Probing the Secondary Structure of Expansion Segment ES6 in 18S Ribosomal RNA†

Gunnar Alkemar‡,§ and Odd Nygård*,‡

School of Life Sciences, Södertörns högskola, S-141 04 Huddinge, Sweden, and Department of Cell Biology, Arrhenius Laboratories E5, Stockholm University, S-106 91 Stockholm, Sweden

Received October 21, 2005; Revised Manuscript Received March 29, 2006

ABSTRACT: Expansion segment ES6 in 18S ribosomal RNA is, unlike many other expansion segments, present in all eukaryotes. The available data suggest that ES6 is located on the surface of the small ribosomal subunit. Here we have analyzed the secondary structure of the complete ES6 sequence in intact ribosomes from three eukaryotes, wheat, yeast, and mouse, representing different eukaryotic kingdoms. The availability of the ES6 sequence for modification and cleavage by structure sensitive chemicals and enzymatic reagents was analyzed by primer extension and gel electrophoresis on an ABI 377 automated DNA sequencer. The experimental results were used to restrict the number of possible secondary structure models of ES6 generated by the folding software MFOLD. The modification data obtained from the three experimental organisms were very similar despite the sequence variation. Consequently, similar secondary structure models were obtained for the ES6 sequence in wheat, yeast, and mouse ribosomes. A comparison of sequence data from more than 6000 eukaryotes showed that similar structural elements could also be formed in other organisms. The comparative analysis also showed that the extent of compensatory base changes in the suggested helices was low. The in situ structure analysis was complemented by a secondary structure analysis of wheat ES6 transcribed and folded in vitro. The obtained modification data indicate that the secondary structure of the in vitro transcribed sequence differs from that observed in the intact ribosome. These results suggest that chaperones, ribosomal proteins, and/or tertiary rRNA interactions could be involved in the in vivo folding of ES6.

Eukaryotic 16S-like rRNAs are generally longer than their prokaryotic homologues. Blocks of additional nucleotides are inserted at fixed positions within the common rRNA core (*1*). These additional sequences, usually termed expansion segments (ESs)[1] (*1*) or variable regions (V) (*2*, *3*), are absent in most eubacteria and archaebacteria (*1*, *4*). Little is known about the function of the expansion segments, but mutations as well as deletions in the expansion could be detrimental to ribosomal function (*6*−*8*). Rearrangements in the position of some expansion on the ribosome, e.g., ES27, have been associated with a shift between functional states (*5*).

One of the expansion segments in 18S rRNA is termed ES6 (*1*) or V4 (*3*). This expansion segment is positioned between helices 23 and 24 in the common core of the eukaryotic 16S-like rRNAs (Figure 1). Unlike most of the expansion segments, ES6 is present in all eukaryotic 16S-like rRNAs, suggesting that it may have a functional role. The average length of ES6 is 250 nucleotides (*3*). However, in taxa such as protists and insects, the expansion segment can contain up to 800 nucleotides (*9*−*11*). The increase in length is mainly due to inserts at the 5′-end of ES6, but inserts at the 3′-end are also found in some taxa, e.g., Kinetoplastida and Euglenoida (*12*, *13*).

ES6 can be divided into two halves, where the two parts exhibit different sequence variability characteristics. For the 5′-half, the variability permitted construction of a consensus secondary structure model, based on covariation analysis (http://www.rna.icmb.utexas.edu/) (*14*) (Figure 1). The consensus model for the 5′-end of ES6 in mouse 18S rRNA is fully compatible with chemical accessibility data obtained for ES6 in both isolated 40S subunits and complete 80S ribosomes (*15*, *16*).

On the other hand, the 3′-half of ES6 is relatively invariable. This has made secondary structure prediction using covariation analysis a challenging project. Several secondary structure models have been suggested for the 3′-section of ES6 (*3*, *17*−*21*). The latest contribution to the plethora of secondary structure models was based on a comparative sequence analysis of more than 3000 ES6 sequences (*12*).

Recently, we proposed that a sequence in expansion segment ES6 interacts with a complementary sequence in ES3. This suggestion was based on both a comparative sequence analysis of sequences from more than 2900 discrete species (*22*) and an experimental analysis of the secondary structure using structure sensitive chemical and enzymatic probes (*23*). However, to our knowledge, no effort has been

† This work was supported by The Swedish Research Council.

* To whom correspondence should be addressed. Phone: +46 8 608 4701. Fax: +46 8 608 4510. E-mail: odd.nygard@sh.se.

‡ Södertörns högskola.

§ Stockholm University.

[1] Abbreviations: CMCT, 1-cyclohexyl-3-(morpholinoethyl)carbodiimide metho-*p*-toluenesulfonate; DMS, dimethyl sulfate; ES, expansion segment; DOC, sodium deoxycholate; DTT, dithiothreitol; EDTA, ethylenediaminetetraacetic acid; RNase, ribonuclease; RT, reverse transcriptase; SDS, sodium dodecyl sulfate.

10.1021/bi052149z CCC: $33.50
Published on Web 06/08/2006

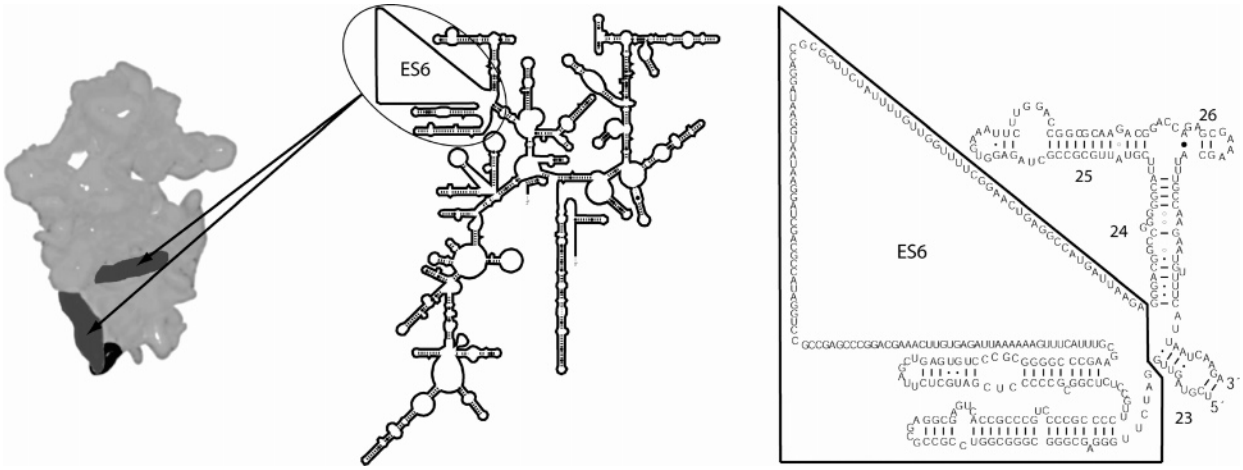

FIGURE 1: Secondary structure model of 18S rRNA [Comparative RNA Web site, http://www.rna.icmb.utexas.edu/ (*14*)], showing the position of expansion sequence ES6 as defined by Gerbi (*1*). The sequence for ES6 in mouse (GenBank accession number X00686) is shown together with the proposed secondary structure of the 5′-section of the expansion segment [http://www.rna.icmb.utexas.edu/ (*14*)]. The adjacent helices in the secondary structure core are numbered according to the system of Wuyts et al. (*12*). The position of ES6 in the 40S ribosomal subunit is according to Spahn et al. (*47*). The 40S subunit is shown from the solvent side.

Table 1: Primers Used for Primer Extension, Sequencing, and Cloning

primer sequence	use	complementary nucleotides
TCGGCATAGTTTATGG	primer extension sequencing (antisense)	1021−1036 in yeast 18S rRNA
CGGCATCGTTTATGG	primer extension	1026−1040 in wheat 18S rRNA, 1080−1094 in mouse 18S rRNA
ATACGAATGCCCCCG	primer extension	879−893 in wheat 18S rRNA, 933−947 in mouse 18S rRNA
AACACTCTAATTTTT	primer extension	811−825 in mouse 18S rRNA, 753−767 in yeast 18S rRNA
CTCGAAGCTCGCCCT	primer extension	348−362 in in vitro-transcribed wheat ES6 (Figure 6)
TCGGCATAGTTTATGG	sequencing (antisense)	1020−1034 in wheat 18S rRNA
GGCGCGCAAATTAC	sequencing (sense)	429−442 in wheat 18S rRNA
ACGGCTACCACATCC	sequencing (sense)	401−415 in yeast 18S rRNA
AATGAAAACATCCTTGGCAAATG	cloning (antisense)	948−970 in wheat 18S rRNA
GGACCTTGGGCCGGG	cloning (sense)	738−752 in wheat 18S rRNA

made to use an experimental approach to analyze the structure of the whole ES6. Therefore, we have analyzed the secondary structure of the entire ES6 in ribosomes from wheat, mouse, and yeast using a range of structure specific RNA-modifying chemical and enzymatic reagents. The experimental results were used to restrict the number of possible secondary structures generated with the RNA folding software MFOLD (version 3.1) (*24*, *25*).

MATERIALS AND METHODS

Materials. Dimethyl sulfate (DMS), 1-cyclohexyl-3-(2-morpholinoethyl)carbodiimide metho-*p*-toluenesulfate (CMCT) and spermidine were from Sigma-Aldrich. Ribonucleases V1, restriction enzyme BstEII (Eco 065), and Sephadex G25 fine were from Amersham Biosciences. SuperScript reverse transcriptase was from Life Technologies, Inc. T7 RNA polymerase and RNasin were from Promega. The pYES2.1 TOPO TA cloning kit was from Invitrogen. Proteinase K was from Roche.

Buffers. Ribosome buffer consisted of 30 mM Hepes-KOH (pH 7.6), 70 mM KCl, 0.25 M sucrose, 2 mM $MgCl_2$, and 5 mM mercaptoethanol. Buffer A consisted of 70 mM KCl, 5 mM $MgCl_2$, 20 mM Tris-HCl (pH 7.6), and 1 mM DTT.

Preparation of Ribosomes. Wheat monosomes were prepared from wheat germ cell free extracts (*26*, *27*). The extracts were treated with 1% DOC (weight by volume) and 1% Triton X-100 (by volume) and layered on a discontinuous sucrose gradient, containing 10 mL of 10% sucrose (weight by volume) in buffer A superimposed on 5 mL of 28% sucrose (weight by volume) in buffer A. The gradients were centrifuged in an SW 27 rotor for 16 h at $90000g_{av}$. The monosomes in the pellet were dissolved in ribosome buffer and stored in aliquots at −80 °C.

Yeast monosomes were prepared from pelleted yeast BY4743 cells (*28*). The cells were dissolved in a buffer containing 20 mM Hepes-KOH (pH 7.6), 2 mM $Mg(CH_3COO)_2$, 100 mM KCl, 1 mM DTT, 1 mM PMSF, and 4000 units of RNasin/mL. An equal volume of glass beads was added, and the cells were crushed by vigorous shaking on a vortex mixer. The glass beads were removed by centrifugation for 5 min at $8000g_{av}$, and the supernatant was treated with DOC and Triton X-100, each at a final concentration of 1% (weight by volume). The samples were centrifuged for 20 min at $15000g_{av}$ and the supernatants collected and treated with KCl at a final concentration of 0.5 M. The supernatants were layered onto a 0.75 M sucrose cushion in buffer A and the ribosomes pelleted by centrifugation for 4 h at $150000g_{av}$. The pelleted salt-washed ribosomes were dissolved in ribosome buffer and stored in aliquots at −80 °C.

Mouse liver monosomes were prepared as previously described (*29*). The isolated salt-washed monosomes were

Table 2: Composition and Base Pair Variability for Helices E23_A, E23_11, E23_12, and E23_14 in ES6

helix[a]	base pair[b]	base pair composition[c] (%)[d]						noncomplementary base pair[c,e] (%)[d]
		A-U	G•U	G-C	C-G	U•G	U-A	
E23_A	1	0.0	0.0	0.0	3.0	95.9	0.1	
	2	96.6	0.5	0.0	0.0	0.0	0.0	
	3	0.0	0.0	0.0	97.4	0.1	0.0	
	4	0.0	0.0	0.0	0.3	1.8	83.5	C°A, 11.0; U°U, 2.3
	5	0.0	0.1	0.0	0.6	78.1	1.3	G°G, 17.4
	6	0.0	0.0	0.0	0.0	3.1	94.5	
	7	0.2	89.2	6.8	0.0	0.0	0.0	
	8	92.9	0.5	0.0	0.0	0.0	0.0	A°A, 4.1
E23_11	1[f]	0.0	0.0	0.0	0.0	0.0	6.6	C°A, 3.6
	2[g]	0.0	0.0	0.0	14.0	3.4	54.7	C°A, 23.6
	3	0.0	0.0	0.0	7.0	92.5	0.1	
	4	0.0	0.0	0.2	86.8	10.1	1.2	
	5[h]	96.4	0.0	0.0	0.0	0.0	0.0	
E23_12	1	0.1	0.0	0.0	96.7	1.7	0.0	
	2[i]	69.5	1.9	0.6	0.0	0.0	0.0	A°C, 23.6; A°G, 2.0
	3	22.0	28.5	45.4	0.1	0.0	0.7	
	4	1.2	2.0	91.3	0.4	0.0	2.4	
	5	0.0	0.0	0.2	94.7	2.0	0.3	
	6	1.6	0.4	1.9	0.3	0.2	0.9	
E23_14	1	90.8	1.4	1.1	0.1	0.0	0.4	A°C, 1.4
	2[j]	0.3	0.1	0.2	75.9	15.7	2.5	
	3	1.2	0.2	22.7	30.8	3.2	38.0	
	4	5.6	0.9	17.1	17.3	16.6	35.1	U°U, 2.1
	5	2.1	0.7	2.2	51.9	19.4	13.8	
	6	2.4	13.3	70.3	0.5	1.0	2.7	C°U, 1.7
	7	5.6	25.3	56.6	0.9	0.7	3.0	G°A, 2.6; U°U, 1.1
	8	1.1	0.7	1.3	4.5	11.7	66.7	U°U, 3.3
	9	0.5	0.6	2.7	29.2	20.0	29.4	U°U, 2.7
	10	1.2	0.1	0.5	77.7	3.3	7.7	C°A, 3.1
	11	0.8	0.2	0.3	9.4	43.4	41.6	
E23_8	1	0.3	0.1	0.3	0.3	0.7	97.1	
	2	0.2	0.0	0.2	2.8	1.0	94.8	
	3	98.7	0.0	0.2	0.0	0.0	0.0	
E23_9	1	1.5	0.8	0.6	5.7	0.9	88.2	
	2	93.9	0.1	3.0	0.2	0.0	1.5	
E23_10	1	1.6	0.0	97.4	0.0	0.0	0.4	
	2	0.1	0.0	0.1	2.3	94.6	2.0	

[a] See Figure 4 for helix numbering. [b] Base pairs are numbered from the 5′-end of the 5′-strand of each helix. [c] The first nucleotide is that in the 5′-strand. [d] Expressed as the percentage of the total number of sequences. The calculation is based on a total of 6362 sequences for helices E23_A, E23_8, E23_9, E23_10, and E23_12, 6357 sequences for helix E23_11, and 6242 sequences for helix E23_14. [e] The most abundant noncomplementary base pairs are shown in cases where <95% of the base pairs are complementary. [f] A total of 89% of the sequences lack the 5′-base in base pair 1. [g] Corresponds to base pair 1 in helix E23_11 in Figure 4. [h] A total of 92.4% of the sequences lack one or both bases in base pair 5. [i] Corresponds to base pair 1 in helix E23_12 in yeast (Figure 4). [j] A total of 3.2% of the sequences lack one or both bases in base pair 2.

dissolved in ribosome buffer and stored in aliquots at −80 °C.

Preparation and Cloning of the Wheat ES6 Sequence. RT-PCR of the ES6 sequence in wheat was essentially as described by Larsson and Nygård (*30*), using the primers listed in Table 1. The PCR product was cloned into the pYES2.1 vector and transformed into TOP10F′ competent *Escherichia coli* cells, using the pYES2.1 TOPO TA cloning kit. The amplified and purified plasmid containing the wheat ES6 sequence was linearized using BstEII, a restriction enzyme that cleaves the plasmid 3′ of the inserted ES6 sequence. Contaminating nucleases were removed by treatment with proteinase K [0.05 mg of proteinase K in 400 μL of buffer containing 10 mM Tris (pH 8.0), 50 mM NaCl, 5 mM EDTA, and 0.5% (weight by volume) SDS] for 60 min at 37 °C, followed by phenol extraction (*31*). The linearized plasmid with the cloned wheat ES6 sequence (pYES2.1/wheat-ES6) was precipitated with 2.5 volumes of ethanol in the presence of 0.1 volume of 3 M $NaCH_3COO$ (pH 5.2), collected by centrifugation, dissolved in H_2O, and stored at −20 °C.

Preparation of Wheat ES6 Transcripts in Vitro. For in vitro transcription, purified and linearized plasmid pYES2.1/wheat-ES6 (5 μg) was incubated in a buffer containing 10 mM DTT, 0.5 mM ATP, 0.5 mM CTP, 0.5 mM GTP, 0.5 mM UTP, 100 units of RNasine, and 40 units of T7 RNA polymerase in a final volume of 100 μL. After incubation for 2 h at 37 °C, 50 units of DNase I was added and the incubation continued for 10 min at 37 °C. The volume of the reaction mixtures was increased to 200 μL by addition of 10 mM DTT, and the samples were extracted twice with phenol and chloroform (1:1). The in vitro-transcribed ES6 RNA was precipitated with 2.5 volumes of ethanol in the presence of 0.1 volume of 3 M $NaCH_3COO$ (pH 5.2), dissolved in ribosome buffer, and purified by gel filtration on a Sephadex G-25 fine column (1 cm × 2 cm) equilibrated in ribosome buffer. Fractions containing rRNA were precipitated as described above, pelleted by centrifugation, and dissolved in ribosome buffer. The RNA was heated at 42 °C for 10 min to allow secondary structure formation (*32*) before chemical and enzymatic modification.

Modification of rRNA. Chemical modification and enzymatic cleavage of rRNA were as previously described (*30*) using the single-strand specific chemical reagents DMS and CMCT and the double-strand specific enzyme RNase V1 (*33*, *34*).

Identification of Modification Sites. The positions of the modified bases and/or cleavage sites were identified using primer extension as previously described (*35*). The primers are listed in Table 1. The primer extension products were analyzed on 8% (weight by volume) acrylamide sequencing gels in an Applied Biosystems 377 DNA sequencer.

Table 3: Helix Lengths

helix[a]	percentage of sequences that could form a helix with the indicated maximum number of base pairs[b]										
	1	2	3	4	5	6	7	8	9	10	11
E23_A	0.0	0.0	0.7	1.9	2.0	5.8	24.4	65.2			
E23_11	0.0	1.9	28.5	63.8	5.7						
E23_12	0,0	0.3	1.3	26.1	67.9	3.5	0.5	nd[d]			
E23_14[c]	nd[d]	nd[d]	nd[d]	nd[d]	0.6	1.4	3.0	5.0	10.3	25.3	53.1
E23_8	0.3	2.2	97.5								
E23_9	3.2	96.8									
E23_10	1.4	98.6									

[a] See Figure 4 for helix numbering. [b] The calculation is based on a total of 6362 sequences for helices E23_A and E23_12, 6357 sequences for helix E23_11, and 6242 sequences for helix E23_14. [c] One hundred twenty sequences (1.9%) of the 18S rRNA sequences in the database lack helix E23_14 according to our sequence alignment. [d] Not determined.

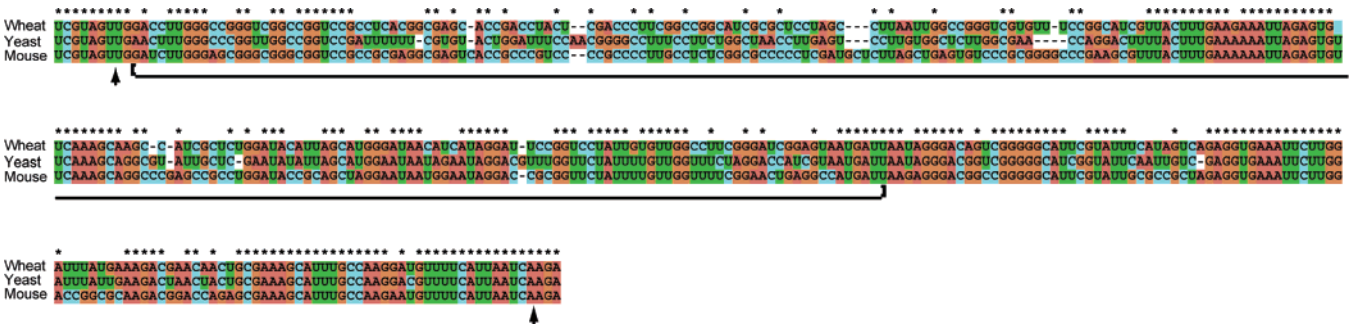

FIGURE 2: Alignment of ES6 sequences from wheat, yeast, and mouse GenBank accession numbers AY049040, J01353, and X00686, using Clustal W (*37*, *38*). Asterisks show the conserved nucleotides. The horizontal line shows the location of ES6 within the aligned sequences. The arrowheads mark the 5′- and 3′-ends of the in vitro-cloned wheat ES6 sequence.

Secondary Structure Predictions. Secondary structure models for ES6 were generated using the energy minimizing software MFOLD (version 3.1) (*24*, *25*). The experimentally obtained modification data were used to restrict the number of possible RNA structures generated by the software (*30*).

Sequencing of the ES6 Sequence in Yeast and Wheat. Yeast and wheat 18S rRNA was sequenced by RT-PCR essentially as described by Larsson and Nygård (*30*) using the forward and reverse primers listed in Table 1.

Base Pair Composition and Helix Length. The nucleotide sequence of ES6 from more than 6000 organisms was downloaded from the European ribosomal RNA database (http://www.psb.ugent.be/rRNA/) (*36*). The present alignment used in the database was kept with minor adjustments to correct for alignment errors. Incomplete sequences lacking sequence information covering the ES6 region were removed before calculation of base pair composition and the number of base pairs in the helices listed in Tables 2 and 3.

RESULTS

Several attempts to generate secondary structure models of expansion segment ES6 in eukaryotic 18S rRNA have been made on the basis of phylogenetic sequence comparisons (*3* and references therein; *12*, *20*). However, as there is considerable uncertainty about these structures, a large part of ES6 is usually left without a defined secondary structure in some of the most used secondary structure diagrams (*14*) (Figure 1). Here we have used an experimental approach to obtain structure data that could be used in the construction of secondary structure models of ES6. To gain insight into the applicability of the generated models, the structure of ES6 in ribosomes from three different organisms, wheat (*Triticum aestivum*), yeast (*Saccharomyces cerevisiae*), and mouse (*Mus musculus*), representing different eukaryotic kingdoms, was analyzed. Although the ES6 structures in these three organisms have relatively similar lengths (233, 232, and 240 nucleotides in wheat, yeast, and mouse, respectively), an alignment of the ES6 sequences using Clustal W (*37*, *38*) showed that only approximately half of the nucleotides were conserved (Figure 2).

Structure of ES6 in Situ

The secondary structure of ES6 in wheat, yeast, and mouse ribosomes was analyzed using the structure sensitive chemical and enzymatic reagents CMCT, DMS, and RNase V1. Ribonuclease V1 specifically cleaves double-strand regions, whereas DMS and CMCT modify only single-strand bases (*33*, *34*). The modifications generated by the various reagents were detected by primer extension using fluorescently labeled primers. The primer extension products were separated on an ABI 377 DNA sequencer. The obtained modification data are shown in Figure 3. Lanes containing unmodified 18S rRNA were run in parallel to identify the reagent-independent natural stops.

ES6 in Wheat. The 5′-part of ES6 in wheat (Figure 3A) contained several bases that were accessible for single-strand specific modification. Clustered regions with three or more adjacent accessible bases were found at U14UG16, U38CAC41, and A94UU96. Position A40 in the second cluster also served as a partial natural stop. However, the peak height at this site was clearly increased upon addition of DMS, suggesting that this position was available for single-strand specific modification. Additional reactive bases were found at A46, U66U67, G69, U77, C82U83, A87, U91U92, G108, and U110U111.

Position A94 was modified by CMCT and not by DMS. An additional CMCT-modified adenine was found at position 170 (see below).

Nuclease V1 cleavage sites were found 3′ of the bases in the sequences G35C36, C44G45, C50, C84CUAGCC90, C99CG101, U104CG106, and C112C113. Cleavages occurred in the regions between the single-strand modified bases except at U91, where both CMCT and nuclease V1 accessibility were observed. The ability of RNase V1 to cleave at stacked single-strand bases may explain this observation (*33*, *34*).

The central section of ES6 (Figure 3B) was insensitive to cleavage by RNase V1. However, this part was relatively accessible to modification by the single-strand specific reagents. CMCT and DMS reactive bases were found at A117, U118, A133, A148, A156U157, U161, A166, and A168 and in the cluster A170UU172. A170 was modified by CMCT and not by DMS. As mentioned above, even A94 exhibited this abnormal behavior. As this type of reactivity has, to our knowledge, not previously been reported, it was possible that our previously determined sequence of wheat 18S rRNA (GenBank accession number AY049040) contained errors. This possibility was analyzed by sequencing ES6 from the ribosome population used in the actual experiments. The results confirmed that positions 94 and 170 are adenines (results not shown). Thus, there was no obvious explanation for the abnormal chemical reactivity at these positions.

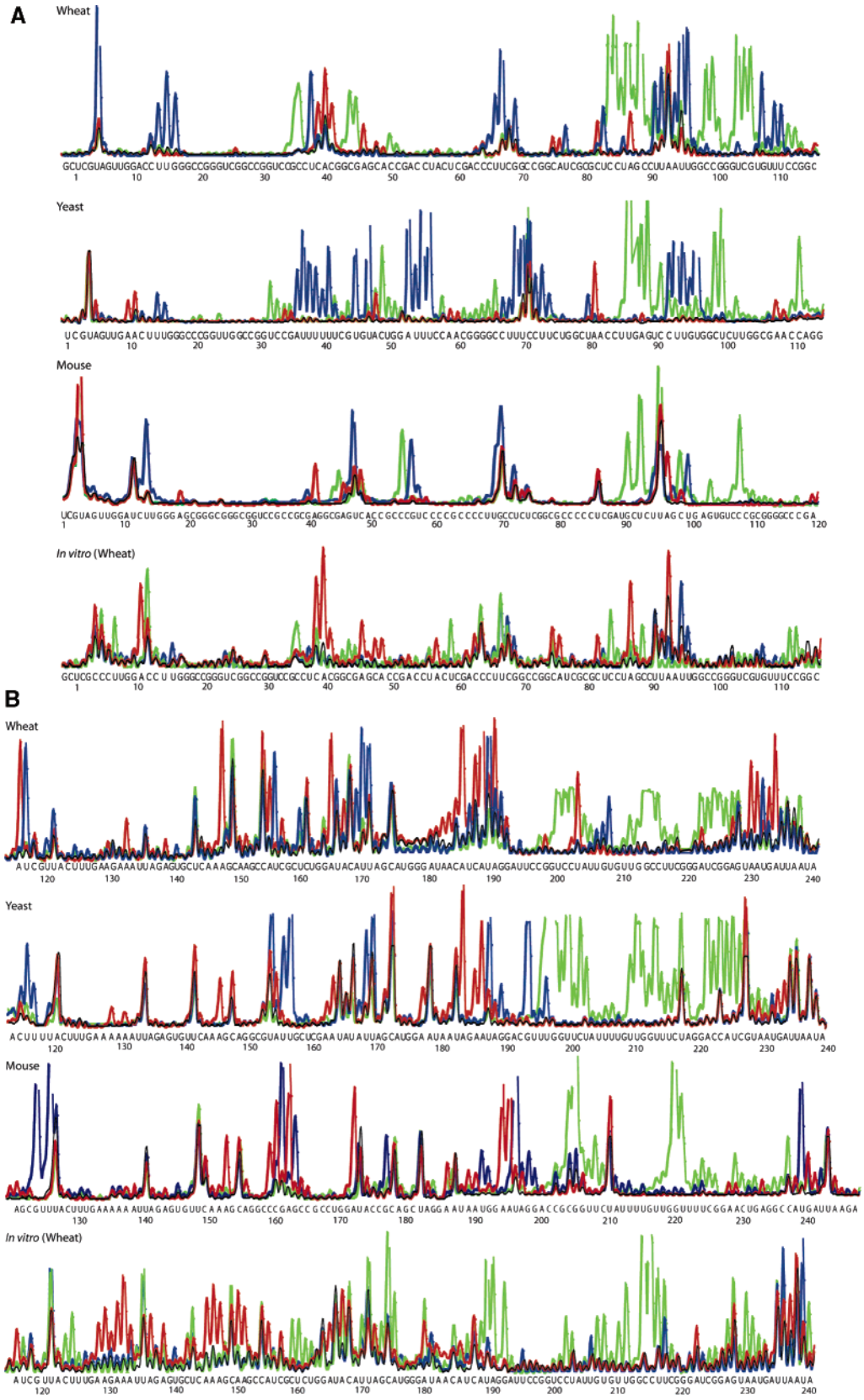

FIGURE 3: Accessibility data obtained by modification and primer extension of rRNA in isolated ribosomes from wheat, yeast, and mouse and of in vitro-transcribed ES6 from wheat. The samples were incubated in the presence of 20 mM DMS (red), 100 mM CMCT (blue), and 2 units of RNase V1 (green). An unmodified but otherwise similarly treated sample was used as a control (black). The ES6 sequence (from 5′ to 3′) of each organism is shown below at the bottom of each graph. Panels A and B show the 5′- and 3′-parts of ES6, respectively. The numbers indicate the nucleotides from the 5′-end of ES6.

The 3′-section of ES6 (Figure 3B) contained many sites that were accessible for nuclease V1 cleavage. Accessible sites were concentrated to three sequences: U201CCU204, G214CC216, and A223UCGGA228. Bases accessible to single-strand specific modification was mainly located in two reactive clusters, A186UCAUA191 and A231AUGA235

(Figure 3B), with additional modified bases found at A183A184, U204, and U209.

ES6 in Yeast. The 5′-section of ES6 was accessible to double-strand specific cleavage by nuclease V1 3′ of bases U32, C49, C66, U86GAG89, C91, C99U100, and C111 (Figure 3A). Because of the many uridines in this part of ES6 in yeast, this part of the expansion segment wasmainly accessible for modification by CMCT. A majority of the modified bases occurred in clusters containing three or more modified nucleotides. Such clusters were found at A36UUUUUU42, A53UUU56, C67UUUCCUU74, and C92UUGU96. Additional modified bases were located at A10A11, U14U15, U45, U47, U80A81, and A108. Nucleotide C71 was also part of a natural stop, but the peak height was clearly increased after addition of CMCT. We have previously noted that cytosines can be modified by CMCT at the pH used in our experiments (*15*, *23*). The observation that A36 and A53 were accessible for modification with CMCT made us suspect that the yeast ES6 sequence retrieved from GenBank differed from that in the yeast strain used in our experiments. Therefore, ES6 in the BY4743 strain was sequenced using RT-PCR. The sequencing confirmed that the bases at positions 36 and 53 were indeed adenines (data not shown).

As previously mentioned, two CMCT reactive adenines were also found in ES6 from wheat ribosomes (see above). The CMCT reactive adenosines seem to be located in similar sequence contexts, i.e., immediately followed by two CMCT reactive uracils. As adenines normally do not react with CMCT, the abnormal reactivity of the four modified adenines was not used in modeling the secondary structure of ES6.

The middle section of ES6 contained no RNase V1 reactive sites. However, nucleotides accessible to single-strand specific modification were found at U117U118, A129, A131, A146, U154AUU157, A164, U169U170, and A178 (Figure 3B).

The 3′-section of ES6 was exposed to nuclease V1 cleavage in the sequences U197GGUUCU203, U210UGGUUUCUA219, and A222CCAUC227 (Figure 3B). Single-strand specific modifications were found at A184, A186AU188, G190, G194, U197, and G233A234 (Figure 3B).

ES6 in Mouse. In comparison to those of wheat and yeast, the 5′-section of mouse ES6 contained relatively few CMCT and DMS reactive bases (Figure 3A). The reactive bases were also mainly found as single sites and not within larger clusters. Accessible bases were found at A41, U48, G57, U69G70, A97, and U100. Nuclease V1 cleavages were detected 3′ of bases C44G45, C55, G91, U93, C99, G103, and C108, indicating that these bases were involved in helical structures.

The middle section of ES6 (Figure 3B) contained two clusters of CMCT and/or DMS reactive bases located in the sequences G124UU126 and C160CGAG164. Additional modified bases were found at positions A153, A173, G178, and A187 (Figure 3B). No nuclease V1 cleavage sites were detected in this part of the mouse ES6 sequence.

The RNase V1 sensitive sites found in the 3′-section of ES6 (Figure 3B) were mainly concentrated in two sequences, C205GGUU209 and G220GU222. Three additional cleavage sites were detected 3′ of U225, A230, and C238. Clusters of bases available for single-strand specific modification was seen at A190AUG193, A195AUAGG200, and U240GA242 (Figure 3B).

Secondary Structure Models

The analyzed ES6 rRNA sequences from wheat, yeast, and mouse were folded with the RNA folding software MFOLD (version 3.1), which is based on the Zuker energy minimization algorithm (*24*, *25*). The CMCT and DMS modification data obtained from the experiments were used to restrict the number of secondary structure models generated for the ES6 sequence. Due to similarities in the modification and cleavage patterns, comparable secondary structure models were obtained for the investigated species (Figure 4). These secondary structure models contained five homologous hairpins, A−E (Figure 4).

Hairpin A. Hairpin A consists of two helices, E23_1 and E23_2 [helices are numbered according to the system of Wuyts et al. (*12*)], and an apical loop of four to five nucleotides (Figure 4). Helix E23_1 contains 8−10 bp and one to two unpaired nucleotides. Due to the lack of single-strand specific modifications, this part of ES6 was modeled as a helix, although there were no RNase V1 cleavages to support the helical nature. The modifications seen at the terminus of helix E23_1 in wheat and yeast (Figure 4A,B) could indicate that these bases were involved in dynamic unstable base pairs (*39*). Helix E23_2 contains 8−12 bp and an internal asymmetrical loop. The helix was partly accessible to cleavage by nuclease V1, while the nucleotides in the internal loop were available for single-strand specific modification in the three species that were investigated. The apical loop of helix E23_2 consists of four to five nucleotides. In mouse, this loop is a tetraloop of the common GNRA type, while in wheat, the loop is of the UMAC type, which was recently shown to structurally belong to the GNRA family of loops (*40*). All nucleotides in the apical loop were accessible for modification in yeast and wheat, while only one nucleotide was reactive in mouse. The two modified bases, A36 and U42, found in yeast could potentially form a dynamic instable A·U pair (*39*).

The sequence between hairpin A and helix 23 in the 18S rRNA core was considered to form a single-strand region due to the presence of DMS and/or CMCT reactive bases (Figure 4).

Hairpin B. Hairpin B consists of two helices, E23_4 and E23_7, and an apical loop of five to six nucleotides (Figure 4). The apical loop was accessible to single-strand specific modification in the three investigated species. Helix E23_4 differed slightly in appearance and accessibility for single- and double-strand specific modification between the investigated species (Figure 4). The helix was inaccessible to both single- and double-strand specific modification in mouse while being accessible for RNase V1 in both wheat and yeast. The single-strand specific modification of U74 in the terminus of helix E23_4 in yeast could result from dynamic breathing as previously discussed (*39*).

Helix E23_7 was accessible to nuclease V1 cleavages at structurally homologous positions in the three experimental organisms (Figure 4), while several bases in the internal loop of the helix and in the loop between helices E23_4 and E23_7 were accessible for DMS and CMCT modification.

Hairpin B is connected to hairpin A via a short single-strand linker sequence consisting of four to seven nucleotides. The linker sequence was available for single-strand specific modification at homologous positions in the investigated species.

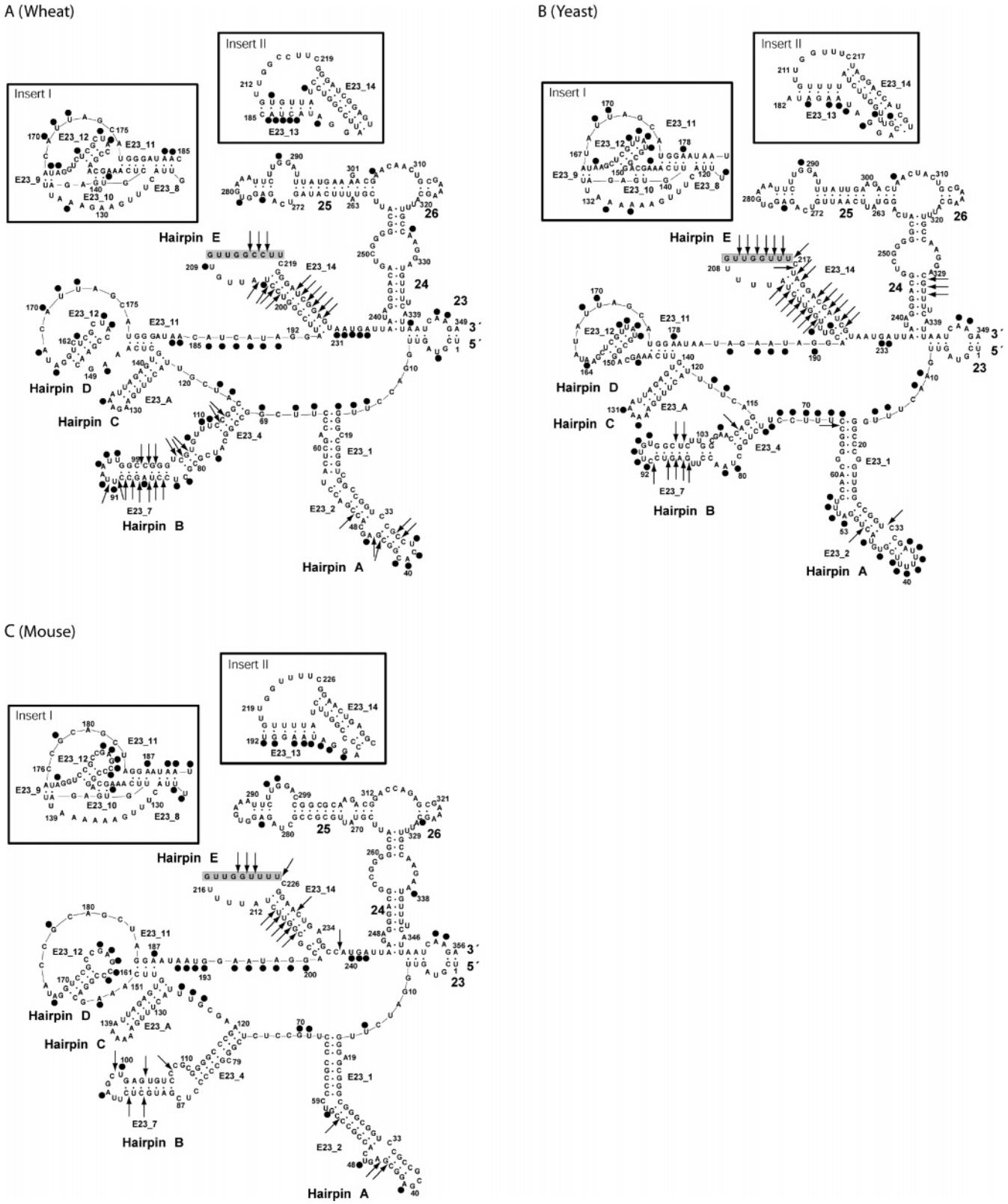

FIGURE 4: Secondary structure models for ES6 in wheat, yeast, and mouse. The basic layout and helix nomenclature were adopted from Wuyts et al. (*12*). Helix E23_A is equivalent to helix E21_6 described by De Rijk et al. (*45*). Modification data are depicted as follows. Filled circles represent bases modified by DMS or CMCT. Arrows indicate RNase V1-dependent cleavages in the backbone. The gray box contains the nucleotides suggested to be involved in base pairing with bases in ES3 (*23*). Insets I and II shows the secondary structure models of part of ES6 according to Wuyts et al. (*12*). Inset I covers nucleotides 120−185 for wheat, 118−183 for yeast, and 125−192 for mouse. Inset II covers nucleotides 185−230 for wheat, 182−229 for yeast, and 192−237 for mouse.

Hairpin C. Hairpin C consists of a helix with 8 bp and an apical pentaloop containing only purines. The loop was accessible for single-strand modification at homologous positions in wheat and yeast, whereas the loop wasinaccessible to modification in mouse. Helical stem E23_A (Figure 4) was insensitive to both single- anddouble-strand specific modification in the investigated species.

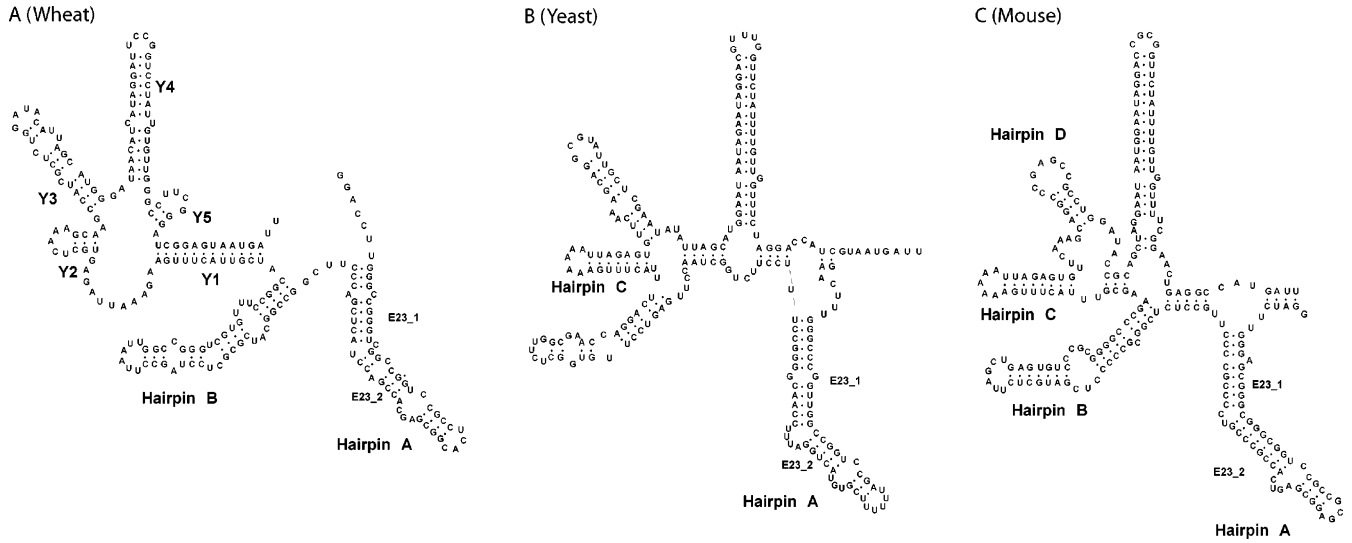

FIGURE 5: Most stable secondary structures of expansion segment ES6 from wheat (A), yeast (B), and mouse (C), folded without restrictions, using the RNA folding software MFOLD (version 3.1) (*24*, *25*).

Hairpin C is connected to hairpin B with a single-strand region that consisted of five to six nucleotides. Several of the nucleotides were accessible for single-strand specific modification (Figure 4).

Hairpin D. Hairpin D varied considerably in nucleotide composition in the investigated species. The hairpin consists of a short helix, E23_12, with 3−5 bp and a loop of variable length. Many of the homologous bases in the apical loop were highly reactive to modification by DMS and CMCT. Thus, the data support the single-strand nature of these nucleotides.

In wheat, the helical region of hairpin D contained an unpaired nucleotide, C162, that functioned as a reagent-independent stop for the reverse transcriptase (Figure 3B). This bulged nucleotide probably reduced the stability of the adjacent base pairs, thereby possibly explaining the CMCT reactivity of U161. An alternative explanation was that the A·U pair could involve an amino-2- or amino-4-carbonyl hydrogen bond, leaving the N3 position in uracil accessible for CMCT modification. No nuclease V1 cleavage sites were detected in helix E23_12.

The single-strand region between hairpin D and helix E23_11 contains 12−14 nucleotides in the investigated species (Figure 4). The occurrence of DMS and/or CMCT accessible sites at homologous positions in wheat and yeast suggested that this was a single-strand sequence, although only G178 was accessible for modification in mouse.

Two adenosines located at structurally homologous positions on each side of hairpin D were available for DMS modification in the examined species.

Internal Helix E23_11. Helix E23_11 consists of 4 bp in wheat and yeast and 3 bp in mouse. No nuclease V1 cleavages supported the presence of this helix. One complication in the construction of this helix was the reactivity at the N1 position of A178 in yeast and its homologous position, A187, in mouse. However, this base could be reactive if these A·U base pairs were in the reverse Hoogsteen conformation. Alternatively, the closing base pair could exhibit dynamic breathing, as previously discussed (*39*).

Hairpin E. Hairpin E (E23_14) contains 9−11 bp and a terminal loop of 13−15 unpaired nucleotides. The helical stem was accessible for double-strand specific cleavages at structurally homologous positions in all three species that were investigated, while the apical loop was inaccessible to single-strand specific modification with the exception of one CMCT sensitive uracil (U209) in wheat. Surprisingly, the loop was accessible to cleavage by the double-strand specific RNase V1 (Figure 4).

Hairpin E was connected on its 3′-side to helix 24 in the 18S rRNA core via a single-strand linker region containing seven to eight nucleotides. The nucleotides in this link were accessible for single-strand specific modification at structurally homologous positions in wheat, yeast, and mouse.

The region between internal helix E23_11 and the 5′-side of hairpin E contained a region of 13 or 14 nucleotides that was suggested to be single-stranded due to the accessibility of homologous bases to single-strand specific modification in the species that were investigated.

Structure of in Vitro Transcribed ES6 from Wheat

During modeling of the secondary structure, we noted that folding of the ES6 sequences with and without application of folding restrictions, obtained using structure sensitive chemical and enzymatic probes, resulted in different secondary structure models (compare Figures 4 and 5). Moreover, the most stable structure models generated by the energy minimizing software were partly species specific (Figure 5). Some structural features of these models were, however, also present in the models of ES6 generated on the basis of the in situ experiments. In mouse ES6, four of the identified hairpins were identical, hairpins A−D (Figures 4 and 5), whereas in wheat and yeast, only two identical or closely similar hairpins (hairpins A and B in wheat and hairpins A and C in yeast) could be identified (Figures 4 and 5).

Recent reports show that prokaryotic ribosomal proteins display RNA chaperone activity (*41*). Thus, there was an obvious possibility that the ES6 sequence could be prevented from forming stable but functionally misfolded states by chaperones available in the nucleus during synthesis, processing, and folding of the rRNA. To see if folding of the expansion segment was influenced by environmental factors in the cell, the structure of wheat ES6, transcribed, purified, and folded in vitro, was analyzed.

For this purpose, the wheat ES6 sequence together with helices 24−26 (Figure 1) was transcribed in vitro and allowed to adopt a stable secondary structure (*32*). Short vector sequences were added at both ends of the in vitro-transcribed wheat ES6 sequence. As the two vector sequences were self-complementary, they could potentially form an additional helix homologous to helix 23 in Figure 1. Thus, the in vitro-transcribed sequence could form a structure similar to the in vivo structure. An analysis of the secondary structure of the transcribed and folded RNA sequence showed that the RNA was accessible for single- and double-strand specific modification at positions 51 and 63, respectively.

The modification data for the 5′-part of the in vitro transcribed sequence are shown in Figure 3A. Here, sites available for DMS and CMCT modification were found at A11, C39AC41, A46, C48A49, A57, U67, A76, C82, A87, A93, U95, and U96. Spread out nuclease V1 cleavages were observed on the 3′-side of U7, C36, U59, U67, C84, C89, C112, and C113.

The middle part of the in vitro-transcribed sequence (Figure 3B) was more accessible to cleavage by nuclease V1 than the 5′-part of the transcript. Sites available for cleavage were detected at C124U125, C160UC162,A168, and G174CA176. DMS reactive sites were found at A117, A129A130, A132A133, A137, A139, A146AA148, A151A152, A156, A166, and A168. None of the bases in this part of the transcript were available for CMCT modification.

The 3′-part of the in vitro-transcribed sequence was mainly accessible to nuclease V1 cleavage (Figure 3B). Accessible sites were seen at A181, C185, A189UA191, G193, C203, G208U209, U211U212, C215CUU218, A228G229, and A231A232. Three single-strand reactive bases were found at positions A181, A238, and U240.

The secondary structure model created using the experimental data suggested that the part of the transcript derived from the rRNA core folded into a secondary structure similar to that originally predicted by covariation analysis (*14*, *42*) and subsequently confirmed by crystallography of the prokaryotic 30S subunit (*43*).

The ES6 sequence of the transcript adopted a structure closely similar to that obtained by folding the ES6 sequence without considering any of the experimental results. As can be seen in Figures 5 and 6, both models exhibited similar structural elements. However, due to the modification pattern of the in vitro-transcribed ES6 sequence, hairpin Y3 became somewhat shorter. This allowed formation of a slightly longer hairpin Y4 and an additional short helix, helix Y2_1 at the expense of hairpin Y5 (Figure 6). The latter hairpin is absent from the structure model of in vitro-transcribed and folded RNA (Figures 5 and 6). The presence of helix Y1 and the helical regions of the two large hairpins, Y3 and Y4, were supported by nuclease V1 cleavages (Figure 6). Small internal helix Y2_1 was not accessible to modification by single- or double-strand specific reagents, whereas the single-strand regions and the apical loops of hairpins Y2 and Y3 were available for single-strand specific modifications.

DISCUSSION

ES6 is the largest expansion segment present in eukaryotic 16S-like rRNAs (*12*). Its ubiquitous presence suggests that ES6 may have a functional role in the ribosome (*12*). Several

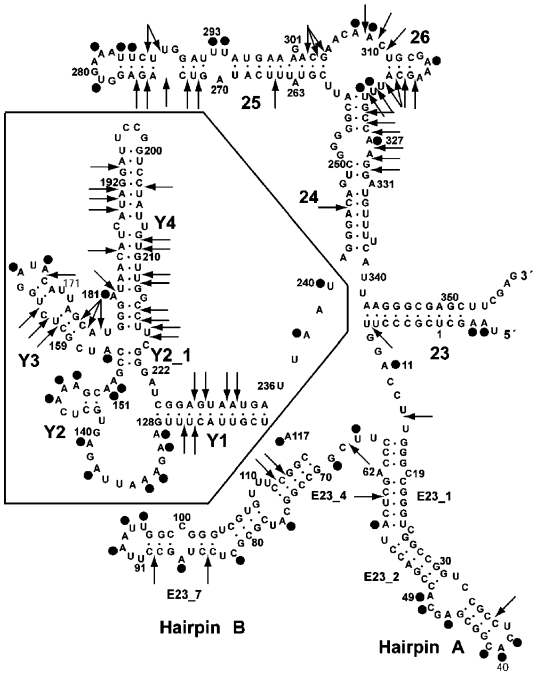

FIGURE 6: Proposed model for the in vitro-transcribed and folded wheat rRNA sequence. Helices and modification data are indicated as in Figure 4.

attempts have been made to construct a consensus secondary structure model for ES6 (*3*, *12*, *17*−*21*), but a large section is usually left without defined structure in secondary structure diagrams based on comparative sequence analysis (*14*). We have analyzed the structure of ES6 in ribosomes from wheat, yeast, and mouse using secondary structure sensitive chemical and enzymatic reagents (*15*, *30*). On the basis of the obtained modification data, a common secondary structure model for ES6 in the three organisms could be constructed. This model contained five hairpins, hairpins A−E, and one internal helix, E23_11 (Figure 4).

Correlation to Previously Suggested Secondary Structure Models. Hairpins A and B have previously been predicted using comparative sequence analysis and are commonly accepted (*14*, *42*). The single-strand regions in these two hairpins were accessible for DMS and/or CMCT modification (Figure 4). This is in agreement with previous observations using isolated 40S subunits and complete 80S ribosomes (*15*, *16*). The double-strand nature of all helical elements in hairpins A and B, except helix E23_1, was confirmed by RNase V1 cleavage (Figure 4).

Hairpin E with its long terminal loop has been proposed in earlier structure predictions (*3*, *12*, *18*−*21*, *23*, *44*). With the exception of U209 in wheat, the 13−15 nucleotides in the loop were inaccessible for modification by DMS or CMCT. In contrast, there were several sites in the loop that were available for cleavage by nuclease V1. The low reactivity to chemical modification and the relatively extensive exposure to nuclease V1 indicate that the loop might be stacked or involved in base pairing with another region in the 18S rRNA. We have recently suggested a tertiary interaction between bases in the 3′-part of the apical loop (positions G210UUGGCCUU218 in wheat) and bases in expansion segment ES3 (5′-AAGGCUGAC-3′) on the basis of comparative sequence analysis (*22*) and on structural probing data (*23*). In previous structure predictions of ES6

(*3*), bases in the 5′-part of the apical loop (positions A205UUGUGUUG213 in wheat) were proposed to form a long pseudoknot by base pairing to the single-strand sequence between helix E23_11 and hairpin E (positions U182AACAUCAU190 in wheat). In a re-evaluation of the ES6 structure using base covariation analysis, Wuyts et al. (*12*) proposed the general existence of a shorter form of the pseudoknot (involving positions A205UUGUG210 and C185AUCAU190 in wheat) in eukaryotes (helix E23_13 in Figure 4, inset II). The short pseudoknot could acquire base pairs at the expense of the helix length in hairpin E23_14 and vice versa, giving a variable length of the single-strand apical loop of hairpin E23_14 (*3*, *12*). This variability was suggested to be associated with the ribosome switching between allosteric states during protein synthesis (*12*).

The shorter form of the pseudoknots is compatible with the proposed interaction between ES3 and ES6 (*22*, 23). However, both the long form and the short form of the pseudoknot involve some of the most DMS or CMCT reactive bases in the ES6 sequences that were investigated (Figures 3 and 4). All bases in the sequence A186UCAU190 as well as U209 in wheat were accessible for modification by single-strand specific reagents. The reactivity of the N1 position of A186 and A189, both involved in A·U base pairs in the pseudoknot, could be explained if the A·U base pairs are in the reverse Hoogsteen conformation. However, the reactivity of the N3 position of U209 could not easily be explained. Likewise, construction of a homologous pseudoknot in mouse and yeast involved the strongly reactive bases U192G193 and A195AU197 (mouse) and A184 and A186AU188 (yeast). As indicated above, the reactivity of the N1 position of adenines involved in A·U base pairs is compatible with the reverse Hoogsteen conformation. However, the reactivity's of the N3 position of U188 in yeast and U192 and U197 in mouse and the reactivity of the N1 position of G193 in mouse are less compatible with these bases being involved in base pairing interactions. One explanation for the high reactivity in these regions could be that the pseudoknot is dynamic. Thus, if the pseudoknot exists, the structure must be open long enough to allow DMS and/or CMCT to react with temporary single-strand bases. The modification data from mouse and wheat cannot rule out the existence of a pseudoknot with highly dynamic properties.

The sequence forming hairpin D varies considerably in both length and base composition with a variability of more than 25% for certain bases in the apical loop (*12*). Despite this variability, many authors have predicted hairpin D as a common structural motif in ES6 (*3*, *12*, *18*, *21*). A helix similar to adjacent helix E23_11 has also been proposed in some earlier structure predictions (*3*, *12*).

The sequence forming hairpin C is highly conserved in all species (*45*). It is therefore not surprising that a homologous hairpin was included in several of the previous secondary structure models of ES6 (*3*, *17*−*21*, *45*).

Recently, Wuyts et al. (*12*) suggested an alternative secondary structure for this sequence based on a comparison of more than 3000 rRNA sequences. According to this alternative model (Figure 4, inset I), three nucleotides, U121UA123 (wheat numbering), form an extra internal helix, helix E23_8, by pairing to nucleotides U182AA184. Bases U136A137 and G140U141 (wheat numbering) form a short interrupted pseudoknot by base pairing to nucleotides U167A168 and G149C150, respectively (helices E23_9 and E23_10 in Figure 4, inset I).

The main disadvantage of this model is that it leaves the majority of the bases in hairpin C as single-strand bases, although only one base in wheat and two in yeast were accessible to single-strand specific modification (Figure 4). Furthermore, helix E23_8 (Figure 4, inset I) is not compatible with the modification data as the reactivity of A183 and A184 in wheat and of A190, A191, and U126 in mouse (Figure 4 inset I) would require considerable destabilization of the short helix. Thus, the only alternatives to E23_A would be the two short (Table 3) helices E23_9 and E23_10.

The main problem with the model presented here is that the existence of E23_A helices could not be confirmed by double-strand specific cleavages (Figure 4). In fact, the whole ES6 sequence between hairpins B and E is inaccessible for RNase V1 cleavage while being exposed to modification by the considerably smaller single-strand specific chemical reagents (Figure 4). Thus, the secondary structure model for this part of ES6 was solely based on the occurrence of single-strand specific modifications.

The presence of helices E23_8−E23_12 in the model by Wuyts et al. (*12*) was based on a comparative analysis of more than 3000 18S rRNA sequences. As a complement to the experimental approach, we have carried out a comparative sequence analysis of putative helix E23_A as well as of helices E23_8−E23_14 using the more than 6000 18S rRNA sequences currently available from the European small ribosomal subunit RNA database (*42*) (Tables 2 and 3). Helix E23_13 was not included in the analysis as the experimental data made the presence of this helix less likely (see above). The current sequence alignment used in the database was kept with only minor adjustments to correct obvious alignment errors (*36*). Accordingly, the composition and abundance of base pairs in helices E23_8−E23_14 were relatively similar to that reported by Wuyts et al. (*12*) despite the increased number of sequences used here.

As noted by Wuyts et al. (*12*), the extent of compensatory base changes varies between the putative helices. Helix E23_14 exhibits substantial compensatory base changes, while helices such as E23_A, E23_8, E23_9, and E23_10, the helices characterizing the two alternative structures for the central part of ES6, exhibit limited compensatory base changes (Table 2).

Wuyts et al. (*12*) have suggested that the structure of ES6 might differ between phyla. Thus, one possibility was that the capability of forming a helix similar to E23_A was restricted to a subgroup of eukaryotic organisms. However, our analysis shows that a helix homologous to E23_A can be formed in all 18S rRNA sequences currently listed in the European small ribosomal subunit RNA database (*36*). As seen in Table 3, there is a substantial length variation in helix E23_A. However, a helix with 4−8 bp could be formed in 99.3% of the organisms. Thus, helix E23_A would probably offer more stability to the structure than the two alternative short helices, E23_9 and E23_10.

Comparison of the Secondary Structures of Wheat ES6 Transcribed and Folded in Vitro with the Secondary Structure in Situ. Analysis of the structure of in vitro-transcribed and folded wheat ES6 showed that the 3′-half of ES6 differed significantly from that obtained in situ, whereas the secondary structure models for the 5′-half (hairpins A

and B) were similar, although there was a difference in the number and position of modifications recorded in situ and in vitro (Figures 4A and 6). An estimation of the free energy for the two structures suggests that the in vitro structure is more stable than the in situ structure. The increased stability is mainly due to long helices Y1 and Y4 (11 and 16 bp, respectively) (Figure 6).

Suggested helices Y1−Y4 (Figure 6) were incompatible with the structural probing data obtained in situ (Figure 3). Thus, a structure similar to that obtained in vitro does not exist in the ribosome, suggesting that stable helices Y1 and Y4 do not form during folding of the 18S rRNA in vivo. One possible explanation could be the involvement of chaperones in folding of eukaryotic rRNAs, in analogy with the situation in prokaryotes (*41*). If the sequence from G178 to U195 is prevented from an immediate cotranscriptional interaction with the complementary sequence from G200 to U217 (hairpin Y4 in Figure 6), part of the latter sequence (G210UUGGCCUU218) could interact with its complementary sequence in ES3 (*22*, *23*), thereby stabilizing the rRNA in favor of the in situ structure. The stabilizing effect of ES6/ES3 could only roughly be estimated. Calculations using RNAeval (*46*) suggest that the tertiary interaction could contribute to the stability by lowering the free energy by −9 to −11 kcal/mol.

A structural stabilization of ES6 due to binding of ribosomal proteins seems less likely as cryo-EM studies (*47*) suggest that the expansion segment lacks prominent protein interaction partners. This suggestion is supported by the substantial accessibility of ES6 to chemical and enzymatic reagents reported here.

Localization of the Structure Parts in the Cryoelectron Microscopic Image. A three-dimensional structural model of yeast ES6 was created using ERNA-3D (*48*). This model was superimposed on the cryo-EM image of the yeast small ribosomal subunit (*47*) (Figure 7). In the cryo-EM image, ES6 emerges on the solvent side of the platform and appears as two separate electron densities. One of the densities forms one of the "back lobes", while the other makes up the so-called left foot (*47*). We have previously suggested that bases in the apical loop of hairpin E interact with sequences in expansion segment ES3 (*22*, *23*). Spahn et al. (*47*) noted that ES3 and ES6 were in close contract in the left foot of the subunit. On the basis of these observations, we have positioned hairpin E in the left foot in the model of the yeast 40S subunit (Figure 7).

Spahn et al. (*47*) suggested that hairpin A holds a position on the yeast ribosome homologous to that occupied by helix 21 in the prokaryotic 30S subunit, e.g., on the back of the subunit. Assuming this position for hairpin A, the model in Figure 7 suggests that hairpin B might form the back lobe situated above hairpin A. The putative location of hairpin A hidden under hairpin B might explain its poor availability to nuclease V1 cleavage. The V1 nuclease is a relatively large molecule and may therefore encounter steric problems accessing the helical regions of a partly buried hairpin A. Similarly, a more exposed position of hairpin B is in agreement with a higher accessibility to nuclease V1 cleavage.

Hairpin C and internal helix E23_11 have been placed in the upper part of the "left foot" (Figure 7). This region seems to be associated with ribosomal proteins (*47*) that may shield these structures from nuclease V1 cleavage.

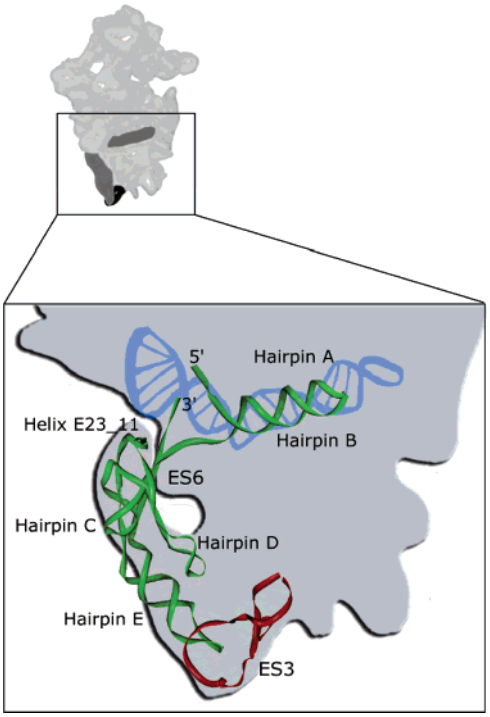

FIGURE 7: Suggested position of the structural elements of ES6 in the yeast 40S ribosomal subunit. A three-dimensional model of ES6 was constructed using ERNA-3D (*48*). The model was superimposed on the contour of the yeast small ribosomal subunit as seen from the solvent side (*47*). The position of hairpin A (blue) is as suggested by Spahn et al. (*47*). ES6 is colored green and ES3 red.

It should be noted that the localization of hairpin D adjacent to hairpin E is very speculative as there are no data to guide the positioning of this hairpin. The high availability of the apical loop for modification only indicated an exposed position on the small ribosomal subunit.

ACKNOWLEDGMENT

We are indebted to Mrs. Birgit Lundberg for technical assistance.

REFERENCES

1. Gerbi, S. A. (1996) Expansion segments: Regions of variable size that interrupt the universal core secondary structure of ribosomal RNA, in *Ribosomal RNA: Structure, evolution, processing, and function in protein biosynthesis* (Zimmermann, R. A., and Dahlberg, A. E., Eds.) pp 71−87, CRC Press, Inc., Boca Raton, FL.
2. Gorski, J. L., Gonzalez, I. L., and Schmickel, R. D. (1987) The secondary structure of human 28S rRNA: The structure and evolution of a mosaic rRNA gene, *J. Mol. Evol. 24*, 236−251.
3. Neefs, J. M., and De Wachter, R. (1990) A proposal for the secondary structure of a variable area of eukaryotic small ribosomal subunit RNA involving the existence of a pseudoknot, *Nucleic Acids Res. 18*, 5695−5704.
4. Raue, H. A., Klootwijk, J., and Musters, W. (1988) Evolutionary conservation of structure and function of high molecular weight ribosomal RNA, *Prog. Biophys. Mol. Biol. 51*, 77−129.
5. Beckmann, R., Spahn, C. M., Eswar, N., Helmers, J., Penczek, P. A., Sali, A., Frank, J., and Blobel, G. (2001) Architecture of the protein-conducting channel associated with the translating 80S ribosome, *Cell 107*, 361−372.
6. Sweeney, R., Chen, L., and Yao, M. C. (1994) An rRNA variable region has an evolutionarily conserved essential role despite sequence divergence, *Mol. Cell. Biol. 14*, 4203−4215.
7. Musters, W., Boon, K., van der Sande, C. A., van Heerikhuizen, H., and Planta, R. J. (1990) Functional analysis of transcribed spacers of yeast ribosomal DNA, *EMBO J. 9*, 3989−3996.
8. Van Nues, R. W., Venema, J., Planta, R. J., and Raue, H. A. (1997) Variable region V1 of *Saccharomyces cerevisiae* 18S rRNA

participates in biogenesis and function of the small ribosomal subunit, *Chromosoma 105*, 523−531.
9. Hancock, J. M., Tautz, D., and Dover, G. A. (1988) Evolution of the secondary structures and compensatory mutations of the ribosomal RNAs of *Drosophila melanogaster*, *Mol. Biol. Evol. 5*, 393−414.
10. Crease, T. J., and Colbourne, J. K. (1998) The unusually long small-subunit ribosomal RNA of the crustacean, *Daphnia pulex*: Sequence and predicted secondary structure, *J. Mol. Evol. 46*, 307−313.
11. Choe, C. P., Hancock, J. M., Hwang, U. W., and Kim, W. (1999) Analysis of the primary sequence and secondary structure of the unusually long SSU rRNA of the soil bug, *Armadillidium vulgare*, *J. Mol. Evol. 49*, 798−805.
12. Wuyts, J., De Rijk, P., Van de Peer, Y., Pison, G., Rousseeuw, P., and De Wachter, R. (2000) Comparative analysis of more than 3000 sequences reveals the existence of two pseudoknots in area V4 of eukaryotic small subunit ribosomal RNA, *Nucleic Acids Res. 28*, 4698−4708.
13. Busse, I., and Preisfeld, A. (2002) Unusually Expanded SSU Ribosomal DNA of Primary Osmotrophic Euglenids: Molecular Evolution and Phylogenetic Inference, *J. Mol. Evol. 55*, 757−767.
14. Cannone, J. J., Subramanian, S., Schnare, M. N., Collett, J. R., D'Souza, L. M., Du, Y., Feng, B., Lin, N., Madabusi, L. V., Muuller, K. M., Pande, N., Shang, Z., Yu, N., and Gutell, R. R. (2002) The Comparative RNA Web (CRW) Site: An online database of comparative sequence and structure information for ribosomal, intron, and other RNAs, *BMC Bioinfor. 3*, 2.
15. Holmberg, L., Melander, Y., and Nygård, O. (1994) Probing the structure of mouse Ehrlich ascites cell 5.8S, 18S and 28S ribosomal RNA *in situ*, *Nucleic Acids Res. 22*, 1374−1382.
16. Holmberg, L., Melander, Y., and Nygård, O. (1994) Probing the conformational changes in 5.8S, 18S and 28S rRNA upon association of derived subunits into complete 80S ribosomes, *Nucleic Acids Res. 22*, 2776−2783.
17. Gonzalez, I. L., and Schmickel, R. D. (1986) The human 18S ribosomal RNA gene: Evolution and stability, *Am. J. Hum. Genet. 38*, 419−427.
18. Ellis, R. E., Sulston, J. E., and Coulson, A. R. (1986) The rDNA of *C. elegans*: Sequence and structure, *Nucleic Acids Res. 14*, 2345−2364.
19. Hendriks, L., De Baere, R., Van Broeckhoven, C., and De Wachter, R. (1988) Primary and secondary structure of the 18S ribosomal RNA of the insect species *Tenebrio molitor*, *FEBS Lett. 232*, 115−120.
20. Nickrent, D. L., and Sargent, M. L. (1991) An overview of the secondary structure of the V4 region of eukaryotic small-subunit ribosomal RNA, *Nucleic Acids Res. 19*, 227−235.
21. Hancock, J. M., and Vogler, A. P. (1998) Modelling the secondary structures of slippage-prone hypervariable RNA regions: The example of the tiger beetle 18S rRNA variable region V4, *Nucleic Acids Res. 26*, 1689−1699.
22. Alkemar, G., and Nygård, O. (2003) A possible tertiary rRNA interaction between expansion segments ES3 and ES6 in eukaryotic 40S ribosomal subunits, *RNA 9*, 20−24.
23. Alkemar, G., and Nygård, O. (2004) Secondary structure of two regions in expansion segments ES3 and ES6 with the potential of forming a tertiary interaction in eukaryotic 40S ribosomal subunits, *RNA 10*, 403−411.
24. Zuker, M. (2003) Mfold web server for nucleic acid folding and hybridization prediction, *Nucleic Acids Res. 31*, 3406−3415.
25. Mathews, D. H., Sabina, J., Zuker, M., and Turner, D. H. (1999) Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure, *J. Mol. Biol. 288*, 911−940.
26. Clemens, M. J. (1984) Translation of eukaryotic messenger RNA in cell-free extracts, in *Transcription and translation: A practical approach* (Hames, B. D., and Higgins, S. J., Eds.) pp 231−270, IRL Press Limited, Oxford, U.K.
27. Seal, S. N., Schmidt, A., and Marcus, A. (1986) The wheat germ protein synthesis system, *Methods Enzymol. 118*, 128−140.
28. Brachmann, C. B., Davies, A., Cost, G. J., Caputo, E., Li, J., Hieter, P., and Boeke, J. D. (1998) Designer Deletion Strains derived from *Saccharomyces cerevisiae* S288C: A Useful set of Strains and Plasmids for PCR-mediated Gene Disruption and Other Applications, *Yeast 14*, 115−132.
29. Sloma, M. S., and Nygård, O. (2001) Chemical accessibility of 18S rRNA in native ribosomal complexes: Interaction sites of mRNA, tRNA and translational factors, *Biol. Chem. 382*, 661−668.
30. Larsson, S., and Nygård, O. (2001) Proposed secondary structure of eukaryote specific expansion segment 15 in 28S rRNA from mice, rats, and rabbits, *Biochemistry 40*, 3222−3231.
31. Brawerman, G., Mendecki, J., and Lee, S. Y. (1972) A procedure for the isolation of mammalian messenger ribonucleic acid, *Biochemistry 11*, 637−641.
32. Noller, H. F., Hoffarth, V., and Zimniak, L. (1992) Unusual resistance of peptidyl transferase to protein extraction procedures, *Science 256*, 1416−1419.
33. Favorova, O. O., Fasiolo, F., Keith, G., Vassilenko, S. K., and Ebel, J. P. (1981) Partial digestion of tRNA−aminoacyl-tRNA synthetase complexes with cobra venom ribonuclease, *Biochemistry 20*, 1006−1011.
34. Lowman, H. B., and Draper, D. E. (1986) On the recognition of helical RNA by cobra venom V1 nuclease, *J. Biol. Chem. 261*, 5396−5403.
35. Holmberg, L., Melander, Y., and Nygård, O. (1992) Ribosome-bound eukaryotic elongation factor 2 protects 5S rRNA from modification, *J. Biol. Chem. 267*, 21906−21910.
36. Wuyts, J., Perrière, G., and Van de Peer, Y. (2004) The European ribosomal RNA database, *Nucleic Acids Res. 32*, D101−D103.
37. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice, *Nucleic Acids Res. 22*, 4673−4680.
38. Higgins, D. G., Thompson, J. D., and Gibson, T. J. (1996) Using CLUSTAL for multiple sequence alignments, *Methods Enzymol. 266*, 383−402.
39. Stebbins-Boaz, B., and Gerbi, S. A. (1991) Structural analysis of the peptidyl transferase region in ribosomal RNA of the eukaryote *Xenopus laevis*, *J. Mol. Biol. 217*, 93−112.
40. Huang, H.-C., Nagaswamy, U., and Fox, G. E. (2005) The application of cluster analysis in the intercomparison of loop structures in RNA, *RNA 11*, 412−423.
41. Semrad, K., Green, R., and Schroeder, R. (2004) RNA chaperone activity of large ribosomal subunit proteins from *Escherichia coli*, *RNA 10*, 1855−1860.
42. Wuyts, J., Van de Peer, Y., Winkelmans, T., and De Wachter, R. (2002) The European database on small subunit ribosomal RNA, *Nucleic Acids Res. 30*, 183−185.
43. Brodersen, D. E., Clemons, W. M., Jr., Carter, A. P., Wimberly, B. T., and Ramakrishnan, V. (2002) Crystal structure of the 30S ribosomal subunit from *Thermus thermophilus*: Structure of the proteins and their interactions with 16S RNA, *J. Mol. Biol. 316*, 725−768.
44. Johansen, T., Johansen, S., and Haugli, F. B. (1988) Nucleotide sequence of the *Physarum polycephalum* small ribosomal RNA as inferred from the gene sequence: Secondary structure and evolutionary implications, *Curr. Genet. 14*, 265−273.
45. De Rijk, P., Neefs, J. M., Van de Peer, Y., and De Wachter, R. (1992) Compilation of small ribosomal subunit RNA sequences, *Nucleic Acids Res. 20*, 2075−2089.
46. Hofacker, I. L., Fontana, W., Stadler, P. F., Bonhoeffer, M., Tacker, P., and Schuster, P. (1994) Fast folding and comparison of RNA secondary structures, *Monatsh. Chem. 125*, 167−188.
47. Spahn, C. M., Beckmann, R., Eswar, N., Penczek, P. A., Sali, A., Blobel, G., and Frank, J. (2001) Structure of the 80S ribosome from *Saccharomyces cerevisiae*−tRNA-ribosome and subunit−subunit interactions, *Cell 107*, 373−386.
48. Mueller, F., Doring, T., Erdemir, T., Greuer, B., Junke, N., Osswald, M., Rinke-Appel, J., Stade, K., Thamm, S., and Brimacombe, R. (1995) Getting closer to an understanding of the three-dimensional structure of ribosomal RNA, *Biochem. Cell Biol. 73*, 767−773.

BI052149Z