Nucleic Acids Research, 1998, Vol. 26, No. 14 3333–3339

Comparative structure analysis of vertebrate ribonuclease P RNA

C. Pitulle, M. Garcia-Paris[1], K. R. Zamudio[1] and N. R. Pace*

Departments of Plant & Microbial Biology and Molecular & Cell Biology, University of California, 111 Koshland Hall, #371, Berkeley, CA 94720-3102, USA and [1]Museum of Vertebrate Zoology and Department of Integrative Biology, University of California, Berkeley, CA 94720-3102, USA

Received March 2, 1998; Revised and Accepted May 18, 1998

DDBJ/EMBL/GenBank accession nos+

ABSTRACT

Ribonuclease P cleaves 5′-precursor sequences from pre-tRNAs. All cellular RNase P holoenzymes contain homologous RNA elements; the eucaryal RNase P RNA, in contrast to the bacterial RNA, is catalytically inactive in the absence of the protein component(s). To understand the function of eucaryal RNase P RNA, knowledge of its structure is needed. Considerable effort has been devoted to comparative studies of the structure of this RNA from diverse organisms, including eucaryotes, primarily fungi, but also a limited set of vertebrates. The substantial differences in the sequences and structures of the vertebrate RNAs from those of other organisms have made it difficult to align the vertebrate sequences, thus limiting comparative studies. To expand our understanding of the structure of diverse RNase P RNAs, we have isolated by PCR and sequenced 13 partial RNase P RNA genes from 11 additional vertebrate taxa representing most extant major vertebrate lineages. Based on a recently proposed structure of the core elements of RNase P RNA, we aligned the sequences and propose a minimum consensus secondary structure for the vertebrate RNase P RNA.

INTRODUCTION

Ribonuclease P (RNase P) is the ubiquitous endonuclease that carries out maturation of the 5′ end of tRNA (see 1 and 2 for reviews). All cellular versions of RNase P from each of the three domains of life—Bacteria, Eucarya and Archaea—are ribonucleoproteins, consisting of one RNA and at least one protein. The RNA component of bacterial RNase P is catalytically active in the absence of protein *in vitro* (3), but the archaeal and eucaryal RNase P RNAs are not independently active. Although RNase P RNAs from diverse organisms exhibit extensive variation in sequence and structure, all contain five universally conserved regions (4) that are considered to reflect similar core structure in the RNAs and possibly homologous function.

The global structure of the bacterial type of RNase P RNA is increasingly well-characterized and there is considerable information on the archaeal counterpart; however, less is known about the eucaryotic version of the RNA. Accumulation of sequences for comparative structure analysis of the eucaryal RNase P RNA has focused mainly on yeasts (5,6). Animal RNAs for comparison have so far been represented by those from only three types of animals: mammals, one frog and one fish (7–10). Thus, the diversity of sequences available for comparative structure analysis has been insufficient to specify a sequence alignment and a basal structure for the vertebrate RNase P RNA.

To expand the diversity of vertebrate sequences available for comparison, we determined 13 partial sequences of RNase P RNAs from 11 of the main vertebrate lineages. We used PCR to isolate the main portions of the RNase P RNA genes from five frogs (*Ascaphus truei*, *Alytes obstetricans*, *Bufo bufo*, *Discoglossus jeanneae* and *Pelobates cultripes*), two salamanders (*Ensatina eschscholtzii* and *Euproctus asper*), two lizards (*Phrynosoma douglasi* and *Uta stansburiana*), one snake (*Lachesis muta*) and one fish (*Oncorhynchus* sp.). Although previously examined, the corresponding partial genes from *Homo sapiens*, *Xenopus laevis* and *Danio rerio* also were obtained by PCR and re-examined. We aligned the sequences consistently with the proposed universally conserved core of the RNA (4), and inferred a minimum-consensus secondary structure model of the vertebrate RNase P RNA.

MATERIALS AND METHODS

DNA isolation

DNA from frogs and salamanders was extracted as described previously (11). For the three reptiles, 20 mg of frozen tissue from each species was incubated overnight at 55°C in 1 ml of 100 mM EDTA, 10 mM Tris–HCl (pH 8), 10 mM NaCl, 1% SDS, 8 mg dithiothreitol and 2 mg proteinase K, and DNA was purified by phenol–chloroform extraction (12). Genomic DNAs were precipitated from 0.3 M NaCl with 2.5 vol ethanol and resuspended in 200 µl of 10 mM Tris–HCl (pH 8) and 1 mM EDTA. In the case of *H.sapiens*, DNA was isolated from 80 µl of

*To whom correspondence should be addressed. Tel: +1 510 643 2571; Fax: +1 510 642 4995; Email: nrpace@nature.berkeley.edu

+AF004371–AF004377, AF044325–AF044328, AF044330–AF044331, AF044737

whole blood treated with heparin (20 U/ml) as described earlier (13). Salmon testis DNA was purchased from Sigma.

PCR amplification of RNase P RNA genes

Each 100 µl PCR reaction contained 8 U of *Taq* DNA polymerase (Promega), 100 ng of each primer (below), 300 ng genomic DNA, 0.2 mM of each dNTP, 1.5 mM $MgCl_2$, 50 mM KCl, 10 mM Tris–HCl (pH 9 at 25°C) and 0.1% Triton X-100. All PCR reactions started with an initial denaturation step for 2 min at 94°C, followed by 20–25 amplification cycles (below), and a final extension step for 10 min at 72°C. To avoid non-specific PCR products, we chose the highest annealing temperature that resulted in PCR fragments of the expected size. PCR products were visualized by agarose gel electrophoresis. In case of low yield or if more than one PCR product was obtained, a fragment (~5 µl) of agarose containing the desired DNA fragments was excised from the gel and used in a second PCR amplification. The following primers (nucleotide positions correspond to the individual RNase P RNA) and cycle profiles (denaturation-primer annealing-extension) were used for the respective taxa: *X.laevis*, primers XEN-F (positions 1–33, 5′-AGATGTGGAGGGAAGCTCAG-TGCTGAGGCTACA-3′) and XEN-R (positions 273–241, 5′-CT-CAGCCATTGAGCTTGAATGTGCCAGCTGTGG-3′), 20 cycles (2 min at 92°C, 45 s at 72°C). The profile for reamplification was 20 cycles (1 min at 92°C, 45 s at 67°C, 30 s at 72°C); *D.rerio*, primers Euk-Z1 (positions 6–30, 5′-GCGGAAGGAAGCTCA-CTGTAGAGG-3′) and Euk-Z2 (positions 254–277, 5′-CTAGC-ATGTTATGCTAGCTCACAC-3′), 25 cycles (1.5 min at 92°C, 45 s at 55°C, 45 s at 72°C). Reamplification was performed under the same conditions, except the annealing temperature was 50°C; *H.sapiens*, primers Euk-D2 (positions 4–36, 5′-GGGCGGAGGG-AAGCTCATCAGTGGGGCCACGAG-3′) and HeLa-A (positions 336–304, 5′-GCGGAGGAGAGTGGTCTGAATTGGGTTATG-AGG-3′), 25 cycles (2 min at 92°C, 45 s at 72°C); *Oncorhynchus* sp., primers Euk-D (7) (positions 1–26, 5′-GGGCGGAGGGAAG-CTCATCAGTGGGG-3′) and Euk-A (7) (positions 308–278 and 312–282, 5′-CGGAATTCCGGAGAGTGGTCTGAATTGGG-TT-3′), 20 cycles (1.5 min at 92°C, 45 s at 50°C, 1 min at 72°C). Reamplification was performed under the same conditions; *A.truei*, primers Euk-Z1 (positions 1–24) and Euk-A (7) (positions 320–290), 25 cycles (2 min at 92°C, 1 min at 47°C, 45 s at 72°C). Reamplification was performed under the same conditions, except the annealing temperature was 49°C; *A.obstetricans*, primers Euk-Z1 (positions 1–24) and Euk-A (7) (positions 280–310), 25 cycles (2 min at 92°C, 1 min at 49°C, 45 s at 72°C); *B.bufo*, primers Euk-Z1 (positions 1–24) and Euk-A (7) (positions 391–361), 25 cycles (2 min at 92°C, 1 min at 48°C, 45 s at 72°C); *D.jeanneae*, primers Euk-Z1 (positions 1–24) and Euk-A (7) (positions 406–376), 25 cycles (2 min at 92°C, 1 min at 48°C, 45 s at 72°C); *E.asper*, primers Euk-Z1 (positions 1–24) and Euk-A (7) (positions 303–273), 25 cycles (2 min at 92°C, 1 min at 48°C, 45 s at 72°C). Reamplification was performed under the same conditions; *E.eschscholtzii*, primers Euk-Z1 (positions 1–24) and Euk-A (7) (positions 326–296), 25 cycles (2 min at 92°C, 1 min at 50°C, 45 s at 72°C); *L.muta*, primers Euk-Z1 (positions 1–24) and Euk-A (7) (positions 330–300 and 348–318), 25 cycles (2 min at 92°C, 1 min at 49°C, 45 s at 72°C); *P.cultripes*, primers Euk-D2 (positions 1–33) and Euk-A (7) (positions 310–280), 25 cycles (2 min at 92°C, 1 min at 49°C, 45 s at 72°C); *P.douglasi*, primers Euk-Z1 (positions 1–24) and Euk-A (7) (positions 331–301), 25 cycles (2 min at 92°C, 1 min at 50°C, 45 s at 72°C); *U.stansburiana*, primers Euk-Z1 (positions 1–24) and Euk-A (7) (positions 328–298), 25 cycles (2 min at 92°C, 1 min at 49°C, 45 s at 72°C).

In case of sequence ambiguities or heterogeneity, the PCR reaction was repeated under the same conditions, except that 2.5 U of *Pfu* DNA polymerase (Stratagene) and 1× PCR buffer (Stratagene) were used.

Cloning of PCR products

PCR fragments were cloned directly using the pGEM®-T Vector System (Promega), the TA Cloning® Kit (Invitrogen) or the TOPO TA Cloning® Kit (Invitrogen), following the protocols provided by the manufacturers. PCR products derived from *Pfu* DNA polymerase were blunt-end ligated into pGEM®5Zf(+) vector for 1 h at 37°C in 10 µl containing 50 ng of *Eco*RV digested pGEM®5Zf(+), 70 ng of PCR product, 1 Weiss U T4 DNA ligase (Promega) and 1× ligase buffer (Promega). In the case of the pGEM®-T Vector and pGEM®5Zf(+) vector, each ligation reaction was purified over a Chromabond-100 TE spin column (Clontech) prior to the transformation of *Escherichia coli* DH5α by electroporation with an *E.coli* Pulser Transformation Apparatus (Bio-Rad). Preparation of the *E.coli* DH5α cells and electroporation were performed as described (14) using 20 ng of recombinant plasmid DNA per 50 µl of cell suspension. Recombinants were selected by blue/white color of colonies.

Plasmid isolation and sequencing

Plasmid isolation was performed using the QIAprep™ Plasmid Kit (Qiagen). Recombinant plasmid DNA (200 ng) was sequenced by cycle sequencing (24 cycles: 30 s at 96°C; 15 s at 45°C, 4 min at 60°C) with AmpliTaq®FS (Perkin Elmer) using 20 ng primer T7 (5′-GTAATACGACTCACTATAGGG-3′) and primer M13-rev (5′-AGGAAACAGCTATGA-3′) or primer SP6 (5′-ATTTAGGT-GACACTATAG-3′). Sequences were analyzed on an ABI 377 or ABI 373A automated DNA sequencer (Applied Biosystems) according to the manufacturer's protocols. Sequences were deposited in GenBank under the accession numbers AF004373 (*A.truei*), AF004374 (*A.obstetricans*), AF044737 (*B.bufo*), AF004372 (*D.jeanneae*), AF004375 (*E.asper*), AF044328 (*E.eschscholtzii*), AF004376 (*L.muta*), AF044326 (*L.muta* with insertion), AF004377 (*Oncorhynchus* sp.), AF044331 (*Oncorhynchus* sp. with deletion), AF004371 (*P.cultripes*), AF044327 (*P.douglasi*), AF044325 (*U.stansburiana*) and AF044330 (*X.laevis*).

Sequence alignment, design of secondary structure and covariation

All available RNase P RNA sequences from vertebrates were aligned manually using the GDE sequence editor (15). Comparative analysis of the secondary structures by covariation was performed as described earlier (16–18). The nomenclature of structural elements in eucaryal RNase P follows that described for bacterial RNase P RNA (4,19).

RESULTS

As detailed in Materials and Methods, we amplified, cloned and sequenced 13 partial (~85% of full-length) RNase P RNA genes from 11 animals, more than doubling the previously known number of vertebrate sequences for this RNA and substantially

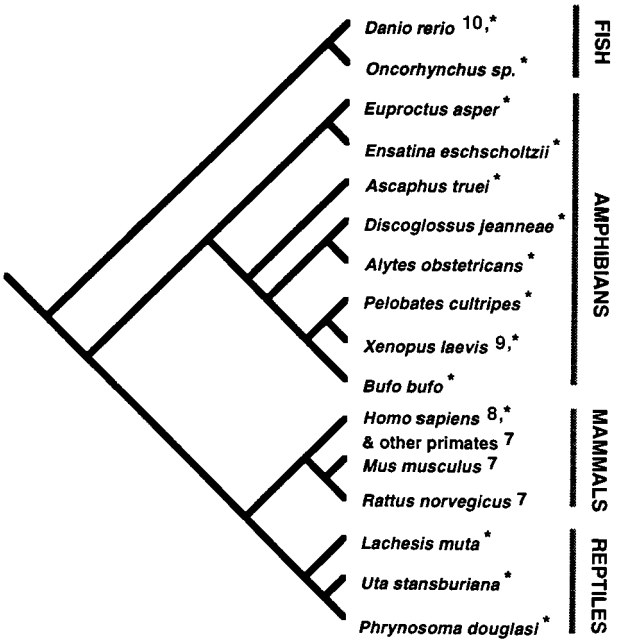

Figure 1. Diagrammatic representation of phylogenetic relationships of vertebrates with available RNase P RNA sequences. Relationships are based on morphological and molecular characters among extant vertebrates (22,29). Taxa were chosen for this study (text) to span a broad extent of vertebrate lineages and thereby delineate the core secondary structure of vertebrate RNase P RNA. References to previously determined sequences are shown as superscripts. Asterisks indicate sequences determined or confirmed in this study.

expanding the breadth of vertebrate diversity sampled for comparative analysis of the RNA. Figure 1 illustrates the phylogenetic distribution of vertebrate RNase P RNA sequences currently available for comparative analysis following this study. As described below, two types of RNase P RNA genes were observed for *Oncorhynchus* sp. and *L.muta*. Based on the sizes of PCR-products and those of known full-length RNAs, the probable sizes for the examined RNase P RNAs range from 303 nt (*E.asper*) to 406 nt (*D.jeanneae*). Taxa represented by the newly determined sequences are especially informative for structural comparisons because they include members of almost all lineages of vertebrates (20–22). Multiple sampling within some of the clades focuses on major phylogenetic splits within these groups. An alignment of known vertebrate RNase P RNA sequences, based on secondary structure, is shown in Figure 2 and available in the RNase P database (http://JWBrown.mbio.ncsu.edu/RNaseP/).

The structures of the vertebrate RNase P RNAs conform to the general eucaryotic model (4) and were consistent with numerous covariations in the amplified regions of the genes. One example of the derived secondary structures of vertebrate RNase P RNA, with a bacterial version of the RNA for comparison, is shown in Figure 3. PCR primers used were complementary to sequences near the 5′ and 3′ ends of the full-length RNA, so the corresponding newly determined sequences are not available for comparative analysis. In the cases of helices P2, P3a and P4 (Fig. 3), only one of the paired sequences represented primer-binding regions, so the complementary sequence was obtained. Although these sequences were excluded from comparative considerations, they allow reasonable inference of the corresponding complementarities and hence reconstruction of the full-length structures (not shown). Figure 4 shows a minimum-consensus structure of the vertebrate RNase P RNA and examples of variable structural elements in the RNAs analyzed. As is the case with other eucaryal RNase P RNAs, regions homologous to the bacterial and archaeal helices P5 and P15/16 (19,23) are not present in the vertebrate RNAs. Helices P3b, P7, P8, P9, P10 and P19 were supported by covariation in at least 2 bp. Length variations in the RNA seem to occur mainly in helix P12 (63 nt in *P.cultripes* to 143 nt in *D.jeannae*). The extent of variation is sufficient that the sequences that constitute this helix are not necessarily clear homologs except for their position in the core structure. Consequently, sequences corresponding to helix P12 are omitted from the alignment in Figure 2. Substantial variation in structure and sequence length in helix P12 also is extensive among the bacterial and archaeal RNase P RNAs. Less extensive variation is seen in helix P3 (38 nt in *E.asper* to 51 nt in *A.obstetricans* and *D.jeanneae*), and in helix P19 (12 nt in *X.laevis* to 35 nt in *A.obstetricans* and *D.jeanneae*).

The vertebrate RNAs display all of the conserved regions (CR I–V, Fig. 4) that contain universally conserved residues in RNase P. The proposed structure of the eucaryal helix P4 shown in Figures 3 and 4 is revised from the previous proposal (4) to accommodate the bulged U residue as seen in the bacterial and archaeal versions of the RNA. This requires the assumption of a G/A pair adjacent to the universally present bulged U residue instead of the G/C pair normally seen in the bacterial or archaeal RNAs, which we believe is justified to maintain the alignment of the CR sequences. All CR sequences examined are identical to the proposed consensus sequences (4) except CR II (AGNNA), which shows some variation. In the case of both *L.muta* RNase P RNAs, CR II is ACCAA, and in *A.truei* and *P.cultripes* CR II is AUCAA. Independent PCR reactions gave the same results. Covariations in the sequence set additionally indicate a 1 bp extension on helix P10/11 compared to previously proposed structures (4) (Figs 3 and 4).

We detected two types of RNase P RNA genes from three of the vertebrates: in two out of nine clones for the *Oncorhynchus* sp. RNA we observed a deletion of positions 117–120 in P9. We also discovered a triple-repeat of positions 187–193 in three out of 15 clones for the *L.muta* gene. This results in an extension of helix P12 in that type of RNA. Minor RNAs were designated *Oncorhynchus* sp. 2 and *L.muta* 2, respectively. These variant types of RNase P genes were all detected in independent PCR amplifications, and presumably represent genomic diversity in alleles or sequence repeats.

The published RNase P RNA sequences of *H.sapiens* (8) and *D.rerio* (10) were confirmed. In the case of the *X.laevis* sequence (9), a few ambiguities were clarified. The detail of the recently proposed secondary structure of helix P3 of human and zebrafish RNase P RNA (4) was adjusted (Fig. 2).

DISCUSSION

Earlier proposed secondary structures for vertebrate RNase P RNAs (4,7,10) were based on consideration of only a few sequences. To provide further examples for comparison, we chose organisms to cover a broad phylogenetic diversity (Fig. 1) and identified 13 additional, partial RNase P RNA sequences from 11 vertebrates using PCR with primers that allowed for the determination of the sequence of ~85% of the molecules. We are confident that all the PCR products represent RNase P RNA genes because of the presence of the conserved regions and the

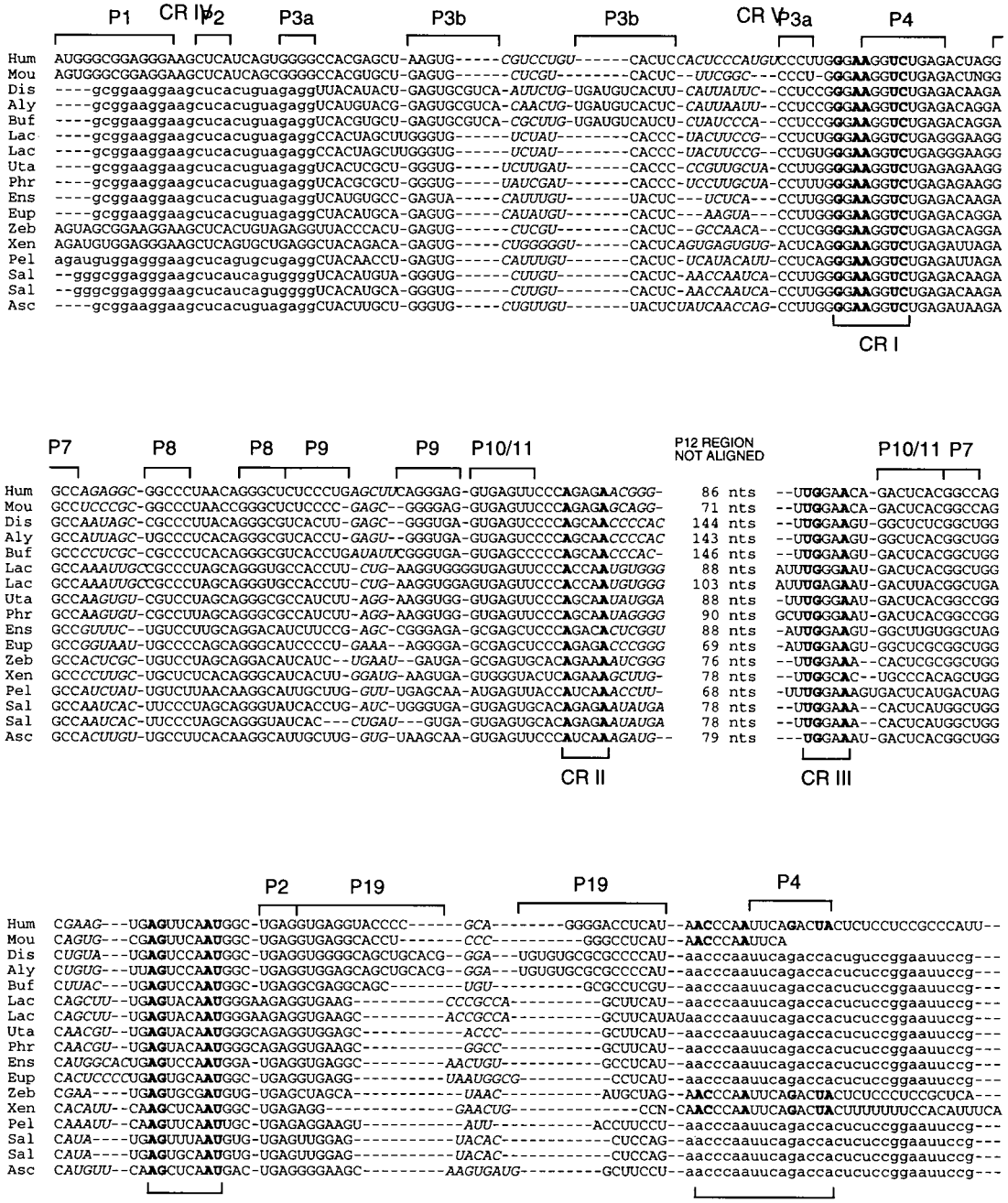

Figure 2. Alignment of RNase P RNAs including 16 partial sequences investigated or confirmed in this study. Brackets above the alignment indicate helices in the corresponding RNA structures. Sequences derived from PCR primers are in lower case letters. Regions that could not be aligned with confidence are in italics. Bold letters in brackets below the alignment indicate universally conserved nucleotides that define the conserved regions (CR) I–V. The sequences in the variable region P12 could not be aligned so were omitted from the alignment and represented by the total number of nucleotides. Abbreviations are: Hum, *H.sapiens* (8,*); Mou, *Mus musculus* (7); Dis, *D.jeanneae*; Aly, *A.obstretricans*; Buf, *B.bufo* (*); Lac, *L.muta* (*); Uta, *U.stansburiana* (*); Phr, *P.douglasi* (*); Ens, *E.eschscholtzii* (*); Eup, *E.asper* (*); Zeb, *D.rerio* (10,*); Xen, *X.laevis* (9,*); Pel, *P.cultripes* (*); Sal, *Onchorhynchus* sp. (*); Asc, *A.truei* (*). (Asterisks indicate organisms for which sequences were determined or confirmed by this study.)

consistency of the sequences with the proposed secondary structures based on entire sequences previously reported. The vertebrate RNase P RNAs conform to the generalized eucaryal RNase P secondary structure (4). Although the CR elements (4) are present in all vertebrate RNase P RNAs, the second position of CRII (AGNNA) seems to be variable and should, therefore, in this group be defined as ANNNA.

Despite the fact that the vertebrate RNAs lack some structural elements present in the catalytically active bacterial RNA, the new sequences offer no new clues as to why the eucaryal RNase P RNAs are catalytically inactive without the protein components of the native holoenzyme. For instance, the eucaryal RNAs lack the bacterial helix P15/16, an internal loop of which interacts with the 3′-terminal CCA of pre-tRNA (24,25). On the other hand,

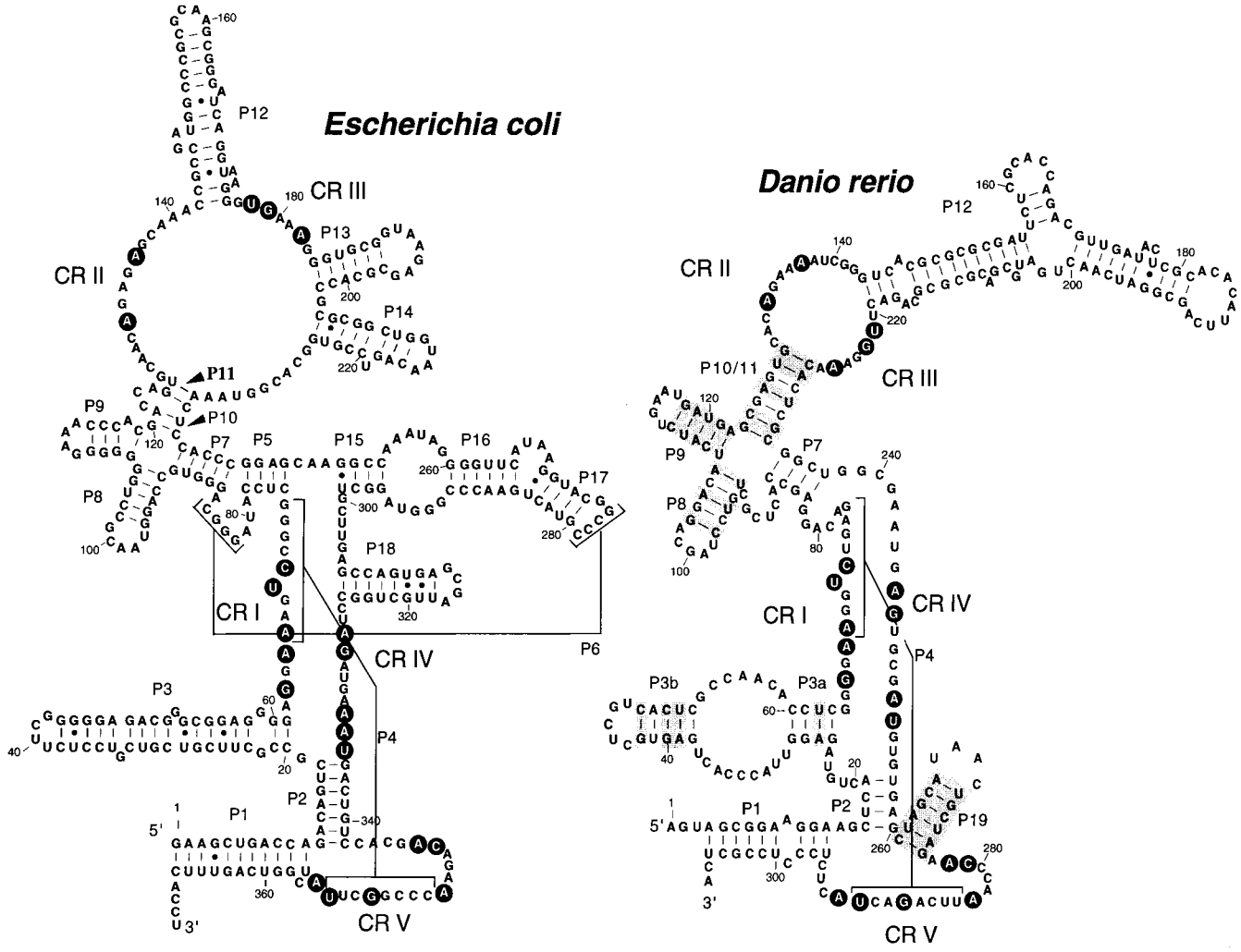

Figure 3. Comparison of proposed secondary structures of RNase P RNAs of *E.coli* (bacterial) and *D.rerio* (vertebrate). The nomenclature for the structural elements of *D.rerio* RNase P RNA is based on putatively homologous bacterial counterparts (4,19). Homologs to the bacterial helices P5, P6, P13, P14, P15, P16, P17 and P18 are not identifiable in the RNase P RNAs of *D.rerio* or other vertebrates. In contrast, the region of P3 is more complex in vertebrate and other eucaryotic RNase P RNAs than in the bacterial version. Watson–Crick base pairs are indicated by dashes, and G/U base pairs by dots. Gray shading indicates base pairs confirmed by covariation among the vertebrate RNAs. The structure of helix P12 of the *D.rerio* RNA shown is based on results with the computer program Mfold (30) and is not confirmed by covariation. Highlighted bases are universally conserved components of conserved regions (CRs) in all RNase P RNAs.

some structural elements of eucaryal RNase P RNA are more complex than their bacterial and archaeal counterparts. The eucaryal helix P3, for instance, consists of two short helices (P3a and P3b) connected by an internal loop (Fig. 4). In contrast, in Bacteria and Archaea helix P3 is manifested as a single helix of variable length. It was speculated earlier (4) based on their approximately equivalent positions in tertiary structure models of the bacterial RNAs, that the P3 internal loop in the eucaryotic version of RNase P RNA may compensate for the loss of the bacterial P15/16 loop that participates in binding tRNA. The newly determined sequences are consistent with this idea in their lack of the bacterial helix P15/16 and content of the compound P3.

Helix P19 is present in all vertebrate RNase P RNAs examined so far. In contrast, it is sometimes absent from bacterial and archaeal RNAs. The function of helix P19, if any, is not known. Commonly, however, the terminal loops of such ephemeral helices dock elsewhere in the RNA to bolster the superstructure (19).

The numbers of gene copies for RNase P RNA that occur in the genomes of the vertebrates for which sequences have been determined are not known. Multiple homologs of RNase P RNA genes have been detected in genomic DNA of mouse, rat, rabbit and human (26). Three genes analyzed from mouse are truncated within CRIV and therefore do not contain full length RNase P RNA genes. Although the corresponding RNAs were detected (26), it seems unlikely that these truncated genes code for functional RNase P RNAs. Some sequence degeneracies were detected in the present study in multiple PCR reactions. Twenty percent of the RNase P RNA genes obtained from *Oncorhynchus* sp. show a deletion of positions 117–120 in the region of P9. In the case of *L.muta*, 20% of the RNase P RNA genes obtained have a triple repeat of nucleotides 187–193 in P12. Those types of repeats also have been discovered recently in helix P12 of several cyanobacterial RNase P RNAs (27). Such sequence idiosyncrasies might indicate the occurrence of more than one gene-copy for

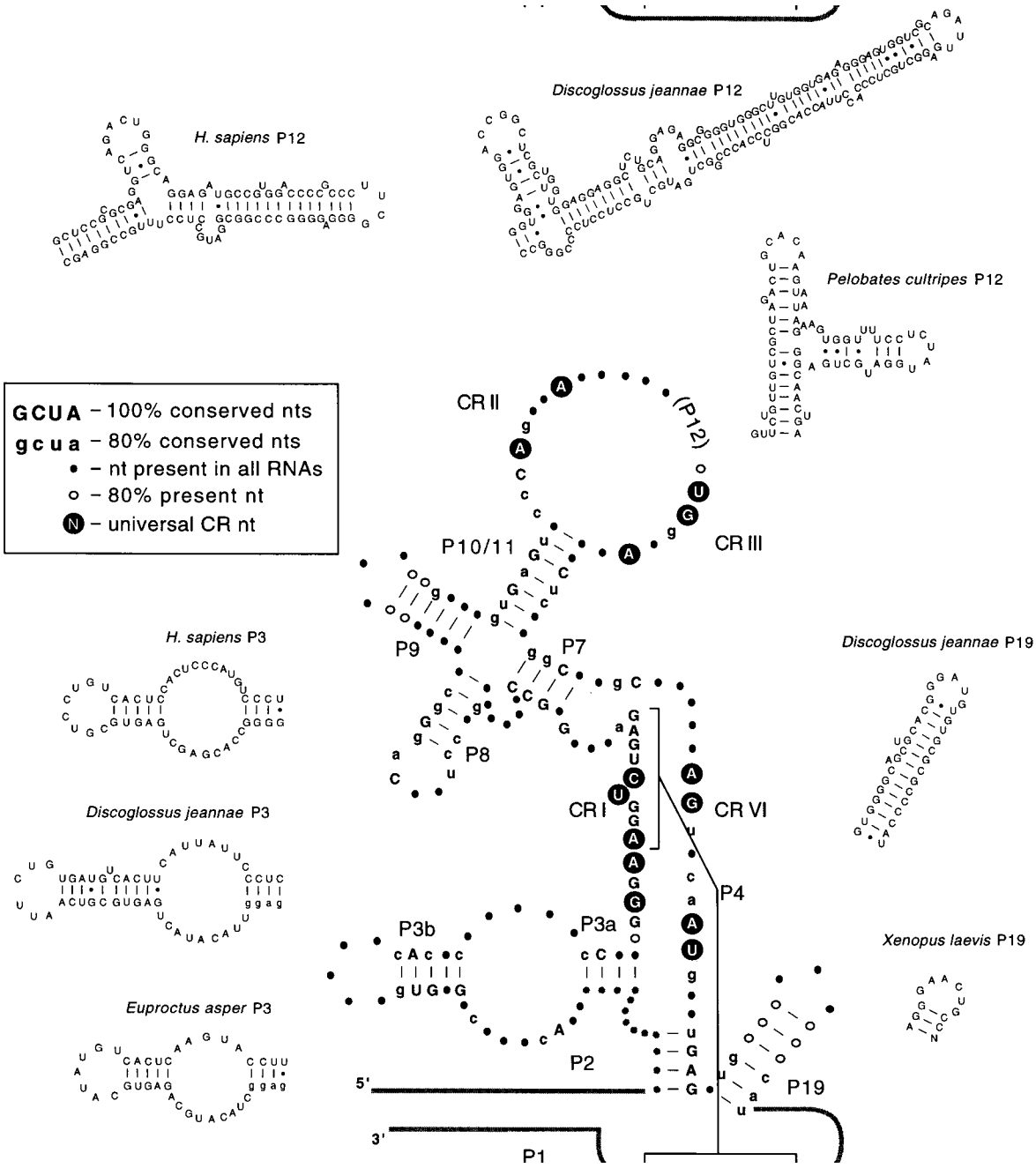

Figure 4. Variation in vertebrate RNase P RNA secondary structure. The proposed minimum consensus secondary structure for vertebrate RNase P RNA is shown in the center. The nomenclature for the structure is based on putatively homologous bacterial counterparts (4,19). Different possible structures for P3, P12 and P19 from the organisms identified illustrate size and structural variation among vertebrate RNase P RNAs. Structures for P12 helices, which because of sequence length variation cannot be aligned for detection of covariations, are based on results using the computer program Mfold (30). Watson–Crick base pairs are indicated by dashes, and G/U base pairs by dots. Gray lines indicate regions that bind PCR primers. Highlighted bases are universally conserved regions in all RNase P RNAs.

RNase P RNA in these vertebrates or represent two alleles. Appropriate Southern hybridization studies have not been conducted to determine copy-numbers and organization of these genes.

Partial or complete RNase P RNA sequences from mammals, reptiles, amphibia and fish provide broad representation of this RNA among vertebrates. We were not able to obtain the corresponding gene from bird (two chickens, *Gallus domesticus* and *Callipela californica*, and two hawks, *Buteo magnirostis* and *Zonotrichia atricapilla* tested), and from turtle (*Clemmys marmorata* and *Chelonia mydas* tested), presumably because of primer-incompatibility. Since birds are derived from deeply within the reptiles, however, we believe that the available sequences broadly represent the vertebrate RNase P RNA as summarized in the proposed minimum consensus secondary

structure (Fig. 3). Although these findings significantly advance our understanding of the structure of the vertebrate RNase P RNA, we note that the vertebrate RNA does not necessarily represent that of animals in general; most animals (>95%) are invertebrates (28). Therefore, further perspective on the evolution of animal RNase P will require comparison of the RNase P RNAs from more deeply divergent lineages of animals than represented by vertebrates.

ACKNOWLEDGEMENTS

We thank A.Barth and M.Dunaway for the gifts of *D.rerio* and *X.laevis* DNAs, respectively. Many thanks to J.W.Brown, G.E.Fox, J.K.Harris and M.A.Tanner for helpful discussions. This work was supported by an NIH grant to N.R.P.

REFERENCES

1 Altman,S., Kirsebom,L. and Talbot,S. (1993) *FASEB J.*, **7**, 7–14.
2 Frank,D.N. and Pace,N.R. (1998) *Annu. Rev. Biochem.*, in press.
3 Guerrier-Takada,C., Gardiner,K., Marsh,T., Pace,N.R. and Altman,S. (1983) *Cell*, **35**, 849–857.
4 Chen,J.-L. and Pace,N.R. (1997) *RNA*, **3**, 557–560.
5 Sbisà,E., Pesole,G., Tullo,A. and Saccone,C. (1996) *J. Mol. Evol.*, **43**, 46–57.
6 Tranguch,A. and Engelke,D. (1993) *J. Biol. Chem.*, **268**, 15045–15055.
7 Altman,S., Wesolowski,D. and Puranam,R.S. (1993) *Genomics*, **18**, 418–422.
8 Baer,M., Nilsen,T.W., Costigan,C. and Altman,S. (1990) *Nucleic Acids Res.*, **18**, 97–103.
9 Doria,M., Carrara,G., Calandra,P. and Tocchini-Valentini,G.P. (1991) *Nucleic Acids Res.*, **19**, 2315–2320.
10 Eder,P.S., Srinivasan,A., Fishman,M.C. and Altman,S. (1996) *J. Biol. Chem.*, **271**, 21031–21036.
11 Miller,S.A., Dykes,D.D. and Polesky,H.F. (1988) *Nucleic Acids Res.*, **16**, 215.
12 Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
13 Kawasaki,E.S. (1990) In Innis,M.A., Gelfand,D.H., Sminsky,J.J. and White,T.J. (eds), *PCR Protocols: A Guide to Methods and Applications*. Academic Press Inc., pp. 146–152.
14 Calvin,N.M. and Hanawalt,P.C. (1988) *J. Bacteriol.*, **170**, 2796–2801.
15 Maidak,B.L., Larsen,N., McCaughey,M.J., Overbeek,R., Olsen,G.J., Fogel,K., Blandey,J. and Woese,C.R. (1994) *Nucleic Acids Res.*, **22**, 3485–3487.
16 Fox,G.E. and Woese,C.R. (1975) *Nature*, **256**, 505–507.
17 Gutell,R.R. (1993) *Curr. Opin. Struct. Biol.*, **3**, 313–322.
18 James,B.D., Olsen,G.J. and Pace,N.R. (1989) *Methods Enzymol.*, **180**, 227–239.
19 Haas,E.S., Brown,J.W., Pitulle,C. and Pace,N.R. (1994) *Proc. Natl Acad. Sci. USA*, **91**, 2527–2531.
20 Ford,L. and Cannatella,D. (1993) *Herpetological Monographs*, **7**, 94–117.
21 Gauthier,J., Kluge,A.G. and Rowe,T. (1988) *Cladistics*, **4**, 105–209.
22 Rieppel,O. and de Braga,M. (1996) *Nature*, **384**, 453–455.
23 Haas,E.S., Armbruster,D.W., Vucson,B.M., Daniels,C.J. and Brown,J.W. (1996) *Nucleic Acids Res.*, **24**, 1252–1259.
24 Kirsebom,L.A. and Svard,S.G. (1994) *EMBO J.*, **13**, 4870–4876.
25 Oh,B.-K. and Pace,N.R. (1994) *Nucleic Acids Res.*, **22**, 4087–4094.
26 Li,K. and Williams,R.S. (1995) *J. Biol. Chem.*, **270**, 25281–25285.
27 Vioque,A. (1997) *Nucleic Acids Res.*, **25**, 3471–3477.
28 Wilson,E.O. (1992) *The Diversity of Life*. Belknap Harvard University Press, Cambridge, MA.
29 Pough,F.H., Heiser,J.B. and McFarland,W.N. (1989) *Vertebrate Life*. MacMillan Publishing Company, New York (3rd edition).
30 Zuker,M. (1989) *Science*, **244**, 48–52.