J Mol Evol (1990) 30:514–521

Journal of Molecular Evolution

Comparison of Primary and Secondary 26S rRNA Structures in Two *Tetrahymena* Species: Evidence for a Strong Evolutionary and Structural Constraint in Expansion Segments

J. Engberg,[1] H. Nielsen,[1] G. Lenaers,[2] O. Murayama,[3] H. Fujitani,[4] and T. Higashinakagawa[3]

[1] Department of Biochemistry B, Panum Institute, University of Copenhagen, 3 Blegdamsvej, DK-2200 Copenhagen, Denmark
[2] Departement de Biologie Moleculaire et Cellulaire, Laboratoire Arago, C.N.R.S. UA 117, Université Paris VI, F-66650 Banyuls-sur-Mer, France
[3] Department of Developmental Biology, Mitsubishi Kasei Institute of Life Sciences, Machida, Tokyo, Japan
[4] Laboratory of Molecular Biology, School of Veterinary Medicine, Azabu University, Sagamihara, Kanagawa 229, Japan

Summary. We have determined the nucleotide sequence of the 26S large subunit (LSU) rRNA genes for two *Tetrahymena* species, *T. thermophila* and *T. pyriformis*. The inferred rRNA sequences are presented in their most probable secondary structures based on compensatory mutations, energy, and conservation criteria. The majority of the nucleotide changes between the two *Tetrahymena* LSU rRNAs and the positions of a relatively large deletion and of the processing cleavage sites resulting in the generation of the hidden break are all located within the so-called divergent domains or expansion segments. These are regions within the common core of secondary structure where expansions have taken place during the evolution of the rRNA of higher eukaryotes.

The dispensable nature of some of the expansion segments has been taken as evidence of their nonfunctionality. However, our data show that a considerable selective constraint has operated to preserve the secondary structure of these segments. Especially in the case of the D2 and D8 segments, the presence of a considerable number of compensatory base changes suggests that the secondary structure of these regions is of functional importance. Alternatively, these expansion segments may have maintained characteristic folding patterns because only such structures are being tolerated within otherwise functionally important regions.

Key words: 26S rRNA expansion segments — Secondary structure — Compensatory mutations — *Tetrahymena*

Introduction

Comparative analyses of the primary and secondary structures of rRNA are particularly useful in two respects: they constitute the basis for inferring phylogenetic relationships, which may lead to a better understanding of the primary lines of descent, and they may lead to the identification of particular regions of rRNA sequences that are functionally important in the ribosome (Gerbi 1985; Noller et al. 1986; Woese 1987; Raué et al. 1988).

Apart from differences in opinion among molecular evolutionists about which statistical treatment of the data is likely to produce the most reliable evolutionary distances, the actual alignment of sequences used for the calculations is a matter of controversy. The construction of secondary structures of rRNA sequences has proven very useful for phylogenetic studies, as they can be used as topographical markers for sequence alignments (Michot and Bachellerie 1987; Lenaers et al. 1988).

The small subunit (SSU) rRNA sequence has been widely used as a measure of evolutionary distances (cf. Sogin et al. 1986a) but the SSU rRNA data alone may lead to confusing indications depending on the outgroup used, as the SSU rRNAs appear to have

Offprint requests to: J. Engberg

evolved at different rates in different lineages. In contrast, the large subunit (LSU) rRNAs appear to have evolved at similar rates in the kingdoms of Animalia, Plantae, and Fungi (Gouy and Li 1989). The LSU rRNAs display a conserved core structure interspersed with 12 expansion segments, which have retained an invariant location relative to the universal core of secondary structure. These segments have undergone large variations in size and sequence during the evolution of ribosomes of higher eukaryotes, but some conservation of secondary structures has been observed between related species (Michot and Bachellerie 1987; Lenaers et al. 1989).

As large primary sequence differences exist within these structures, the evidence for their alleged conservation is mainly based on folding patterns, whereas evidence based on compensatory base changes is lacking.

We have attempted to investigate the structural–functional relationship of these expansion segments by sequencing the coding regions of the macronuclear copy of the LSU rRNA genes of two *Tetrahymena* species. *Tetrahymena* is uniquely suited for this kind of analysis because its macronucleus contains a completely homogeneous population of rRNA genes (due to an amplification of the single gene copy of the micronuclear genome, reviewed by Engberg 1985; Yao 1986) and because a considerable amount of phylogenetic information is available on the *Tetrahymena* species. The latter information is based mainly on SSU rRNA sequence data (Sogin et al. 1986b) and helped us to choose two species for the LSU rRNA sequence comparison that would be likely to give phylogenetically meaningful data.

Materials and Methods

Cell Strains and Ribosomal DNA (rDNA) Recombinant Clones. The cultivation of *Tetrahymena thermophila,* strain B1868VII (micronucleate, rDNA intron$^+$) and *Tetrahymena pyriformis,* strain GL-C (amicronucleate, rDNA intron$^-$), as well as the isolation and identification of rDNA clones containing the 26S LSU rRNA coding units have been described previously (Engberg et al. 1980; Higashinakagawa et al. 1981; Niles et al. 1981).

In the case of *T. thermophila,* detailed restriction enzyme maps were worked out for the rDNA clones pRP4, pRP10, pRP14, and pGY17 (cf. Engberg et al. 1980), and the appropriate restriction fragments were gel-purified, end-labeled, and the sequence of both strands was determined according to the Maxam and Gilbert procedure. Sequencing of the ends of all restriction fragments was verified by sequencing overlapping fragments. In the case of *T. pyriformis,* the ~10-kb large rDNA fragment extending from the centrally located KpnI site to the terminus of the rDNA molecule was cloned into a pBR322-derived vector. This plasmid is called pTprKT. For sequencing, the pTprKT plasmid was cleaved with appropriate restriction enzymes and the rDNA fragments subcloned into pUC19. From a complete set of overlapping subclones, deletion clones were prepared using the exonuclease III/mung bean nuclease procedure. The sequence of both strands was determined according to the chain termination procedure.

The two 26S LSU rRNA gene sequences are being submitted to the EMBL data bank.

Nuclease Protection, Primer Extension, and Reverse Sequencing Procedures. Mapping of the 3′ end of 26S LSU rRNA on the rDNA sequences of *T. thermophila* and *T. pyriformis* has been published previously (Niles et al. 1981; Din et al. 1982), as has the sequence of the LSU rDNA intron from *T. thermophila* (Wild and Sommer 1980). Mapping of the 3′ end of the 26S α-fragment of *T. thermophila* was carried out according to conventional nuclease protection protocols (cf. Engberg et al. 1980), using the 540-bp BglII/HindIII fragment (prepared from the pRP10 plasmid). The fragment was labeled at the BglII site and hybridized to total rRNA, which was prepared as described previously (Pedersen et al. 1985). The Maxam–Gilbert sequencing ladder of the protected end-labeled fragment was used for size determinations. Oligonucleotides complementary to nucleotides (nt) 25–47 and 1643–1662 were used in primer extension and reverse sequencing experiments, thereby mapping the 5′ ends of the 26S α and β fragments, respectively. An oligonucleotide complementary to nt 205–228 was used to verify the rRNA sequence of different *Tetrahymena* species in the D1a region. Oligonucleotides were kindly provided by Dr. Otto Dahl, Department of Chemistry II, University of Copenhagen, and used for primer extension and reverse sequencing according to the protocols by Geliebter (1987).

Secondary Structure. The secondary structure model for the LSU rRNA of *Tetrahymena* was constructed by comparison with former models (Michot et al. 1984; Leffers et al. 1987; Gutell and Fox 1988; Lenaers et al. 1988). The Zuker program (Zuker and Stiegler 1981) was used to determine the folding of the various expansion segments, by taking into consideration the recent data (Lenaers et al. 1989) on *Prorocentrum micans,* the *Tetrahymena* data of the present paper, and the partial rDNA sequence data of other *Tetrahymena* species from the literature. We have used a new numbering system for the stem regions, which starts with the 5′ end of 5.8S rRNA and labels all the succeeding stems regardless of their size and the size of the bulges they include. Furthermore, the stems of expansion segments have received a name that includes the number of the variable domain, followed by a letter.

Results and Discussion

Primary Structure

The entire sequences of the extrachromosomal rDNA molecules of *T. thermophila* and *T. pyriformis* have been completed recently (Engberg and Nielsen, unpublished; Murayama, Fujitani, and Higashinakagawa, unpublished). The clones used for the sequencing in the present work are described in the Materials and Methods. The 5′ end of 26S rRNA was determined by reverse sequencing methods using an oligonucleotide primer complementary to nt 25–47 as described in Fig. 1A. The mapping of the 3′ end has been described previously (Niles et al. 1981; Din et al. 1982). In *T. thermophila,* the 5′ and 3′ ends of 26S rRNA are located 3760 nt apart on the rDNA molecule. The primary transcript corre-

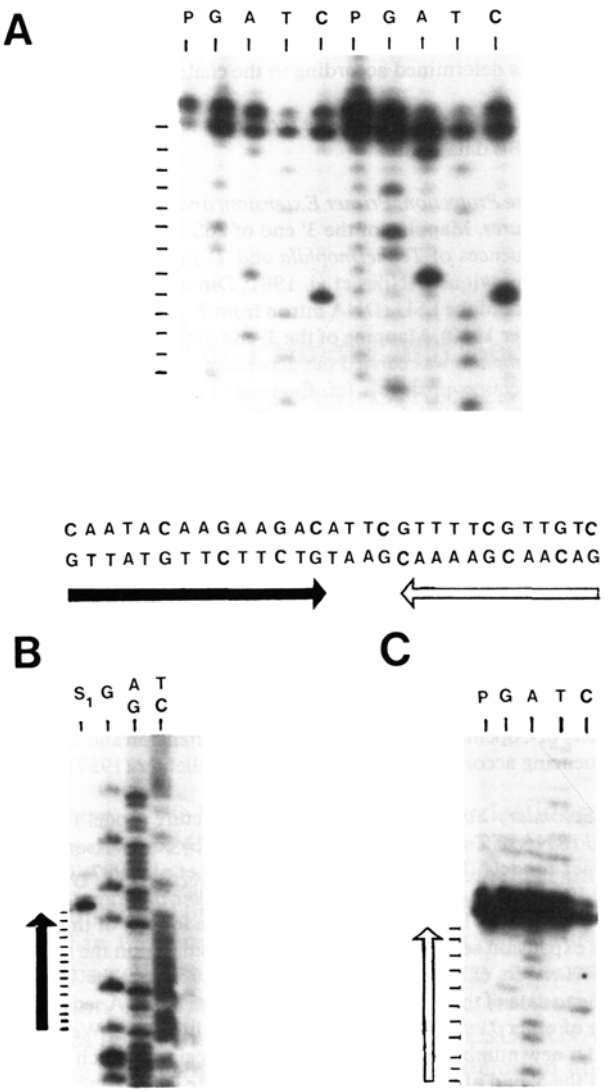

Fig. 1. Mapping the 5′ end of LSU rRNA and defining the boundaries of the hidden break. **A** The oligonucleotide complementary to LSU rRNA nt 25–47 was radioactively labeled and annealed to total RNA isolated from *T. pyriformis* (left panel) or *T. thermophila* (right panel), followed by extension and reverse sequencing according to the protocol of Geliebter (1987). The extension products were fractionated on an 8% urea polyacrylamide sequencing gel. The lower of the two prominent bands that are present in all lanes are believed to represent the run-off extension products, and the upper bands probably represent the result of a nontemplate-directed addition of one extra nucleotide to the run-off products. G, A, T, and C, nucleotides of the cDNA; P, primer extension. **B** The 540-bp BglII/HindIII fragment, which covers the hidden break region, was isolated and hybridized to total *T. thermophila* RNA. The S1 nuclease-resistant material was analyzed on 8% urea polyacrylamide sequencing gels along with the Maxam–Gilbert sequencing products of the same end-labeled BglII/HindIII fragment. The black arrow illustrates the terminal 3′ end region of the nuclease-protected fragment (after correction for one nucleotide (cf. Sollner-Webb and Reeder 1979). **C** The oligonucleotide complementary to LSU rRNA nt 1643–1662 was used for primer extension and reverse sequencing procedures exactly as described above. The white arrow illustrates the terminal 5′ end region of the extended primer.

sponding to this region includes a 413-nt large self-splicing group I intervening sequence (reviewed by Cech 1986), as well as a small segment of 4 nt, which is removed during processing (see below), whereby the 26S rRNA is split into two fragments (26S α and 26S β) having the sizes of 1593 nt and 1750 nt, respectively. Because these fragments form S1 nuclease-resistant hybrid molecules with the corresponding rDNA fragments (cf. Engberg et al. 1980), no additional intervening sequences or hidden breaks appear to be present in the 26S rRNA. In *T. pyriformis* (which has no rDNA intron), the 5′ and 3′ ends of 26S rRNA are located 3343 nt apart on the rDNA molecule. The hidden break (cf. Eckert et al. 1978) was determined, by S1 mapping, to be similar in size and location to that of *T. thermophila* (results not shown), but it has not been mapped definitively at the nucleotide level. Thus, the sizes of the 26S rRNAs (α + β + hidden gap) of the two *Tetrahymena*s are almost the same: 3347 nt and 3343 nt. The overall G+C content of the two LSU rRNAs is about 48%, which is strikingly higher than that of the nontranscribed rDNA spacer region and the genome overall (about 25%, cf. Engberg 1985; Karrer 1986) but similar to that of the rDNA of other ciliates and yeast (cf. Lenaers et al. 1989).

Hidden Break

The boundaries of the hidden break (i.e., the distance between the 3′ end of 26S α and the 5′ end of 26S β) were mapped by S1 nuclease and reverse sequencing procedures, as described in Fig. 1B and C. The reverse sequencing experiment showed a faint sequence pattern that could be recognized as the rDNA sequence continuing beyond the position of the prominent stop. When the 5′ end of 26S β was mapped using the S1 nuclease procedure (results not shown), an obvious discontinuity corresponding to the strong stop in the reverse sequencing reaction was observed. These experiments indicate that the strong stop depicted in Fig. 1C was not caused by the pausing of the reverse transcriptase. From the relative intensities of the bands on either side of the strong stop, an estimated 10% of the 26S rRNA molecules did not contain the hidden break. This estimate is identical to that obtained by Eckert et al. (1978) on the basis of in vivo pulse-labeling studies. The combined mapping data of Fig. 1B and C demonstrate that the size of the hidden break is 4 nt in *T. thermophila*. The position on the secondary structure map of the breakpoints is illustrated in Fig. 2, which is an enlargement of the expansion segment D7A from Fig. 3 (see later). The generation of the hidden break in *T. thermophila* can be described as a deletion of the terminal loop of the D7Ac domain. As illustrated in Fig. 2, the sequence

of the *T. pyriformis* LSU rRNA allows for an additional base pair to be formed in the D7Ac stem next to the cleavage sites as determined for *T. thermophila*. To test whether the generation of the hidden break in *T. pyriformis* could be described as a loop deletion event too, we determined the 3′ end of the hidden break in this (and several other) species of *Tetrahymena* using the reverse sequencing procedure. In all cases the 3′ cleavage site corresponded exactly to that found for *T. thermophila* (J. Engberg, unpublished). We have not yet determined the position of the 5′ cleavage site in these other *Tetrahymena* species. A centrally located hidden break in the LSU rRNA has been demonstrated in a variety of organisms (reviewed by Raué et al. 1988), but only in the case of insects (Ware et al. 1985) and protists (Lenaers et al. 1989; this work) have the boundaries been mapped at the nucleotide level. In two insect species, *Drosophila* and *Sciara,* the break in LSU rRNA is accompanied by a loss of 19 and 75 nt, respectively, and it is interesting to note that the introduced gap is located at exactly the same position on the secondary structure map in insects as it is in *Tetrahymena*; namely, at the tip of the expansion segment D7A (cf. Ware et al. 1985). In contrast, in dinoflagellates, the hidden break is located in expansion segment D2 (Lenaers et al. 1989). Considering the recent observations of multiple breaks in the LSU rRNA of the trypanosomid protozoan, *Crithidia fasciculata* (Spencer et al. 1987), and the mitochondrial LSU rRNAs of *Chlamydomonas* (Boer and Gray 1988) and *Tetrahymena* (Heinonen et al. 1987), there may be no easy explanation for the significance of the specific locations of the different hidden breaks on the secondary structure map, except that they are located within dispensable regions of the expansion segments (cf. Clark 1987; Spencer et al. 1987; Dover 1988).

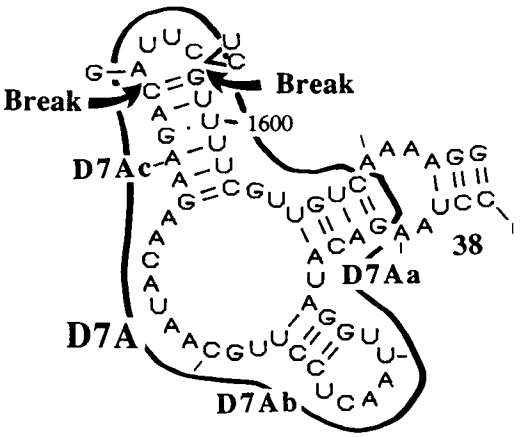

Fig. 2. Location on the secondary structure map of the hidden break in *T. thermophila* LSU rRNA. The D7A region of the LSU rRNA secondary structure map (cf. Fig. 3) is shown to highlight the position of the cleavage sites (arrows) involved in the generation of 26S α and 26S β. The positions of the sites were inferred from the data shown in Fig. 1B and C.

Secondary Structure

Our proposed secondary structure model for the *Tetrahymena* LSU rRNA is shown in Fig. 3. Stems of the conserved core have been assigned numbers beginning at the 5′ end of 5.8S RNA; successive stems are numbered irrespective of their length or the number and size of the bulges they include. This numbering system represents a simplification in that it includes all the stems described in the recent compilation of Gutell and Fox (1988) but reduces the actual number of stems used in previous numbering systems (Michot et al. 1984; Leffers et al. 1987). Furthermore, it can be adapted to any LSU rRNA secondary structure. Stems of expansion segments have received a name including the number of the expansion segment, followed by a letter.

The proposed secondary structure of the conserved core takes into account all the stems of the Gutell and Fox model with some minor modifications. Two of these have been discussed in connection with the *P. micans* LSU rRNA secondary structure model (Lenaers et al. 1989) and deal with a new definition of stem 52 and the presence of two base pairings located between stems 68 and 69, all of which are supported by compensated mutations. Others concern minor refinements of stems 42, 43, and 83. We have previously reported on the folding of some of the expansion segments of the *T. thermophila* LSU rRNA (Lenaers et al. 1988). The now available *T. pyriformis* sequence has permitted further refinements of the secondary structures. For example, searching for base-pair changes in Fig. 3 revealed 26 compensatory mutations between the two sequences; 11 of these increase the helix stability, 4 conserve it, and 11 reduce it. Most of the compensatory base changes occur within the expansion segments D2 and D8. Additional data from the literature confirm our proposed folding of D2: As part of a comprehensive analysis of the phylogenetic relationships among ciliates, Nanney and coworkers have sequenced a region of the LSU rRNA from 19 different *Tetrahymena* species (Preparata et al. 1989). This region corresponds to our D2 domain, and out of the 43 mutated positions observed in their study, 28 are compensated and the others are located in single-stranded regions. The positions of the compensatory base changes on the secondary structure map are illustrated in Fig. 4. In general, the sequence comparisons depicted in Figs. 3 and 4 show that the secondary structure of all the expansion segments of the two *Tetrahymena* LSU rRNAs are very conserved; only a few insertions (or deletions) have occurred and none give rise to drastic folding changes.

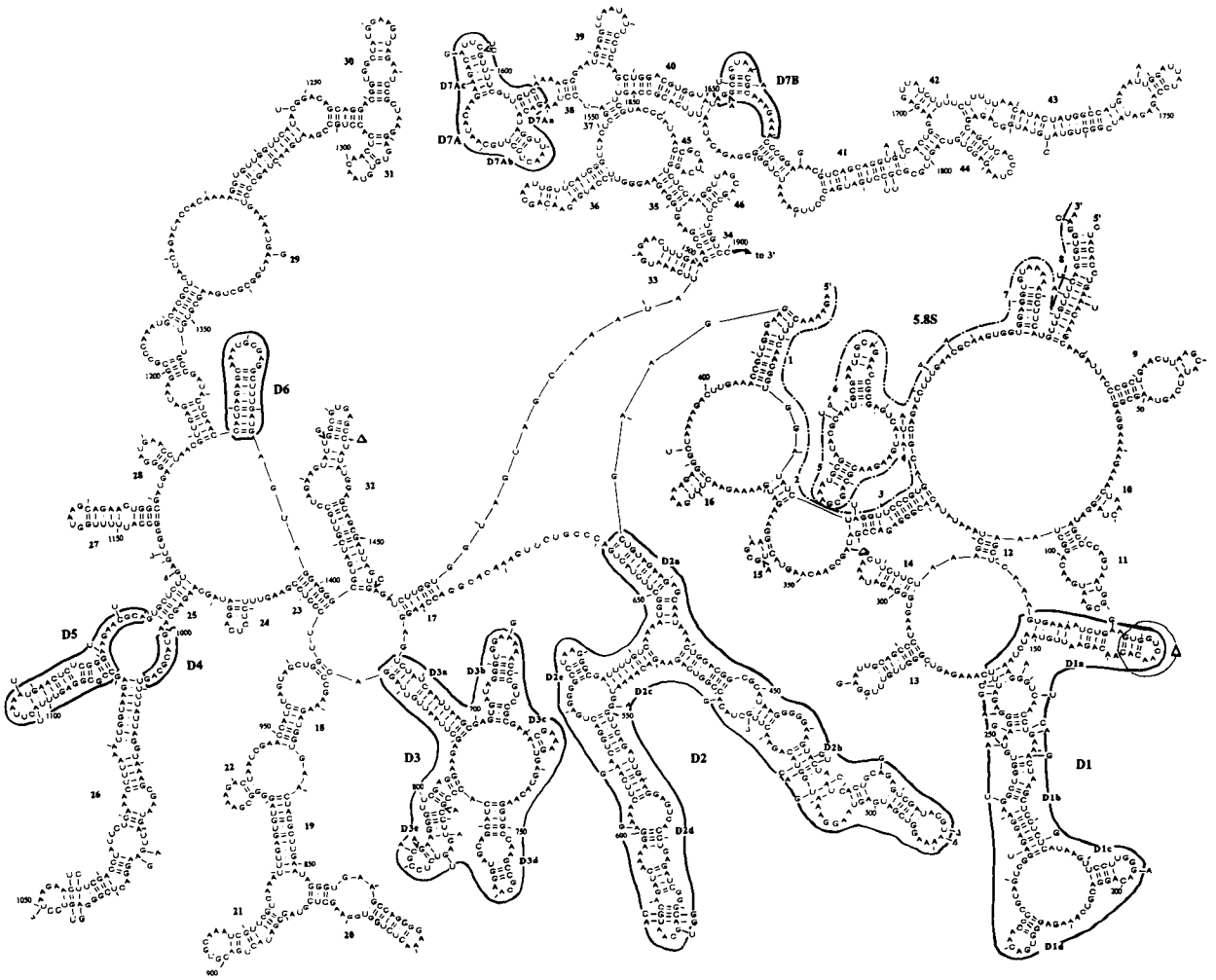

Fig. 3. Secondary structure map of the LSU rRNA from *T. thermophila* and *T. pyriformis*. The sequences were inferred from the gene sequences. Selected portions of the LSU rRNA have been sequenced by us and others (Baroin et al. 1988) using the reverse sequencing method and were found to be in total agreement with the corresponding DNA sequences. The secondary structure model is based on Michot et al. (1984), Gutell and Fox (1988), and Lenaers et al. (1989). Stems of the conserved core have been assigned numbers beginning at the 5′ end of 5.8S RNA as described in the text. The figure shows the *T. thermophila* sequence with the modifications corresponding to the *T. pyriformis* sequence. Nucleotide changes are positioned by hyphens and single insertions by an arrowhead. Double insertions (D7a and D10) have a point in the middle of the arrowhead. Deletions are indicated by a capital delta. The 11-nt-long deletion in the D1 domain has been encircled and denoted by a large delta. The 5.8S RNA molecule is shown with a dashed line. IVS, intervening sequence.

In D1, the terminal base pairing of the D1a stem has been deleted (or inserted), probably as a result of a single event. Sequence data from the D1 region of the LSU rRNA of *Tetrahymena pigmentosa* showed the D1a region to be present in this species (J. Engberg, unpublished). This suggests that the corresponding region was lost in the case of *T. pyriformis,* as *T. thermophila* and *T. pigmentosa* are representatives of the two major groupings of the *Tetrahymena* complex, whereas *T. pyriformis* has branched out from the line leading to *T. thermophila,* according to the phylogenetic tree based on SSU rRNA sequence data (Sogin et al. 1986a). Otherwise, the secondary structure of D1 is perfectly conserved. Mutations have occurred in loops or are compensatory base changes. The observed difference is 11% (15 nt out of 138 nt). In the D2 domain, 16 mutations have occurred within 238 nt (7% difference), and 3 nt are deleted, which may be the result of a double deletion in the terminal loop of the D2b stem and a single deletion inside the same stem. The D3 as well as the D10 domains have two deletion events. The mutations in D3 are located in the D3e stem, which was noted to be not alignable in other protist sequences (Lenaers et al. 1988). The observed difference is 3% (4 nt out of 138 nt) in D3 and 9% (6 nt out of 68 nt) in D10. The D8 domain has a perfectly conserved folding; 16 mutations occur within 122 nt (13% difference). In D12, only two mutations occur within 104 nt (2% difference). The

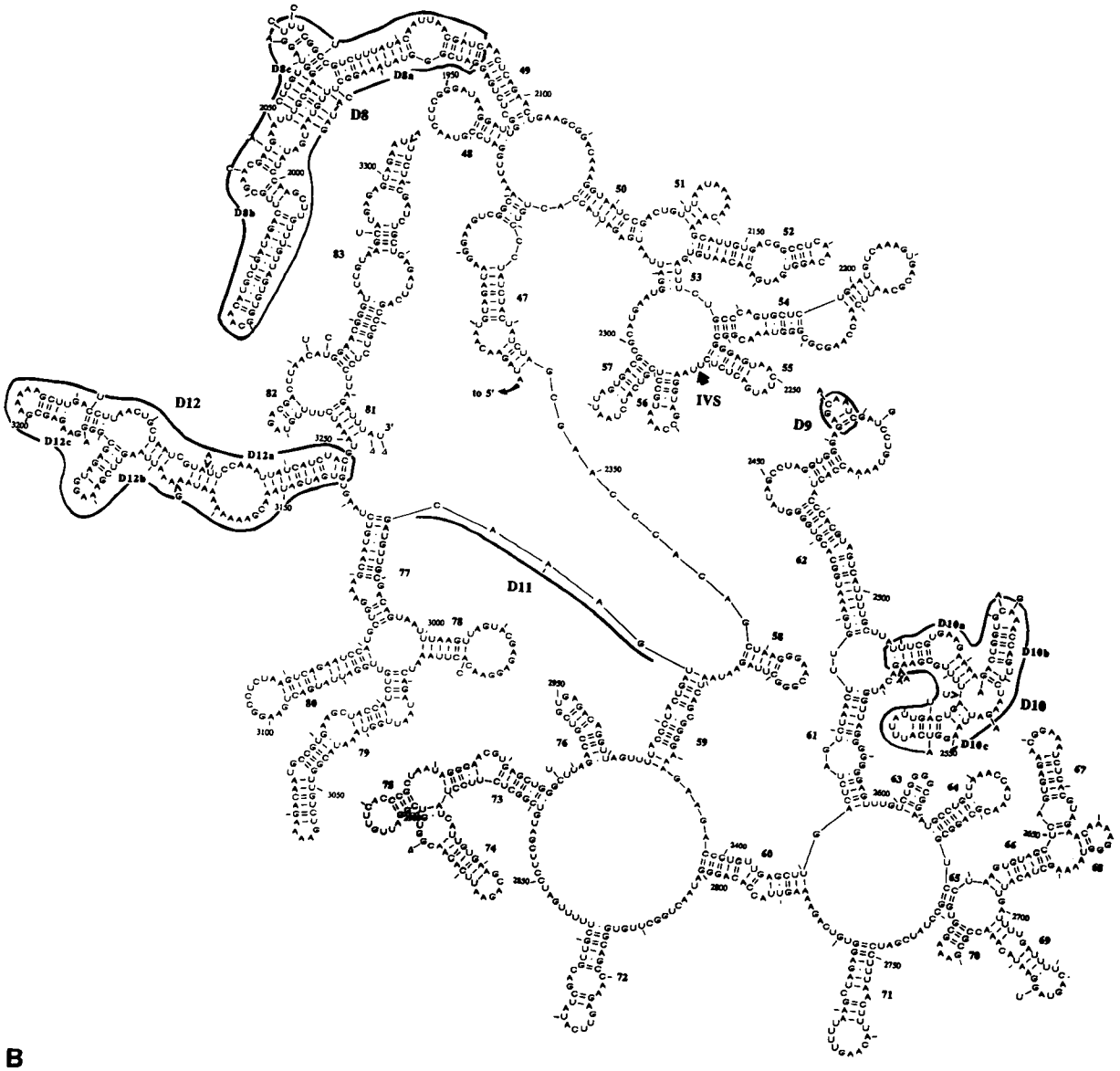

Fig. 3. Continued.

two insertions are this time located within the D12a stem.

In conclusion, the *T. pyriformis* sequence differs from that of *T. thermophila* by 58 transitions, 38 transversions, 9 point deletions plus 1 large one (11 nt in D1), 10 single insertions (6 of which are As), and 2 insertions of 2 consecutive nucleotides; however, the secondary structure of the expansion segments seems to be very conserved.

Evolutionary Aspects

It has been suggested that the expansion segments of eukaryotic rRNA represent the remains of ancient internally transcribed sequences (ITS) that separated functional domains of the ancestral gene (cf. Clark 1987; Spencer et al. 1987). Alternatively, expansion segments might represent later additions to the rRNA structure (cf. Gerbi 1985). Neither of these opposing views have predictive value with respect to the extent of sequence divergence of the ITS regions relative to that of the expansion segments, as mutational changes may accumulate at different rates in different regions of the genome (Hancock and Dover 1988) with the effect of masking the time of origin of a particular region. The sequence differences of the ETS, ITS1, and ITS2 regions between the two *Tetrahymenas* are 11%, 3%, and 10%, respectively (J. Engberg, unpublished), which is similar to the differences observed among the expansion segments

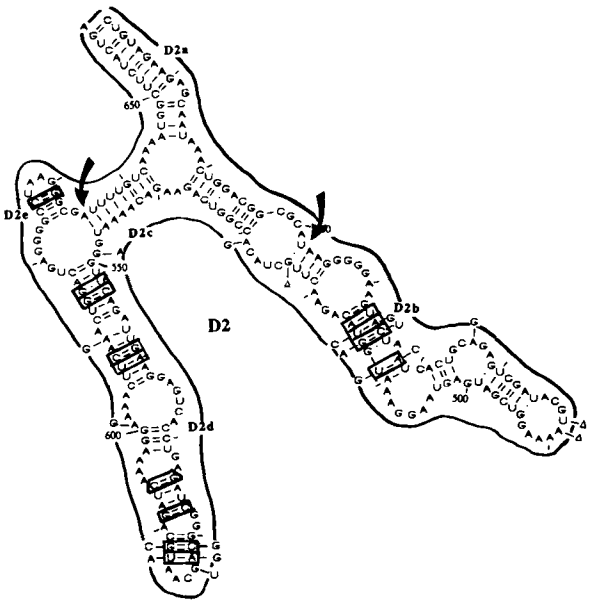

Fig. 4. Compilation of compensated base changes found within the D2 domain of LSU rRNA from 19 different *Tetrahymena* species. A portion of the LSU rRNA corresponding to the sequence between the two arrows was determined from 19 different *Tetrahymena* species by Preparata et al. (1989). When the differences among the sequences were positioned on our secondary structure map, 28 out of the 43 mutations observed were found to be compensated, and the remainder were located in single-stranded regions. The positions where compensatory base changes were observed have been indicated by boxes. In the box corresponding to positions 582 and 589, changes from A–T to G–C to T–A were observed among the different species.

of the two species. Because knowledge of mutational rates in the different regions is lacking, these data do not resolve the question of origin of the expansion segments.

Our data, which indicate that evolutionary constraints have operated to maintain the secondary structure of the expansion segments, may imply that these regions, although originally functionless, have been recruited for lineage-specific functions (cf. Raué et al. 1988). Alternatively, the maintained secondary structures may be a reflection of how the expansion segments have evolved. It has been speculated that slippage-like processes can generate internal repetitions which, in turn, would affect the RNA structure (Hancock and Dover 1988; Hancock et al. 1988). A better understanding of the above matters may be gained from studies involving functional tests of the rRNA.

Functional Aspects

Recently, Yao and coworkers have described a very elegant and promising approach to identifying functional regions of rRNA in *Tetrahymena* (Sweeney and Yao 1989). By injecting cloned copies of rDNA containing a selectable marker into developing macronuclei, they were able to recover transformants in which the injected rDNA had been processed into macronuclear rDNA and had completely replaced the host macronuclear rDNA. Using this system, they tested whether viable transformants could be obtained bearing rDNA in which foreign DNA had been inserted in vitro. Some of the 26S rRNA mutants failed to generate transformants but others succeeded. In one case a 119-bp linker insertion within the HindIII site at position 3204 (cf. Fig. 3) resulted in cells producing 26S rRNA containing the linker insertion and growing with a normal doubling time. The insertion in this rDNA mutant maps at the very tip of the D12 segment (stem D12C, cf. Fig. 3). On the other hand, it was not possible to obtain viable transformants containing the same fragments inserted into the HindIII site at position 2001 (stem D8b, Fig. 3). Clearly, this experimental approach is capable of defining structurally and functionally important domains within the LSU rRNA, and it is hoped that the proposed secondary structure map presented here will be helpful in the design of future mutants.

Acknowledgments. We are grateful to Dr. Jean-Pierre Bachellerie for providing the opportunity to draw the secondary structure, to Drs. Otto Dahl and Thorkild Christensen for providing oligonucleotides, to Prof. Dave Nanney for providing sequence data before publication, and to Mrs. Else Uhrenfeldt for typing this manuscript. The technical assistance of Mr. Franz Frenzel and Mrs. Naoko Hashimoto is greatly acknowledged.

The work was supported by the Danish Natural Science and Medical Science Research Councils, the Carlsberg Foundation, the Centre National de la Recherche Scientifique, France, and the Ministry of Education, Science and Culture, Japan.

This project was part of a program under the Centre for Medical Biotechnology, University of Copenhagen.

References

Baroin A, Perasso R, Qu LH, Brugerolle G, Bachellerie J-P, Adoutte A (1988) Partial phylogeny of unicellular eukaryotes based on rapid sequencing of a portion of the 28S ribosomal RNA. Proc Natl Acad Sci USA 85:3474–3478

Boer PH, Gray MW (1988) Scrambled ribosomal RNA gene pieces in *Chlamydomonas reinhardtii* mitochondrial DNA. Cell 55:399–411

Cech TR (1986) Ribosomal RNA gene expression in *Tetrahymena*: transcription and RNA splicing. In: Gall JG (ed) The molecular biology of ciliated protozoa. Academic Press, New York, pp 203–225

Clark CG (1987) On the evolution of ribosomal RNA. J Mol Evol 25:343–350

Din N, Engberg J, Gall JG (1982) The nucleotide sequence at the transcription termination site of the ribosomal RNA gene in *Tetrahymena thermophila*. Nucleic Acids Res 10:1503–1513

Eckert WA, Kaffenberger W, Krohne G, Franke WW (1978) Introduction of hidden breaks during rRNA maturation and ageing in *Tetrahymena pyriformis*. Eur J Biochem 87:607–616

Engberg J (1985) The ribosomal RNA genes of *Tetrahymena*. Structure and function. Eur J Cell Biol 36:133–151

Engberg J, Din N, Eckert WA, Kaffenberger W, Pearlman RE (1980) Detailed transcription map of extrachromosomal ribosomal RNA genes in *Tetrahymena thermophila*. J Mol Biol 142:289–313

Geliebter J (1987) Dideoxy nucleotide sequencing of RNA and uncloned cDNA. Focus 9:5–8

Gerbi SA (1985) Evolution of ribosomal RNA. In: MacIntyre RJ (ed) Molecular evolutionary genetics. Plenum, New York, pp 419–517

Gouy M, Li W-H (1989) Molecular phylogeny of kingdoms Animalia, Plantae, and Fungi. Mol Biol Evol 6:109–122

Gutell RR, Fox EF (1988) A compilation of large subunit RNA sequences presented in a structural format. Nucleic Acids Res [suppl] 16:r175–r269

Hancock JM, Dover GA (1988) Molecular coevolution among cryptically simple expansion segments of eukaryotic 26S/28S rRNAs. Mol Biol Evol 5:377–391

Hancock JM, Tautz D, Dover GA (1988) Evolution of the secondary structures and compensatory mutations of the ribosomal RNAs of *Drosophila melanogaster*. Mol Biol Evol 5:393–414

Heinonen TYK, Schnarre MN, Young PG, Gray MW (1987) Rearranged coding segments, separated by a transfer RNA gene, specify the two parts of a discontinuous large subunit ribosomal RNA in *Tetrahymena pyriformis* mitochondria. J Biol Chem 262:2879–2887

Higashinakagawa T, Saiga H, Shintani N, Narushima-Iio M, Mita T (1981) Localization of putative transcription initiation site on the cloned rDNA fragment of *Tetrahymena pyriformis*. Nucleic Acids Res 9:5905–5916

Karrer KM (1986) The nuclear DNA's af holotrichous ciliates. In: Gall JG (ed) The molecular biology of ciliated protozoa. Academic Press, New York, pp 85–110

Leffers H, Kjems J, Østergaard L, Lassen N, Garrett RA (1987) Evolutionary relationships amongst Archebacteria. J Mol Biol 195:43–61

Lenaers G, Nielsen H, Engberg J, Herzog M (1988) The secondary structure of large-subunit rRNA divergent domains, a marker for protist evolution. BioSystems 21:215–222

Lenaers G, Maroteaux L, Michot B, Herzog M (1989) Dinoflagellates in evolution. A molecular phylogeny analysis of large subunit ribosomal RNA. J Mol Evol 29:40–51

Michot B, Bachellerie J-P (1987) Comparisons of large subunit rRNA's reveal some eukaryotic-specific elements of secondary structure. Biochimie 69:11–23

Michot B, Hassouna N, Bachellerie J-P (1984) Secondary structure of mouse 28S rRNA and general model for the folding of the large rRNA in eukaryotes. Nucleic Acids Res 12:4259–4279

Niles EG, Cunningham K, Jain R (1981) Structure of the *Tetrahymena pyriformis* rRNA gene. J Biol Chem 256:12857–12860

Noller HF, Asire M, Barta A, Douthwaite S, Goldstein T, Gutell RR, Moazed D, Normanly J, Prince JB, Stern S, Triman K, Turner S, Van Stolk B, Weaton V, Weises B, Woese CR (1986) Studies on the structure and function of ribosomal RNA. In: Hardesty B, Kramer G (eds) Structure, function and genetics of ribosomes. Springer Verlag, Berlin, pp 143–163

Pedersen N, Hellung-Larsen P, Engberg J (1985) Small nuclear RNA's in the ciliate *Tetrahymena*. Nucleic Acids Res 13: 4203–4224

Preparata RM, Meyer EB, Preparata FP, Simon EM, Vossbrinck CR, Nanney DL (1989) The ribosomal phylogenies of the tetrahymenine ciliates. J Mol Evol 28:427–441

Raué HA, Klootwijk J, Musters W (1988) Evolutionary conservation of structure and function of high molecular weight ribosomal RNA. Progr Biophys Mol Biol 51:77–129

Sogin ML, Elwood HJ, Gunderson JH (1986a) Evolutionary diversity of eukaryotic small-subunit rRNA genes. Proc Natl Acad Sci USA 83:1383–1387

Sogin ML, Ingold A, Karlok M, Nielsen H, Engberg J (1986b) Phylogenetic evidence for the acquisition of ribosomal RNA introns subsequent to the divergence of some of the major *Tetrahymena* groups. EMBO J 5:3625–3630

Sollner-Webb B, Reeder R (1979) The nucleotide sequence of the initiation and termination sites for ribosomal RNA transcription in *X. laevis*. Cell 18:485–499

Spencer DF, Collings JC, Schnarre MN, Gray MW (1987) Multiple spacer sequences in the nuclear large subunit ribosomal RNA gene of *Crithidia fasciculata*. EMBO J 6:1063–1071

Sweeney R, Yao M-C (1989) Identifying functional regions of rRNA by insertion mutagenesis and complete gene replacement in *Tetrahymena thermophila*. EMBO J 8:933–938

Ware VC, Renkawitz R, Gerbi SA (1985) rRNA processing: removal of only nineteen bases at the gap between 28 α and 28 β rRNAs in *Sciara coprophila*. Nucleic Acids Res 13:3581–3597

Wild MA, Sommer R (1980) Sequence of a ribosomal RNA gene intron from *Tetrahymena*. Nature 283:693–694

Woese CR (1987) Bacterial evolution. Microbiol Rev 51:221–271

Yao M-C (1986) Amplification of ribosomal RNA genes. In: Gall JG (ed) The molecular biology of ciliated protozoa. Academic Press, New York, pp 179–201

Zuker M, Stiegler P (1981) Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information. Nucleic Acids Res 9:133–148

Received September 29, 1989/Revised November 24, 1989