

Neutral endopeptidase, NEP (EC 3.4.24.11), initially characterized in the kidney brush border (Kerr & Kenny, 1974) was also subsequently shown to be present in the brain (Malfroy et al., 1978). The enzyme, localized in the vicinity of opioid receptors (Waksman et al., 1986), is involved in the physiological inactivation of the endogenous opioid peptides methionine- and leucine-enkephalin (Roques et al., 1980), and is therefore often called "enkephalinase".

NEP belongs to an important physiological group of Zn metallopeptidases (Vallee & Galdes, 1984), such as angiotensin converting enzyme, ACE (Das & Soffer, 1975), and collagenase (Seltzer et al., 1977). Two of the enzymes from this group, carboxypeptidase A (Rees et al., 1983) and thermolysin, TLN (Holmes & Matthews, 1982), have been crystallized, alone or in the presence of various substrates or inhibitors. A simplified model of the active site of Zn metallopeptidases, derived from these crystallographic studies, was used by Ondetti et al. (1977) for the rational design of ACE inhibitors such as captopril, which act as potent antihypertensive drugs. Taking into account the similarities in the enzymatic activity of NEP and TLN (Kerr & Kenny, 1974; Llorens et al., 1980; Benchetrit et al., 1987), potent inhibitors of NEP have also been developed, leading to a new class of potential analgesic drugs (Chipkin, 1986; Roques & Fournié-Zaluski, 1986).

Recently, the amino acid sequence of NEP (749 amino acids), deduced from its complementary DNA (Devault et al.,

† This research is supported by grants from Université René Descartes, the Centre National de la Recherche Scientifique, and the Institut National de la Santé et de la Recherche Médicale.
‡ UER des Sciences Pharmaceutiques et Biologiques.
§ Université Paris VI.
∥ Université de Montréal.

0006-2960/88/0427-0592$01.50/0