Kurland, C. G. (1977) in *Molecular Mechanisms of Protein Biosynthesis* (Weissbach, H., & Pestka, S., Eds.) pp 81-116, Academic Press, New York.
Lake, J. A. (1976) *J. Mol. Biol. 105*, 131-159.
Liljas, A. (1982) *Prog. Biophys. Mol. Biol. 40*, 161-228.
Lowry, O. H., Rosebrough, N. J., Farr, A. L., & Randall, R. J. (1951) *J. Biol. Chem. 193*, 265-275.
Maassen, J. A., Thielen, A. P. G. M., & Möller, W. (1983) *Eur. J. Biochem. 134*, 327-330.
Maassen, J. A., Schop, E. N., & Möller, W. (1984) *Eur. J. Biochem. 138*, 131-134.
Matheson, A. T., Möller, W., Amons, R., & Yaguchi, M. (1980) in *Ribosomes, Structure, Function and Genetics* (Chambliss, G., Craven, G. R., Davies, J., Davis, K., Kahan, L., & Nomura, M., Eds.) pp 297-332, University Park Press, Baltimore.
Möller, W. (1974) in *Ribosomes* (Nomura, M., Tissières, A., & Lengyel, P., Eds.) pp 711-731, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
Möller, W., Castleman, H., & Terhorst, C. P. (1970) *FEBS Lett. 8*, 192-196.
Möller, W., Groene, A., Terhorst, C., & Amons, R. (1972) *Eur. J. Biochem. 25*, 5-12.
Möller, W., Schrier, P. I., Maassen, J. A., Zantema, A., Schop, E., Reinalda, H., Cremers, A. F. M., & Mellema, J. E. (1983) *J. Mol. Biol. 162*, 553-573.
Österberg, R., Sjöberg, B., Liljas, A., & Pettersson, I. (1976) *FEBS Lett. 66*, 48-51.
Schiller, P. W. (1975) in *Biochemical Fluorescence, Concepts* (Chen, R. F., & Edelhoch, H., Eds.) Vol. 1, pp 285-303, Marcel Dekker, New York.
Schop, E. N., & Maassen, J. A. (1982) *Eur. J. Biochem. 128*, 371-375.
Schrier, P. I., Maassen, J. A., & Möller, W. (1973) *Biochem. Biophys. Res. Commun. 53*, 90-98.
Strycharz, W. A., Nomura, M., & Lake, J. A. (1978) *J. Mol. Biol. 126*, 123-140.
Subramanian, A. R. (1975) *J. Mol. Biol. 95*, 1-8.
Terhorst, C., Möller, W., Laursen, R., & Wittmann-Liebold, B. (1972) *FEBS Lett. 28*, 325-328.
Traut, R. R., Lambert, J. M., Boileau, J. M., & Kenny, J. W. (1980) in *Ribosomes, Structure, Function and Genetics* (Chambliss, G., Craven, G. R., Davies, J., Davis, K., Kahan, L., & Nomura, M., Eds.) pp 89-110, University Park Press, Baltimore.
Van Agthoven, A. J., Maassen, J. A., Schrier, P. I., & Möller, W. (1975) *Biochem. Biophys. Res. Commun. 64*, 1184-1191.
Wong, K.-P., & Paradies, H. H. (1974) *Biochem. Biophys. Res. Commun. 61*, 178-184.
Yamamoto, K., Sekine, T., & Kanaoka, Y. (1977) *Anal. Biochem. 79*, 83-94.
Zantema, A., Maassen, J. A., Kriek, J., & Möller, W. (1982a) *Biochemistry 21*, 3069-3076.
Zantema, A., Maassen, J. A., Kriek, J., & Möller, W. (1982b) *Biochemistry 21*, 3077-3082.

Probing the Conformation of 18S rRNA in Yeast 40S Ribosomal Subunits with Kethoxal†

James J. Hogan,‡ Robin R. Gutell, and Harry F. Noller*

ABSTRACT: Yeast 40S ribosomal subunits have been reacted with kethoxal to probe the conformation of 18S rRNA. Over 130 oligonucleotides were isolated by diagonal electrophoresis and sequenced, allowing identification of 48 kethoxal-reactive sites in the 18S rRNA chain. These results generally support a secondary structure model for 18S rRNA derived from comparative sequence analysis. Significant reactivity at positions 1436 and 1439, in a region shown to be base paired by comparative analysis, lends support to the earlier suggestion [Chapman, N. M., & Noller, H. F. (1977) *J. Mol. Biol 109*, 131-149] that part of the 3'-major domain of 16S-like rRNAs may undergo a biologically significant conformational rearrangement. Modification of positions in 18S rRNA analogous to those previously found for *Escherichia coli* 16S rRNA argues for extensive structural homology between 30S and 40S ribosomal subunits, particularly in regions thought to be directly involved in translation.

Ribosomes from the cytoplasm of eucaryotes are distinguishable from those of procaryotic cells or organelles by their size, shape, and chemical composition (Wool, 1979). They are usually referred to as 80S ribosomes, dissociable into 40S and 60S subunits, in contrast with the 70S, 30S, and 50S subunits of procaryotes. They possess morphological characteristics that distinguish them from eubacterial or archaebacterial ribosomes (Lake et al., 1982) and have a higher content of protein relative to RNA (Wool, 1979). The large ribosomal RNAs of 80S ribosomes are also significantly larger than their eucaryotic counterparts; 18S rRNAs contain about 1800 nucleotides in contrast to about 1540 for procaryotes, while 26S or 28S rRNAs can have from 500 to over 1500 more nucleotides than the 2900 found in procaryotic 23S rRNAs (Noller, 1984). Thus, 70S and 80S ribosomes seem to represent two distinct classes.

Do the two classes represent fundamental differences in ribosome architecture, mechanism, and evolution, or are there underlying similarities between them? Sequence homology between yeast (Rubtsov et al., 1980), maize (Messing et al., 1984), *Xenopus* (Salim & Maden, 1981), *Dictyostelium*

† From the Thimann Laboratories, University of California, Santa Cruz, California 95064. *Received November 9, 1983.* This work was supported by Grant GM-17129 from the National Institutes of Health.
‡ Present address: Gen-Probe, Inc., La Jolla, CA 92121.

0006-2960/84/0423-3322$01.50/0

(McCarroll et al., 1983), and rat liver (Chan et al., 1983) 18S rRNAs and *Escherichia coli* 16S rRNA (Brosius et al., 1978; Carbon et al., 1979) suggests conservation of common structural features between the two classes of ribosomes. This is borne out by the fundamental similarity between procaryotic and eucaryotic rRNA secondary structure (Woese et al., 1983); the additional nucleotides in eucaryotic rRNA are for the most part accounted for by insertions at a few positions.

In this study, we explore the topography of 18S rRNA in 40S ribosomal subunits from the yeast *Saccharomyces cerevisiae* by means of the single-strand-specific RNA probe kethoxal (Noller, 1974). This approach serves two purposes; we obtain experimental evidence bearing on a secondary structure model for 18S rRNA derived by comparative sequence analysis (Woese et al., 1983), and we compare the accessibility of specific kethoxal-reactive sites in the 18S rRNA with that of their homologues in *E. coli* 16S rRNA (Noller, 1974; Hogan & Noller, 1978). Since many of the latter are thought to participate directly in translation (Chapman & Noller, 1977; Herr et al., 1979; Brow & Noller, 1983), it is reasonable to expect that corresponding positions in 18S rRNA should also be accessible, if indeed their accessibility implies availability for interaction with functional ligands (e.g., tRNA) that are common to the translational apparatus of all organisms. We have located 48 kethoxal-reactive guanines in the 18S rRNA sequence. There is good agreement between these data and the proposed secondary structure model. Furthermore, sites that are homologous to the reactive sites in *E. coli* 30S subunits are also reactive in yeast 40S subunits, implying that the rRNAs of the two classes of ribosomes have very similar structural organizations on their exposed surfaces, including those regions that have been implicated in ribosomal function.

Materials and Methods

Buffers. The compositions of the buffers used are as follows: (buffer A) 50 mM Tris-HCl,[1] 10 mM KCl, 10 mM β-mercaptoethanol, and 12.5 mM $MgCl_2$, pH 7.65; (buffer B) 50 mM Tris-HCl, 880 mM KCl, 20 mM β-mercaptoethanol, and 12.5 mM $MgCl_2$, pH 7.65; (buffer C) 50 mM Tris-HCl, 1.75 M KCl, 20 mM β-mercaptoethanol, and 20 mM $MgCl_2$, pH 7.65; (buffer D) 100 mM potassium cacodylate and 10 mM $MgCl_2$, pH 8.0; (buffer E) 200 mM NaOAc, 10 mM EDTA, and 1 mM sodium borate, pH 5.0.

Saccharomyces cerevisiae A364A was grown at 30 °C in low-phosphate media for ^{32}P labeling and in methionine-free media for specifically labeling the methylated nucleotides with L-[*methyl*-^{14}C]methionine as described by Rubin (1975). The cultures were harvested at A_{660} = 3.0 by centrifugation at 8000 rpm for 15 min in a Sorvall SS-34 rotor. Cell lysis with minimal RNase damage to ribosomes was accomplished by vortexing the cell pellet for 3 min with 3 volumes of buffer A and 2 volumes of 0.3-mm glass beads. Glass beads and cell wall material were removed by centrifugation at 14000 rpm for 20 min in a Sorvall SS-34 rotor. The supernatant was then layered onto a 4-mL cushion of 2.0 M sucrose in buffer A. Ribosomes were pelleted by centrifugation at 45000 rpm for 24 h in an International A321 rotor. The pellet was resuspended in 0.5 mL of buffer A and then mixed with an equal volume of buffer C. The resultant increase in the KCl concentration to 0.88 M induces 80S ribosomes to dissociate into 40S and 60S subunits. For separation of subunits, a 15–30% sucrose gradient in buffer B was centrifuged in a Beckman SW27 rotor at 24000 rpm for 14 h at 10 °C. Incubation of the subunits for 3 min at room temperature prior to layering on the gradient and running the gradient at 10 °C help to inhibit dimerization of 40S subunits. The 40S subunits were precipitated with 0.7 volume of ethanol at 0 °C and pelleted by centrifugation in a Sorvall HB-4 rotor at 10000 rpm for 30 min.

The 40s subunits were immediately resuspended in 0.8 mL of buffer D to which 0.1 volume of 37 mg/mL kethoxal in 20% ethanol was added and incubated for 1 h at 37 °C. Kethoxal-modified 40S subunits were precipitated with ethanol and resuspended in buffer E. The 18S rRNA was phenol extracted 3 times, washed with ether, and then precipitated with 3 volumes of absolute ethanol at −60 °C. Precipitated rRNA was redissolved in 0.2 mL of water and lyophilized.

Diagonal electrophoresis was performed as previously described (Noller, 1974). RNase T_1 oligonucleotides originating from the kethoxal-reactive regions appear as off-diagonal spots, while the bulk of the material remains on the diagonal line. All of the off-diagonal oligonucleotides were subjected to at least one additional purification step prior to sequence analysis. Spots running faster than the blue dye (xylene cyanol FF) were subjected to electrophoresis on DEAE paper at pH 3.5, and spots running slower than the blue dye were purified by thin-layer homochromatography on DEAE-cellulose plates with homomix c (Barrell, 1971). Spots running close to the origin were purified by a "phosphatase-diagonal" (Dahlberg, 1968) step: A strip containing these spots was cut out and moistened with a 1 mg/mL solution of bacterial alkaline phosphatase (Sigma) and incubated for 2 h at 37 °C in a humid temperature-controlled desiccator. The strip was washed down with water, dried, sewn onto a sheet of DEAE paper, and reelectrophoresed as before. This allowed resolution of slow-running 3′-phosphorylated oligomers from non-phosphorylated ones and generally resulted in purification sufficient to allow sequencing. Phosphorylated oligonucleotides could be distinguished by their increased electrophoretic mobility in the second dimension of this procedure.

Sequence analysis was done essentially as described by Barrell (1971). Purified oligonucleotides were subjected to digestion with RNases A, T_2, and U_2, as well as snake venom phosphodiesterases, and the products analyzed by paper electrophoresis (Barrell, 1971). Posttranscriptionally modified nucleotides (either ^{32}P or [^{14}C]methyl labeled) were obtained by RNase T_2 digestion and identified by two-dimensional thin-layer chromatography.

Oligonucleotides were located in the yeast 18S rRNA gene sequence (Rubtsov et al., 1980) by comparison with a catalog of predicted RNase T_1 products obtained by computer analysis (B. Weiser, unpublished results). In cases where a given oligonucleotide sequence occurs many times in the 18S rRNA sequence, the correct site of modification could often be deduced by consideration of mobility shifts. For example, if the 5′ partner of an oligomer is Gp, its second dimension mobilitiy will be double that of the first dimension; if it is UGp, the mobility increase will be 4-fold. U or G each give a 2-fold mobility increase; three A residues or 8–10 C's have approximately the effect or a U or G. Thus, certain oligomer placements may be confidently eliminated if they are predicted to exhibit mobility shifts greater than the experimentally observed one. Multiple hits may also be inferred by similar reasoning. For example, a double kethoxal hit on adjacent guanines will increase the mobility shift compared with that observed with a single hit. Interpretation is usually rather

[1] Abbreviations: Tris-HCl, tris(hydroxymethyl)aminomethane hydrochloride; EDTA, ethylenediaminetetraacetic acid.

Table I: Comparative Sequence Evidence for the Eukaryote-Specific Helix 483–491/496–504[a]

Dictyostelium	CAAUACCUAUCCUUUUU·GGAGGG··CAAUUG
yeast	CGAUACAGGGCCCAUUCGGGU·CUUGUAAUUG
maize	CAAUACCGGGCGCGUUAGUGU·CUGGUAAUUG
Xenopus and rat[b]	CAAUACAGGACUCUUUCGAGGCCCUGU·AUUG

[a] Sequences are from Rubtsov et al. (1980), Salim & Maden (1981), Chan et al. (1983), McCarroll et al. (1983), and Messing et al. (1984). Overlining indicates complementary sequences believed to be involved in intramolecular base pairing. [b] The sequences of *Xenopus* and rat are identical in this region.

clear, since the observed mobilities differ by at least a factor of 2 in such cases.

Results

Provisional Changes in the Yeast 18S rRNA Sequences. Comparison of the yeast 18S rRNA sequence (Rubtsov et al., 1980) with the over 20 published 16S-like rRNA sequences reveals some apparent deletions in the yeast sequence. Three of these are sufficiently anomalous that we have made provisional alterations in the sequence for the purpose of this paper. At least one of these discrepancies has previously been noted (Salim & Maden, 1981). The specific changes and rationales for making them follow. In each case, the numbering refers to that of the changed sequence, as shown in Figure 1.

(*1*) *Positions 987–996.* There is a 10-nucleotide gap in the published sequence of yeast 18S rRNA with respect to all-published 16S-like rRNAs following position 986. Very likely, this can be accounted for by a missing 10-nucleotide *Sau*3A restriction fragment here, as suggested by Salim & Maden (1981). We suggest the provisional sequence (G)AUCGAAGACG(AUC)$_{999}$ for the missing region, on the basis of the consensus of the other five available eucaryotic 18S rRNA sequences, which all have identical sequences in this region except for the initial A residue. The presence of A at position 987 is suggested both by the *Sau*3A recognition sequence and by the fact that the other lower eucaryotes have A at this position.

(*2*) *Positions 1040–1041.* A gap of two nucleotides seems to exist in this region of the yeast sequence, on the basis of alignment of the conserved flanking sequences CCGACU$_{1037}$ and GAUC$_{1044}$, present with only minor variations in all eucaryotic 18S rRNAs. The missing nucleotides are likely to be GG, found in the other lower eucaryotes and in maize 18S rRNA. Again, this apparent deletion occurs at the junction of *Sau*3A restriction fragments.

(*3*) *Position 1181.* This apparent deletion occurs in the invariant eucaryotic sequence GCGGCUUAAUUUGACUCAAC$_{1195}$. The oligomer CUUAAUUUG is found in RNase T$_1$ digests of yeast 18S rRNA, whereas the sequence CUAAUUUG is not found (C. Woese, personal communication). Furthermore, all known 16S-like rRNAs, except for the mammalian mitochondria, have the same number of nucleotides in this region. We have therefore placed a U residue at position 1181.

Secondary Structure Model for Yeast 18S rRNA. We have developed a secondary structure model for yeast 18S rRNA on the basis of comparative sequence analysis. Much of the structure is similar to that of bacterial 16S rRNA, and the evidence bearing on these features of the structure has been presented by Woese et al. (1983). The structure shown in Figure 1 extends the analysis to include most of the yeast 18S rRNA sequence. Two regions, around positions 520 and 630, are shown as unstructured (a and b in Figure 1), because the high phylogenetic variability of these regions of the available eucaryotic sequences is not consistent with any obvious univeral base-pairing scheme. Other regions that differ from the 16S structure tend simply to be extensions of helices whose procaryotic counterparts are well characterized. These include, for example, the region 1060, 1360, and 1700 helices. Many of the helices show eucaryote-specific variations, such as introduction of mispairs or bulges into otherwise regular structures.

Table II: Helices in Common between Proposed Secondary Structure Models for Yeast 18S rRNA

	total helices	helices by domain[a] 5′	central	3′-major	3′-terminal
this work	69	27	13	21	8
Mankin et al. (1981)	38	11	10	11	6
Stiegler et al. (1981)	43	12	8	15	8
Zwieb et al. (1981)	33	7	8	12	6

[a] Domains are as defined by Noller & Woese (1981).

An example of a structural feature unique to eucaryotic 18S rRNA is the helix 483–491/496–504 (Figure 1). About 45 nucleotides are inserted in this region of eucaryotic RNAs, compared with bacterial RNAs. Comparative analysis of eucaryotic sequences from this region (Table I) reveals a helix supported by several compensating base-pair changes. The form and extent of this structure vary somewhat in detail between the five organisms listed in Table I. A-G pairing may occur in the *Dictyostelium*, *Xenopus*, and rat versions but not in yeast or maize.

Similar structure has been derived for rat 18S rRNA (Chan et al., 1984), and a comparison of these two eucaryotic structures provides a detailed illustration of these points. Secondary structural models for yeast 18S rRNA have been proposed earlier by Mankin et al. (1981), Stiegler et al. (1981), and Zwieb et al. (1981). The number of helices in common between the model shown in Figure 1 and the structures proposed by these other investigators is summarized in Table II.

Kethoxal Probe Experiments. Yeast 40S ribosomal subunits, either ^{32}P or [^{14}C]methyl labeled, were modified with kethoxal and analyzed by diagonal electrophoresis (Noller, 1974) in four independent experiments. A representative ^{32}P diagonal analysis is shown in Figure 2. Off-diagonal spots that were reproducibly observed in the four runs are identified in the schematic drawing. Following repurification of the oligonucleotides contained in each spot, they were each subjected to sequence analysis, as described under Materials and Methods. The sequences are summarized in Table III.

In all, over 130 oligomer sequences are listed. Most of these are too short to allow unambiguous placement in the 18S rRNA sequence (Rubtsov et al., 1980). Some of the oligomers can be placed precisely, either because they possess unique sequences or by a combination of sequence and mobility-shift data (see Materials and Methods). Where possible, their positions in the 18S rRNA sequence and the number of occurrences of each oligomer sequence in the 18S rRNA are

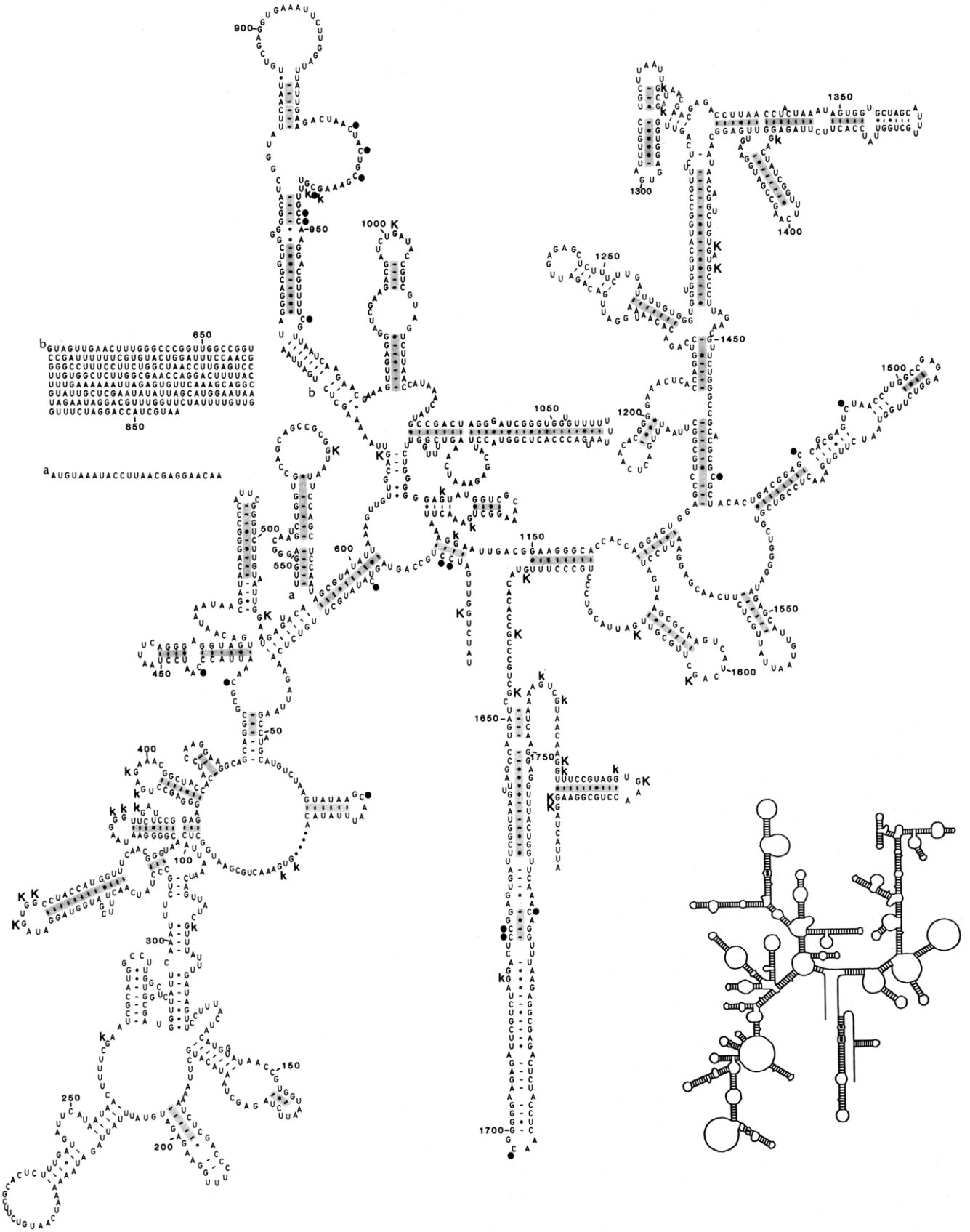

FIGURE 1: Secondary structure model for yeast 18S rRNA, derived by comparative sequence analysis (Woese et al., 1983). The sequence shown is that reported by Rubtsov et al. (1980) with some minor changes, as noted in the text. Major kethoxal-reactive sites are indicated by K; minor sites, by k (Table IV). Cytidines shown to be reactive toward dimethyl sulfate by Mankin et al. (1981) are indicated by (●). Additional residues in 18S rRNA not yet structured are indicated at ca. positions 540 (a) and 630 (b). Helices considered proven by comparative sequence analysis are shaded.

listed in Table III. In all, positions of 48 kethoxal-reactive guanines have been identified. The mobility shift for each spot is also listed in Table III; this value is a measure of the G+U content of the partner oligonucleotide(s) released by RNase T_1 digestion after the first electrophoresis dimension. Thus, in cases where there are relatively few occurrences of a given

Table III: Sequences of Oligonucleotides from Kethoxal-Reactive Sites in Yeast 18S rRNA[a]

oligonucleotide no.	sequence	no. of occurrences	mobility shift	position	oligonucleotide no.	sequence	no. of occurrences	mobility shift	position
1	UG_{OH}	26			43a	$AAACUG_{OH}$	1	≥4	81 + 83
2	UG_{OH}	26			43b	AACGp	3	≥4	
3	UG_{OH}	26			44	$AUAACG_{OH}$	1	4	1317 + 1319
4	UG_{OH}	26			45a (56b)	$AAACUG_{OH}$	1	2	83
5	UG_{OH}	26			45b	CAAGp	2	2	
6	Gp				45c	AG_{OH}	35		
7	Gp				46a	$UACACACCG_{OH}$	1	≥4	1631
8	Gp				46b (53b)	$UAACAAG_{OH}$	1	≥4	1759 + 1762
9a (25a)	Gp				46c (17d)	AAAGp	2	≥4	941 + 942
9b	UG_{OH}	26			47	UUG_{OH}	9	≥4	
10 (52)	Gp				48	$(U,\psi)G_{OH}$			
11	UCG_{OH}	7	2		49	UGp	26		
12	Gp				50a (79)	UGp	26		
13	CUG_{OH}	4	1+		50b	$U(U,C)G_{OH}$	5	3+	
14	Gp				51a	$(U,C)G_{OH}$	5	2+	
15 (32)	Gp			510	51b	$CUUG_{OH}$	2	2+	1603
16	CUG_{OH}	4	1		51c	UGp	26		
17a	$CUAG_{OH}$	2	≥4	1647	52 (10)	$UUCG_{OH}$	2	2	372 + 373
17b	AUG_{OH}	6	≥4		53a	UGp	26		
17c	CCGp	7	≥4		53b (46b)	(U,C)Gp			
17d (46c)	CGp	19			54	CUGp	4		
18a	AUG_{OH}	6	3		55	$UUAG_{OH}$			
18b	CGp	19			56a (35b)	$UAUG_{OH}$	1	1+	1113
19a	$UACCG_{OH}$	1	2	1651	56b (45a)	UGp	26		
19b	$CAUG_{OH}$	4	2		57 (28)	UGp	26		
20a	$CAUG_{OH}$	4	2		58 (20b)	UGp	26		
20b (58)	AUG_{OH}	6	2	1436	59a	$AU(C_{1-2}U)G_{OH}$			
21a	$CAUG_{OH}$	4	1		59b	UAGp	6	≥4	
21b	AUG_{OH}	6	1		59c	CUGp	4	≥4	
22a	$AUCG_{OH}$	1	1		60a (78)	CUGp	4	4	
22b	AAG_{OH}	8	1		60b	$U(U,C)G_{OH}$	5	4	
23	AGp	35			60c	$AU(C_2U)G_{OH}$	2	4	373 + 377
24a (24b)	$AAACG_{OH}$	1	4	396 + 398	61	$AU(C_2U)G_{OH}$	2	3	377
24b (24a)	AGp	35			62	$AUCUG_{OH}$	2	1	1682
25a (9)	$A^*A^*CCUG_{OH}$	1	3	1780 + 1782	63	(AU,C)Gp	5	≥4	
25b	$AAUG_{OH}$	3	3		64 (83a)	$AUUACG_{OH}$	1	3+	1612
26a	$AAUG_{OH}$	3	2+		65	$CCCUUAG_{OH}$	1	2+	1439
26b	$AACG_{OH}$	3	2+		66a (68a)	$CCUACCAUG_{OH}$	1	≥4	334 + 336 + 337
27a (27b)	$AAACG_{OH}$	1	2	398	66b	AUGp	6	≥4	
27b (27a)	AGp	35			67a	$AAUUG_{OH}$	1	1	1141
28 (57)	$A^*A^*CCUG_{OH}$	1	2	1782	67b	$ACUAUCG_{OH}$	1	1	1388
29	AAG_{OH}	8	2		68a (66a)	AUAGp	3	≥4	334 + 336 + 337
30a	$AAUG_{OH}$	3	1		68b	AAUGp	3	≥4	
30b	$AACG_{OH}$	3	1		69a	AAUGp	3	2+	
30c	AAG_{OH}	8	1		69b	AUAGp	3	2+	
31	$AAUG_{OH}$	3	1		70a	AUAGp	3	1	
32 (15)	$AAUG_{OH}$	3	1	510	70b	AAUGp	3	1	
33a	$AAUCG_{OH}$	1	≥4	264	71	UACACACCGp	1	≥4	1631 + 1640
33b	$A\Psi AG_{OH}$	(3)	≥4	123	72	$A^*ACAAUUG_{OH}$	1	2	539, 540
33c	CAGp	5	≥4		73a	$A^*ACAAUUG_{OH}$	1	1	540
33d	AGp	35			73b	$A\Psi CAGp$			
34a	$AUCAG_{OH}$	*b*			74	$UUAAA^*AAG_{OH}$	1	<1	616
34b	$AAUG_{OH}$	3	3+		75 (83b)	UAACAAGp	1	≥4	1769 + 1770
34c	CAGp	5	3+		76	$AUCAUUA_{OH}$	1	2	1794 + 1795
34d	AGp	35			77	$AUCAUUA_{OH}$	1	1	1795
35a	$AUACC^*G_{OH}$	1	2+	1001	78 (60a)	$AAACUUAAAG_{OH}$	1	2	1130
35b (56a)	AGp	35			79 (50a)	UACUGp	1	2	671 + 676
36	AAG_{OH}	8	2		80	UACUGp	1	1	676
37	AAGp	8	≥4		81	$AAAAAA\Psi U^*AG_{OH}$	1	1	751
38a	$AACN^*G_{OH}$				82a	$U^*AAUUCCAG_{OH}$	1	1	577
38b	AAGp	8	3+		82b	(p)UAUCUGp	1	1	6
39	AAGp	8	3		83a (64)	UUGp	9	≥4	
40	AAG_{OH}	8	1		83b (75)	$U(C_2U_2)G_{OH}$	3	≥4	1769 + 1770
41	AAGp	8	1		84a	UACUGp	1	≥4	
42	$AAAUG_{OH}$								

[a] Oligonucleotide numbers refer to Figure 2. Where purification gave rise to more than one oligonucleotide from a given spot, these are given by letters (e.g., 9a and 9b). The number of occurrences of each oligomer in the yeast 18S rRNA sequence (Rubtsov et al., 1980) was compiled by a computer program (B. Weiser, unpublished) with the sequence as shown in Figure 1. Mobility shifts were obtained by measurement of the autoradiographs. Positions of modified sequences were deduced as described in the text and by Noller (1974). "Partner" oligonucleotides are indicated in parentheses (see text). [b] Although the RNase T_1 oligomer AUCAG is not predicted from the published gene sequence for yeast 18S rRNA (Rubtsov et al., 1980), this oligomer has been found in T_1 digests of yeast 18S rRNA (C. Woese, personal communication).

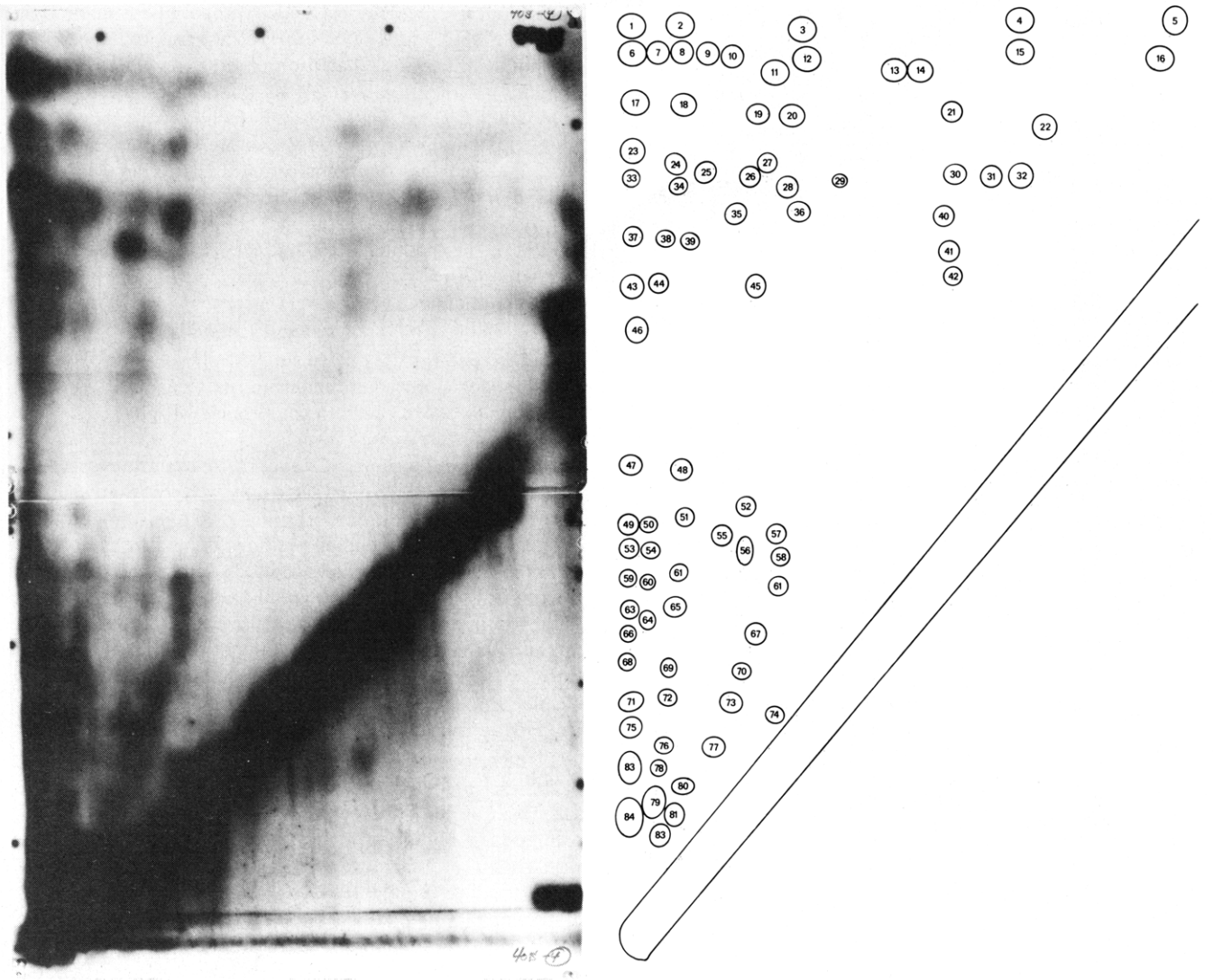

FIGURE 2: Diagonal electrophoresis of kethoxal-modified yeast 18S rRNA (Noller, 1974). RNA was extracted from kethoxal-modified 40S ribosomal subunits, digested to completion with RNase T_1, and subjected to electrophoresis on DEAE paper in 7% formic acid (left to right). Kethoxal was removed from the oligonucleotides by mild alkali, and the resulting newly exposed guanines were cleaved by treatment with RNase T_1. A second dimension of electrophoresis (bottom to top) was performed, in which oligonucleotides originating from sites of modification move to positions above the diagonal line. Oligonucleotide spots reproducibly observed in four independent experiments are numbered in the schematic shown at the right.

oligonucleotide sequence, the position of a kethoxal-reactive guanine can often be confirmed from its surrounding sequence. An example of this type of mobility-shift analysis is discussed in the following paper (Hogan et al., 1984). Locations of the reactive positions in 18S rRNA are summarized in Table IV, which also lists their relative reactivities and availability of the corresponding positions in *E. coli* 16S rRNA in active (Noller, 1974) or inactive (Hogan & Noller, 1978) 30S subunits.

Because of the high uridine content of yeast 18S rRNA, many of the resulting RNase T_1 oligomers run very slowly in the diagonal electrophoresis system. For this reason, it should be kept in mind that Tables III and IV represent a minimal list of kethoxal-reactive sequences and that there may be several more reactive sites that remain to be identified. This also accounts for the skewing of the diagonal pattern to the left (Figure 2), caused by the slow mobility of oligomers in the first dimension.

Positions of kethoxal-reactive guanines are indicated in Figure 1. In 36 out of 48 cases, the kethoxal-reactive sites are found in single-stranded regions of the proposed structure. The 12 cases where reactivities are in apparent conflict with the proposed structure are clearly of interest and are discussed in detail below. Also shown in the figure are the dimethyl sulfate reactive cytosines reported by Kopylov and coworkers (Mankin et al., 1981).

Discussion

This study addresses two questions bearing on the structure of eucaryotic 18S rRNA: (1) How well do chemical probe data on 40S ribosomes agree with the secondary structure model for 18S rRNA, derived by comparative sequence analysis? (2) Is there any correlation between the positions of kethoxal-reactive sites in the small-subunit rRNAs from a procaryote and a eucaryote?

Tables IV summarizes the findings of this study and presents a comparison of kethoxal-reactive sites in yeast 40S subunits and *E. coli* 30S subunits. Although about twice as many sites are identified in Table IV as were found for *E. coli* subunits (Noller, 1974), the number of major sites found for yeast 40S subunits is 23 (Table IV), which compares well with the 25 found for *E. coli*. These numbers are to be taken as approximate minimum estimates of the number of exposed guanines, since the classification of major reactive sites is somewhat arbitrary and certain sites may escape detection due

Table IV: Positions of Kethoxal Reactivity in Yeast 18S rRNA[a]

position	oligonucleotides	reactivity	corresponding reactivity in *E. coli* 16S rRNA: active	corresponding reactivity in *E. coli* 16S rRNA: inactive
6	82b	++		
81	43a	+		
83	43a, 45a, 56b	+		
123	33b	+		134
264	33a	+		204
334	66a, 68a	++		
336	66a, 68a	++		
337	66a, 68a	++		266
372	10, 52	+		301
373	10, 52, 60c	+		302
377	60c, 61	+		
396	24a, 24b	+		
398	24a, 24b, 27a, 27b	+		326
510	15, 32	++	*b*	
539	72	+	*b*	
540	72, 73a	++	*b*	
577	82a	++	530	530
616	74	++		
616	74	++		
616	74	++		
616	74	++		
858	84b	++	*b*	
941	17d, 46c	+		
943	17d, 46c	+		
1001	35a	++	791	791
1113	35b, 56a	+		890
1130	60a, 78	+		
1141	67a	+		917
1317	44	+		
1319	44	+		
1388	67b	+		
1436	20b, 58	++		
1439	65	++		
1603	51b	+		1361
1612	64, 83a	++		1389
1631	46a, 71	+		1392
1640	71	++	1405	
1647	17a	++		
1651	19a	+		
1682	62	+		
1759	46b, 53b	+		
1762	46b, 53b	+	1497	
1769	75, 83b	++	1504	
1770	75, 83b	++	1505	
1780	9, 25a	+		
1782	9, 25a, 28, 57	++	1517	1517
1794	76	++	1529	
1795	76, 77	++	1530	

[a] Oligonucleotide numbers refer to Table III and Figure 2. Sequence positions in 18S rRNA are as given in Figure 1. Reactivity of oligonucleotides is judged as major (++) or minor (+) qualitatively by comparison of relative yields of individual oligonucleotides in four independent experiments. Reactivity of corresponding positions in *E. coli* 16S rRNA in active (Noller, 1974) or inactive (Hogan & Noller, 1978) 30S subunits is indicated, and the positions in the 16S rRNA sequence are listed. [b] Indicates positions in 18S rRNA not found in 16S rRNA.

to inherent limitations of this method. Nevertheless, the number of exposed guanines is clearly of the same order in the two types of ribosomes. Detailed comparison of these results is most easily done in the context of the secondary structure model.

Figure 1 shows a secondary structure for yeast 18S rRNA on the basis of comparative sequence analysis with 16 complete small-subunit rRNA sequences and RNase T_1 catalogs from more than 200 organisms (Woese et al., 1983). The model bears a clear resemblance to that proposed for eubacteria (Woese et al., 1980, 1983; Noller, 1984) and archaebacteria (Gupta et al., 1983). Partition of the structure into three domains (5′, central, and 3′) by characteristic long-range base-paired interactions is evident. Conserved sequences are found in similar or identical positions in the secondary structure, and the common structural "core" (Woese et al., 1983) is highly conserved. The main differences are around positions 240, 500, 530, 700, and 1500 (Figure 1; Chan et al., 1983). The additional 260 nucleotides in yeast 18S rRNA, compared with *E. coli*, are mainly accounted for by insertions around positions 240 and 700. There are, however, many small insertions and differences in detail elsewhere in the structure.

Of the reactive sites listed in Table IV, 36 out of 48 are found in single-stranded regions in the secondary structure, including 19 out of the 23 most strongly hit guanines. Two of the 12 disagreements involved G-U pairs at the ends of proposed helices (positions 123 and 1794), one involves an A-G pair (1682), and 8 of the 12 are only weakly reactive. Two of the latter positions, 1113 and 1141, were also found to be reactive at the corresponding positions in inactive *E. coli* 30S subunits (Hogan & Noller, 1978). This suggests that a fraction of the 40S subunits were in a conformation analogous to the inactive form of 30S subunits as a result of our isolation procedure and may account for many of the less reactive sites listed in Table IV. In fact, nine of the latter sites have reactive counterparts in 16S rRNA from inactive 30S subunits (Table IV; Hogan & Noller, 1978).

Three of the major sites in Table IV, positions 616, 1436, and 1439, are in conflict with the structure in Figure 1. Position 616 lies in a helix that is so far unsupported by comparative evidence (Woese et al., 1983). Its presence in the structure is consistent with psoralen cross-linking results (Wollenzien & Cantor, 1982), as well as several indirect lines of evidence (Noller & Woese, 1981). Pairing in this helix is particularly questionable in eucaryotic 18S rRNAs, and the kethoxal reactivity of position 616 gives further cause for suspicion. It may be that 18S rRNAs have a different structure in this region; no attractive alternative pairing is obvious, however.

Reactivity of positions 1436 and 1439 is interesting, because it implies disruption of a phylogenetically established helical region that was previously found to become more reactive in *E. coli* 16S rRNA upon association of 30S and 50S subunits (Chapman & Noller, 1977). The observed high reactivity of these two positions in yeast 40S subunits in spite of the fact that this helix is amply supported by comparative sequence evidence gives further reason to suspect that this region of the 3′-major domain may undergo biologically significant conformational rearrangement during translation. The low reactivity observed at positions 1317 and 1319 might also be related to such a rearrangement.

Yeast 18S rRNA conformation has also been probed by Mankin et al. (1981), who modified 40S subunits with dimethyl sulfate by the procedure of Peattie & Gilbert (1980). This approach alows identification of unpaired cytosines, and these findings (Figure 1) do not agree completely with our proposed structure. Of the 21 reactive C residues, 13 are unpaired in the secondary structure model, and eight are in apparent conflict. One of the latter is in a G-C pair at the end of a stem at position 1215. Two others are at positions 13 and 14, on the opposite strand from the kethoxal-reactive guanosine at position 1141, which also violates the 12–15/1139–1142 helix. Residues in this helix have also been shown to react with kethoxal in inactive *E. coli* 30S subunits (Hogan

& Noller, 1978). This tends to support the idea that a fraction of the yeast 40S subunits, as isolated, may have a conformation resembling that of the inactive form of *E. coli* 30S subunits.

The remaining reactive sites are in agreement with the proposed structure. Furthermore, there is a striking correlation between the results obtained for 16S and 18S rRNA (Table IV). Six of the strongly reactive 18S sites (510, 540, 671, 676, 751, and 858) are in regions of the structure not present in 16S rRNA and so are excluded from comparison. Eight of the remaining ones have been found to be reactive in the respective analogous positions of 16S rRNA in active *E. coli* 30s subunits (Noller, 1974). Particularly significant is the reactivity of positions 577, 1001, 1640, 1759, 1762, 1769, 1782, 1794, and 1795, whose analogous positions in *E. coli* 16S rRNA have all been implicated in ribosome function.

Position 577 (yeast) corresponds to position 530 (*E. coli*), one of the most kethoxal-reactive guanines in *E. coli* ribosomes. It is equally reactive in 30S subunits and 70S ribosomes but is protected in polysomes (Chapman & Noller, 1977; Brow & Noller, 1983). This loop contains one of the three most highly conserved sequences in the 16S-like rRNAs (Woese et al., 1983), supporting our conclusion that this region of the molecule is directly involved in some crucial aspect of translation.

Position 1001 (yeast) corresponds to position 791 (*E. coli*) and is part of another highly conserved sequence. This position is kethoxal-reactive in 30S subunits but is strongly protected in 70S ribosomes and, when modified, is rejected by 50S subunits in modification/selection experiments (Chapman & Noller, 1977; Herr et al., 1979). This and closely neighboring positions in 16S rRNA are also protected from RNase T_1 and cobra venom RNase by 50S subunits (Santer & Shane, 1977; Vassilenko et al., 1981). The implication is that this region of the molecule is involved in making contact with the large ribosomal subunit.

Position 1640 is located in the nearly universally conserved sequence linking the 3′-major and 3′-terminal domains (positions 1631–1647 in yeast; Figure 1). It is adjacent to the cytidine residue corresponding to C_{1400} (*E. coli*) that has been shown to become cross-linked by mild UV irradiation to the wobble base of the anticodon of certain tRNAs in the ribosomal P site both in *E. coli* (Prince et al., 1982) and in yeast (Ofengand et al., 1982).

Residues 1759, 1762, and 1769 (1494, 1497, and 1504 in *E. coli*) occur in the last of the three most highly conserved sequences in the 16S-like rRNAs, found between positions 1757 and 1770 (1492–1505 in *E. coli*). G_{1497} (*E. coli*), corresponding to G_{1762} (yeast), is protected from kethoxal in *E. coli* 70S ribosomes (Chapman & Noller, 1977). Colicin E3 inactivates *E. coli* ribosomes by cleaving a single phosphodiester bond 5′ to G_{1494} (G_{1759} in yeast) (Bowman et al., 1971; Senior & Holland, 1971). Initiation factor IF3 has been shown to protect residues in this region from nuclease attack when bound to the 49-nucleotide "colicin" fragment (Wickstrom, 1983).

G_{1782} lies in the loop of the 3′-terminal helix, adjacent to two dimethyladenosine residues; bacteria that are resistant to the antibiotic kasugamycin have unmethylated adenosines in the latter two positions (Helser et al., 1972). Positions 1794 and 1795 are located in a region of the molecule that, in bacterial ribosomes, is involved in recognition of mRNA (Shine & Dalgarno, 1974); it is not yet clear whether a similar mechanism holds for eucaryotic ribosomes.

The experimental findings presented here are in general agreement with the secondary structure for yeast 18S rRNA derived from comparative sequence analysis; some interesting anomalies seem to support a previously suggested conformational change in 16S RNA (Chapman & Noller, 1977). The observation that the procaryotic and eucaryotic ribosomes are for the most part reactive at homologous positions (as defined by secondary as well as primary structure) argues strongly that the two classes of ribosomes also share three-dimensional structural homology, particularly in those regions that appear to be directly involved in translation. The fact that both classes of ribosome recognize each other's respective heterologous tRNAs probably reflects this.

Acknowledgments

We are indebted to J. Messing, M. Sogin, and I. Wool for making their sequence results available to us prior to publication and to C. Woese for discussions concerning the secondary structure of yeast 18S rRNA.

Registry No. Kethoxal, 27762-78-3; guanine, 73-40-5.

References

Barrell, B. G. (1971) *Proc. Nucleic Acid Res. Mol. Biol. 2*, 751–779.

Bowman, C. M., Dahlberg, J. E., Ikemura, T., Konisky, J., & Nomura, M. (1971) *Proc. Natl. Acad. Sci. U.S.A. 68*, 964–968.

Brosius, J., Palmer, M. L., Kennedy, P. J., & Noller, H. F. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75*, 4801–4805.

Brow, D. A., & Noller, H. F. (1983) *J. Mol. Biol. 163*, 27–46.

Carbon, P., Ehresmann, C., Ehresmann, B., & Ebel, J.-P. (1979) *Eur. J. Biochem. 100*, 399–410.

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

Chapman, N. M., & Noller, H. F. (1977) *J. Mol. Biol. 109*, 131–149.

Dahlberg, J. E. (1968) *Nature (London) 220*, 548–552.

Gupta, R., Lanter, J., & Woese, C. R. (1983) *Science (Washington, D.C.) 221*, 656–659.

Helser, T. L., Davies, J. E., & Dahlberg, J. E. (1972) *Nature (London), New Biol. 235*, 6–8.

Herr, W., Chapman, N. M., & Noller, H. F. (1979) *J. Mol. Biol 130*, 433–449.

Hogan, J. J., & Noller, H. F. (1978) *Biochemistry 17*, 587–593.

Hogan, J. J., Gutell, R. R., & Noller, H. F. (1984) *Biochemistry* (following paper in this issue).

Lake, J. A., Henderson, E., Clark, M. C., & Matheson, A. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 5948–5952.

Mankin, A. S., Kopylov, A. M., & Bogdanov, A. A. (1981) *FEBS Lett. 134*, 11–14.

McCarroll, R., Olsen, G. J., Stahl, Y. B., Woese, C. R., & Sogin, M. L. (1983) *Biochemistry 22*, 5858–5865.

Messing, J., Carlson, J., Hagen, G., Rubenstein, I., & Oleson, A. (1984) *DNA 3*, 31–40.

Noller, H. F. (1974) *Biochemistry 13*, 4694–4703.

Noller, H. F. (1984) *Annu. Rev. Biochem. 53*, 119–162.

Noller, H. F., & Woese, C. R. (1981) *Science (Washington, D.C.) 212*, 403–411.

Ofengand, J., Gornicki, P., Chakraburtty, K., & Nurse, K. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 2817–2821.

Peattie, D. A., & Gilbert, W. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 4679–4682.

Prince, J. B., Taylor, B. H., Thurlow, D. L., Ofengand, J., & Zimmerman, R. A. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 5450–5454.

Rubin, G. M. (1975) *Methods Cell Biol. 12*, 45–64.

Rubtsov, P. M., Musakhanov, M. M., Zakharyev, V. M., Krayev, A. S., Skryabin, K. G., & Bayev, A. A. (1980) *Nucleic Acids Res. 8*, 5779–5794.
Salim, M., & Madin, B. E. H. (1981) *Nature (London) 291*, 205–208.
Santer, M., & Shane, S. (1977) *J. Bacteriol. 130*, 900–910.
Senior, B. W., & Holland, I. B. (1971) *Proc. Natl. Acad. Sci. U.S.A. 68*, 959–963.
Stiegler, P., Carbon, P., Ebel, J.-P., & Ehresmann, C. (1981) *Eur. J. Biochem. 120*, 487–495.
Vassilenko, S. K., Carbon, P., Ebel, J. P., & Ehresmann, C. (1981) *J. Mol. Biol. 152*, 699–721.
Wickstrom, E. (1983) *Nucleic Acids Res. 11*, 2035–2052.
Woese, C. R., Magrum, L. J., Gupta, R., Siegel, R. B., Stahl, D. A., Kop, J., Crawford, N., Brosius, J., Gutell, R., Hogan, J. J., & Noller, H. F. (1980) *Nucleic Acids Res. 8*, 2275–2293.
Woese, C. R., Gutell, R. R., Gupta, R., & Noller, H. F. (1983) *Microbiol. Rev. 47*, 621–669.
Wollenzien, P. L., & Cantor, C. R. (1982) *J. Mol. Biol. 159*, 151–166.
Wool, I. G. (1979) *Annu. Rev. Biochem. 48*, 719–754.
Zwieb, C., Glotz, C., & Brimacombe, R. (1981) *Nucleic Acids Res. 9*, 3621–3640.

Probing the Conformation of 26S rRNA in Yeast 60S Ribosomal Subunits with Kethoxal†

James J. Hogan,‡ Robin R. Gutell, and Harry F. Noller*

ABSTRACT: The conformation and accessibility of 26S rRNA in yeast 60S ribosomal subunits were probed with kethoxal. Oligonucleotides originating from reactive sites were isolated by diagonal electrophoresis and sequenced. From over 70 oligonucleotide sequences, 26 kethoxal-reactive sites could be placed in the 26S rRNA sequence. These are in close agreement with a proposed secondary structure model for the RNA that is based on comparative sequence analysis. At least seven kethoxal-reactive sites in yeast 26S rRNA are in positions that are exactly homologous to reactive positions in *E. coli* 23S rRNA; each of these sites has previously been implicated in some aspect of ribosomal function.

In the preceding paper (Hogan et al., 1984), we described experiments designed to probe the conformation of 18S rRNA in yeast 40S ribosomal subunits. General agreement was found with a proposed secondary structure model for the 18S rRNA that was derived by comparative sequence analysis. There were, however, significant violations, some of which are in specific regions of the structure that have been previously implicated in possible biologically significant conformational rearrangements.

In this paper, we present analogous studies on the kethoxal reactivity of 26S rRNA in yeast 60S ribosomal subunits. As in the previous study, our intention is to compare the pattern of accessible, kethoxal-reactive guanines in a eucaryotic ribosome with that found for the *Escherichia coli* ribosome and at the same time test a secondary structure model for 26S rRNA derived from comparative sequence analysis. Sequences of over 70 oligonucleotides originating from kethoxal-modified sites have been determined, and 26 of these can be located in the published nucleotide sequence (Georgiev et al., 1981; Veldman et al., 1981) of yeast 26S rRNA. There is general agreement between the proposed secondary structure model and the experimental results. Furthermore, several reactive sites in 26S rRNA are in positions that are exactly analogous to reactive sites in *E. coli* 23S rRNA.

Materials and Methods

Growth and labeling of *Saccharomyces cerevisiae* (strain A364A) and preparation of ribosomal subunits were done as described in the preceding paper (Hogan et al., 1984). Kethoxal modification of 60S subunits, extraction of rRNA, and diagonal electrophoresis were carried out as described previously (Noller, 1974; Hogan et al., 1984). Repurification of oligonucleotides was done as in the preceding paper (Hogan et al., 1984).

Results

Yeast 60S ribosomal subunits were reacted with kethoxal as described in the preceding paper for 40S subunits (Hogan et al., 1984). The extracted 26S rRNA was digested with RNase T_1 and subjected to diagonal paper electrophoresis (Noller, 1974) to isolate and identify oligonucleotides originating from reactive sites. Figure 1 shows the diagonal pattern from one of the four independent experiments. The schematic diagram (Figure 1) indicates the numbering scheme used here in describing the individual oligonucleotides. After repurification, they were subjected to sequence analysis by a variety of further enzymatic digestions, followed by identification of resulting fragments (Barrell, 1971). Identification of methylated nucleotides was confirmed by diagonal electrophoresis with [^{14}C]methyl-labeled ribosomes in a fifth experiment. Sequences deduced from this information are listed in Table I.

Although over 70 oligonucleotides were sequenced, many of these cannot be located unambiguously in the 26S rRNA sequence, due to numerous occurrences of the smaller oligonucleotide sequences. As mentioned in the preceding paper (Hogan et al., 1984), some of the oligonucleotides could be placed on the basis of mobility-shift information. The number of occurrences of each oligomer in the 26S rRNA sequence and their off-diagonal mobility shifts are listed in Table I. An

† From the Thimann Laboratories, University of California, Santa Cruz, California 95064. *Received November 9, 1983.*
‡ Present address: Gen-Probe, Inc., La Jolla, CA 92121.

0006-2960/84/0423-3330$01.50/0