J Mol Evol (2000) 51:404–415
DOI: 10.1007/s002390010103

JOURNAL OF MOLECULAR EVOLUTION

Mitochondrial DNA of *Hydra attenuata* (Cnidaria): A Sequence That Includes an End of One Linear Molecule and the Genes for l-rRNA, $tRNA^{f\text{-}Met}$, $tRNA^{Trp}$, COII, and ATPase8

Genevieve Pont-Kingdon,[1,*] Cecile G. Vassort,[1] Rahul Warrior,[2,†] Ronald Okimoto,[1,‡] C. Timothy Beagley,[1] David R. Wolstenholme[1]

[1] Department of Biology, University of Utah, Salt Lake City, UT 84112, USA
[2] Department of Embryology, Carnegie Institution of Washington, 115 West University Parkway, Baltimore, MD 21210, USA

Received: 9 March 2000 / Accepted: 24 July 2000

Abstract. The 3231-nucleotide-pair (ntp) sequence of one end of one of the two linear mitochondrial (mt) DNA molecules of *Hydra attenuata* (phylum Cnidaria, class Hydrozoa, order Anthomedusae) has been determined. This segment contains complete genes for $tRNA^{f\text{-}Met}$, l-rRNA, $tRNA^{Trp}$, subunit 2 of cytochrome *c* oxidase (COII), subunit 8 of ATP synthetase (ATPase8), and the 5′ 136 ntp of ATPase6. These genes are arranged in the order given and are transcribed from the same strand of the molecule. As in two other cnidarians, the hexacorallian anthozoan *Metridium senile* and the octocorallian anthozoan *Sarcophyton glaucum*, the mt-genetic code of *H. attenuata* is near standard. The only modification appears to be that TGA specifies tryptophan rather than termination. Also as in *M. senile* and *S. glaucum*, the encoded *H. attenuata* mt-$tRNA^{f\text{-}Met}$ has primary and secondary structural features resembling those of *Escherichia coli* initiator $tRNA^{t\text{-}Met}$. As the encoded mt-$tRNA^{Trp}$ cannot be folded into a totally orthodox secondary structure, two alternative forms are suggested. The encoded *H. attenuata* mt-l-rRNA is 1738 nt, which is 451 nt shorter than the *M. senile* mt-l-rRNA. Comparisons of secondary structure models of these two mt-l-rRNAs indicate that most of the size difference results from loss of nucleotides in the *H. attenuata* molecule at a minimum of 46 locations, which includes elimination of six distinct helical elements.

Key words: *Hydra attenuata* — Cnidaria — Mitochondrial genes — Nucleotide sequences — Genetic code — Transfer RNA genes — Large ribosomal rRNA gene

Introduction

The mitochondrial (mt) genomes of most Metazoa (multicellular animals) comprise a single circular molecule of 14 to 20 kb. The more than 90 metazoan mtDNAs that have been sequenced from both vertebrates and invertebrates have been found to have a remarkably constant gene content. With only a small number of exceptions, each contains the genes for 13 energy pathway proteins, the 2 RNA components of mitochondrial ribosomes, and 22 transfer RNAs (tRNAs). However, within the Metazoa, differences in order of these genes occur that are related in extent to evolutionary distance (Wolstenholme 1992a; Okimoto et al. 1992; Boore and Brown 1994, 1995; Wolstenholme and Fauron 1995; Asakawa et al. 1995; Krettek et al. 1995; Beagley et al. 1996, 1998; Zardoya and Meyer 1997; Boore 1999).

A number of unusual features characterize metazoan mt-genomes (Wolstenholme 1992a, 1992b). These include a gene arrangement that is extremely compact due

* *Present address:* Huntsman Cancer Institute, University of Utah, Salt Lake City, UT 84112, USA
† *Present address:* Department of Biological Sciences, University of Southern California, Los Angeles, CA 90089, USA
‡ *Present address:* Department of Poultry Science, University of Arkansas, Fayetteville, AR 72701, USA
Correspondence to: D.R. Wolstenholme; *e-mail:* Wolstenholme@biology.utah.edu

to there being few or no nucleotides between most gene pairs; genetic codes that contain up to five changes in codon specificity relative to the standard (universal) code; unorthodox translation initiation codons; post-transcriptional generation of translation termination codons; genes for tRNAs of diverse structural form; and replication by a unique asymmetrical mode (Wolstenholme 1992a, 1992b; Clayton 1992; Wolstenholme and Fauron 1995; Shadel and Clayton 1997).

The complete nucleotide sequences of the mt-genomes of two cnidarians, the sea anemone *Metridium senile* (class Anthozoa, subclass Hexacorallia, order Actiniaria), and a soft coral, *Sarcophyton glaucum* (class Anthozoa, subclass Octocorallia, order Alcyonacea) have been reported (Pont-Kingdon et al. 1994, 1995, 1998; Beagley et al. 1998; Beaton et al. 1998). Also, the complete gene content and order in the mtDNA of another octocorallian, the sea pansy *Renilla kollikeri* (order Pennatulacea) has been determined (Beagley et al. 1995). Each of these three cnidarian mt-genomes comprise a single circular molecule in the range 17.5 to 19.0 kb. The protein and rRNA genes encoded are similar to those of other metazoan mtDNAs, except that both octocorallian mtDNAs include a gene for a homolog of MutS, an integral protein in the *Escherichia coli* postreplicative mismatch repair pathway (Pont-Kingdon et al. 1995, 1998; Beagley et al. 1995). Also, and in striking contrast to all other metazoan mtDNAs, *M. senile* mtDNA encodes only two tRNAs ($tRNA^{f\text{-}Met}$ and $tRNA^{Trp}$), and *S. glaucum* and *R. kollikeri* mtDNAs encode only a single tRNA ($tRNA^{f\text{-}Met}$). *M. senile* mtDNA has a further unique feature relative to all other completely sequenced metazoan mtDNAs in that two protein genes, COI and ND5, each contain a group I intron. The COI intron includes a gene for a homing endonuclease as has been found for some other group I introns, but the ND5 intron is unusual in that it contains the only copies of the ND1 and ND3 genes present in the *M. senile* mt-genome (Beagley et al. 1996, 1998).

The first exception to the circularity of metazoan mtDNAs was reported for the mtDNA of the cnidarian *Hydra attenuata* (class Hydrozoa, order Anthomedusae), which was found to occur as two linear molecules of approximately 8.1 and 8.2 kb (Warrior and Gall 1985). Later it was shown that mtDNAs of other species of *Hydra* also were in the form of two 8-kb linear molecules, and that mtDNAs of other Hydrozoa and Scyphozoa were in the form of a single 14–18-kb linar molecule (Bridge et al. 1992).

In this paper we report an analysis of the 3231-ntp sequence at one end of the 8.1-kb linear molecule of *H. attenuata.*

Materials and Methods

H. attenuata was obtained from Hans Bode, University of California, Irvine, maintained in the medium of Loomis and Lenhoff (1956) and fed on freshly hatched brine shrimp (*Artemia salina*). MtDNA was obtained and purified as described previously (Warrior and Gall 1985) from a clone of *H. attenuata* derived from a single organism. Following incubation of purified mtDNA (two linear molecules, each of about 8 kb) with T4 DNA polymerase and all four dNTPs to ensure flush ends, ^{32}P-end-labeled (using [γ-^{32}P] ATP and T4 polynucleotide kinase) *Eco*RI linkers were added using T4 DNA ligase. Excess linkers were removed by sepharose CL-4B column chromatography. The resulting mtDNA molecules were cleaved with *Eco*RI (using conditions suggested by the manufacturer) and the products were ligated into *Eco*RI-digested plasmid pUC19 and amplified using as host *E. coli* JM101, then recloned into bacteriophage M13mp19. Other details regarding electrophoresis, cloning, and purification of single-stranded M13 DNAs for sequencing are given or referred to in Pont-Kingdon et al. (1994).

The entire sequence of the *Eco*RI fragment in one M13 plasmid (pRW-E3.2), determined by restriction analysis to comprise the end segment of the 8.1-kb linear mtDNA molecule, was obtained by employing three different strategies. The extension di-deoxyribonucleotide termination procedure (Sanger et al. 1977) was used to determine the sequence of a segment of 3016 ntp that includes one end of the *Eco*RI fragment from pRW-E3.2 (ntp 125–3240, Fig. 1A). This sequence was obtained from sets of deletion clones (Dale et al. 1985) containing overlapping sequences representing the entire sequence of both complementary strands of the 3016-ntp segment. As the 3016-ntp sequence lacked an *Eco*RI site at one end, an attempt was made to obtain an extension of this sequence as follows. A polymerase chain reaction (PCR) was carried out using natural *H. attenuata* mtDNA and a single oligonucleotide [318-1; 5′ CATATTCAATCGCCTTTACTTTC] which is complementary to nt 468–489 (within the l-rRNA gene, Fig. 1A), and *Taq* DNA polymerase (Bethesda Research Laboratories), for 50 cycles of the following regiment: 1 min at 94°C, 1 min at 60°C, 3 min at 72°C. The 3′ ends of the single-stranded product were either poly (A)-tailed, or poly (C)-tailed, using terminal deoxynucleotidyl transferase, and complementary strands were synthesized using either the *Bam*HI, *Bgl*II, *Pst*I restriction site (RS)–containing primer TB 17-1 (5′ CCAGATCTGGATCCTGCAGTTTTTTTTTTTTTTT; Beagley et al. 1996) for poly (A)-tailed strands, or the *Eco*RI, *Bam*HI restriction site (RS)–containing primer TB17-4 (5′ CTGCGAATTCGGATCCTGTGGGGGGGG) for poly (C)-tailed strands, for 5 cycles (1 min at 94°C, 1 min at 37°C, 3 min at 72°C). In each case the double-stranded product was PCR amplified for 30 cycles (1 min at 94°C, 1 min at 60°C, 3 min at 72°C) using oligonucleotide 318-1 and either the RS-containing primer TB17-2 (5′ CCAGATCTGGATCCTGCAG; Beagley et al. 1996) or the RS-containing primer TB17-5 (5′ CTGCGAATTCGGATCCTGT). The resulting products were sequenced in separate reactions using two oligonucleotides that are complementary to nt 192–214 (Fig. 1A) (318-2: 5′ ATAATTACTAAGATGTTTCAGTTC) and to nt 435–454 (Fig. 1A) (720-4; 5′ CATCATACTTTTAATGAATTACC).

The sequence of the end 190 ntp of the linear molecule was determined using the chemical cleavage method of Maxam and Gilbert (1980). The *H. attenuatta* mtDNA *Eco*RI fragment was purified from pRW-E3.2, and 3′ end labeled by a fill-in reaction using the Klenow fragment and [α-^{32}P] dATP. The labeled fragment was cleaved at an internal *Hind*III site (nt-939, Fig. 1A) and the resulting two *Eco*RI–*Hind*III fragments were separated by gel electrophoresis. The smaller fragment (934 ntp), 3′ end labeled within the *Eco*RI site was purified from the gel, subjected to Maxam and Gilbert chemical cleavage, and the products were fractionated on 8% and 15% denaturing acrylamide gels. Sequences were analyzed using Wisconsin GCG programs and Lasergene software from DNASTAR. Secondary structure potentials of DNA sequences upstream from the $tRNA^{f\text{-}Met}$ gene were analyzed by the methods of Zucker et al. (1999) (*www.ibc.wustl.edu/~zuker/dna/form1.cgi* [based on parameters in SantaLucia 1998]). The nucleotide sequence of the 3231-ntp *H. attenuata* mtDNA has been submitted to Genbank under the accession number AF100773.

A.

Transcription

```
GAATTCCGGGGGATAGTGCGGCCGGGGCCCCCTGCGATAGCCGCAGCAGGGGGTGCACCAGTGCGGCGCCTGACTGGGGGGGGGGGGGGGGGGGGGGGGGGGATCGGGTGATCTTC  120
TTTTAGGAGAGAAGAAAACCGAAAAGTTTTTTTGATTCCCTTGTTTTTAATGGGTTGAATGGTTTATAAAAGGTAATTCATTAAAAGTATGATGACTAGGAAAATTTTTAAGAATTTTAA  240
                  tRNA f-Met                                                                                  1-rRNA  5'end
TTATTTGATCTTCTAAAAGAAGAAGAGTAAACTAATGGAAAGTTATTTGGCTCATGACCAGAATATAAGGGTTCGATTCCTTTCTCTTTATTTTTGAATTAAATTTTTAAATTATAAAAA  360
TAGAATAATCAATAAGAGTAATATTAATTAATATTATTATAATCTTTTCATAAATGATTAAATTTAAAATGAACTGAAACATCTTAGTAATTATAAAATAATATTATGAAAGTAAAGGCG  480
ATTGAAATCATAATTTTTTCTTTGGATTTACAAAATTAAATAATTGATTATATCTATAAAATTAGTACTGTGAAGGAAATTAATAAATAATTTATAATAAAAAATGAAAATCATTTACCT  600
TTTGTATAATGGGCTTATAAAATAATAATTAGCAAGCTATAAGGCTAAATAATTTATAAAGTTAATTATTCCCGAAACTTAACGATCTAATTGTGAATAATTTTCAGAATGAACCAATAA  720
ATGATAAAAATTTTTGGATAATTTGCAATTGGCAGTGAAAAACTAATCGAGTTAAGAGATAGCTGGTTTTTCATGAAATTTATATAAGTAATTTTTTTTAAAAAAAAGCAGTTATTTAAT  840
GCAAAGATTATTTAACGAGTAAAAAACAATTTATATCCAACATCAATGAACTCTATAAACATTTAGAAATAAGTTTATGATTTAAACAAAAAGTAAATAAGCTTAGAATCAGCTATATTT  960
TAAAAATTACGTAAAACTTTACTTTAAAAAAATTATAACTTTACAATATAGTTAATCATTATCATTTTTGTGGAAAATAGTAATGAAATTTCTTTTAATATATTATGAAAATTTTAAAAG 1080
AGTAATAATATAAGCATGAGTAAAATTAAATTTATCCTATTTTTTTCTTAATTTAATAAATTGATTAAGGTAATACAATTTCTCAATTTATAAGGAATATATTAAAAATCTTTTCAAGAA 1200
AATTTTAATTTATATTTTGTAAAAATCTAAATCAAGTTAGGATAAATAATAAATTAATTTTTTTTTAAGGAACTCGGCAAAATAAAATATCGACTGTTTACCAAAAACATAGCCCTCTAA 1320
AATATTGAAGGTGAAACCTGCCCAATGATATAAAAATAATTATAACCTAATTTAAATAATATTAAATGGATGCAGTAACTCTGACTGTAATAAAGTAGCATAATTAGTTGTCATTTAATT 1440
GATGGAGAGAATGAATGGTTACACGAATTTTTCACTGTCTTAAAAAAAATTTTTTAAAATTGAAATAATAGTTAAGATGCTATTTAAAATTGTAAGACGAAGAGACCCTATAGAGCTTTA 1560
CTATAAACTTTTTTCTTAAAAATAATATAAAATTATTAAAACTAGAAAGTTTGGTAGTTTAGTTGGGGCGACTGTTTTTTAAAAATAACAAAAATAAGCAATATAATAAATTTATTTATT 1680
GTATAATTAAACAATTTAACAATTACTATAGTAGGCTATAATGACCCGTTATTATAATAAATAATAAATAACGATTAATTGATAAAAGCTACCTTAGGGATAACAGGATAATTTTGTTTT 1800
AGAGTCCTTATCAAAAACAAAGTTTGTCACCTCTATGTTGAATTAAGATATCCTAATAATGCAGTAGTTATTGAGGGTAGGTCTGTTCGACCTTTAAAATCTTACATGATTTGAGTTCAT 1920
TCCGTCGCAAGACAGGAAGGTTTCTATCTACAATTATAAAATTACTTTAGACTTTTAGTACGAAAGGAATAAAGTTTCACTTAGTTAATTAAAAAAACTAATAAAAAAATTTTGAAAGAA 2040
      l-rRNA  3'end  tRNA Trp                                                                              COII
TCTAAAAATTTAAAGTAATTATTTAATAAAGAGGGTTGATTAAGTAGATCAAATAGTCTTCAAAATTATTAGTATTGGTTCGAATCCAATACCCTCTGTTTATGAAAAATTTAATATTTC 2160
                                                                                                    M  K  N  L  I  F  L
TAATAAGTTTTTTCACTTGTTTATTATTTAATAATTTTATCTATAAAGATATTCCTGAAATAAATCAACTTTCTTTCCAAGAAAGTGCTTCATCTTGTAGTAATAATTTAATATTTTTTC 2280
  I  S  F  F  T  C  L  L  F  N  N  F  I  Y  K  D  I  P  E  I  N  Q  L  S  F  Q  E  S  A  S  S  C  S  N  N  L  I  F  F  H
ATGATAATACAATGTTTTATATGATAATTATTCTTGTATTAGTAGGATGACTATTATTTTCAATAATTATTAATAAAAATTATAACAAATTTTTACTAGAAAACAATTTTATTGAAATTG 2400
  D  N  T  M  F  Y  M  I  I  I  L  V  L  V  G  W  L  L  F  S  I  I  I  N  K  N  Y  N  K  F  L  L  E  N  N  F  I  E  I  V
TATGAACATTAATACCTGCTATTATATTAATAATTATAGCTATTCCTTCTCTATTTTTATTATATTCTATTGATGAAAATGTTGAACCTTCTTTAACTGTAAAAATAGTTGGTCATCAAT 2520
  W  T  L  I  P  A  I  I  L  I  I  I  A  I  P  S  L  F  L  L  Y  S  I  D  E  N  V  E  P  S  L  T  V  K  I  V  G  H  Q  W
GATATTGATCTTATGAATATTCTAATTTTTCACCAAATATAGAATTTGATTCTTATATGATTCCTACTACAGATTTAAATCAAGGTGATTTTAGATTACTTGAAACTGATAATGATCTTG 2640
  Y  W  S  Y  E  Y  S  N  F  S  P  N  I  E  F  D  S  Y  M  I  P  T  T  D  L  N  Q  G  D  F  R  L  L  E  T  D  N  D  L  V
TTTTACCAGTTAATACAAAAATTAAATTAATTGTTTCAAGTGCTGATGTTATTCATTGTTGAACAGTACCTTCTTTAGGCGTAAAAGTAGATGCTATTCCAGGTCGTTTAAATCAATTAA 2760
  L  P  V  N  T  K  I  K  L  I  V  S  S  A  D  V  I  H  C  W  T  V  P  S  L  G  V  K  V  D  A  I  P  G  R  L  N  Q  L  N
ATTTTATTATTAATAGACCTGGTAAATTTTTTTGGTCAATGTTCAGAATTATGTGGTCTTAATCATTCTTTTATGCCTATTTCAATTTATTCCGTTTCTCAAGAAAAATTTATAAACTGAT 2880
  F  I  I  N  R  P  G  K  F  F  G  Q  C  S  E  L  C  G  L  N  H  S  F  M  P  I  S  I  Y  S  V  S  Q  E  K  F  I  N  W  S
          ATPase8
CTATTAATAAAACTTAAAATGTCTCAATTAGATTTATCTTTATATTTAAATCACTATTTTGTTTTACTATTATTATTTTTTTTTATTAATAATTTTAATAAGTTTAAGATTTTATAATTTT 3000
  I  N  K  T  ter M  S  Q  L  D  L  S  L  Y  L  N  H  Y  F  V  L  L  L  L  F  F  L  L  I  I  L  I  S  L  R  F  Y  N  F
                                                                                                    ATPase6
TTCTTAGTTTTAAACATAAGAAATAATTTTATTTCAAATAGTTCAGAATCAATTAATAATGATTTTAAACAATTTTCTATATTATATAATATATTGAAATTATAATGTCTTCCTATTTCG 3120
F  L  V  L  N  I  R  N  N  F  I  S  N  S  S  E  S  I  N  N  D  F  K  Q  F  S  I  L  Y  N  I  L  K  L  ter
                                                                                                    M  S  S  Y  F  D
ATCATTTTTTAATTATTAAAATTTTTGGGTTCTTATTTGATTCTCATTTGATATTGATATTATCTATATTTATAATTTTTGTAGCTTTAAATTTTACTTTTATAATCCCTTCAAGAATTC 3240
  H  F  L  I  I  K  I  F  G  F  L  F  D  S  H  L  I  L  I  L  S  I  F  I  I  F  V  A  L  N  F  T  F  I  I  P  S  R  I
```

B.

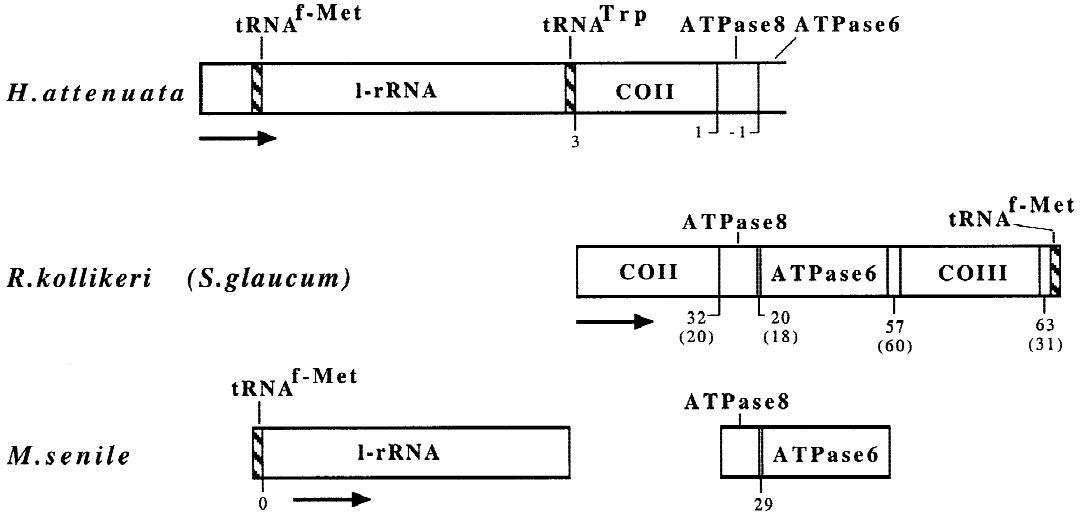

Fig. 1. **A** Nucleotide sequence of a segment of one of the 8-kb linear mtDNA molecules of *Hydra attenuata*. The first 9 ntp (underlined) of the 3240-ntp sequence shown originate from an added *Eco*RI site–containing linker. The remainder of the sequence contains complete genes for COII (cytochrome *c* oxidase subunit II), ATPase8 (ATPase subunit 8), l-rRNA (large ribosomal subunit RNA), $tRNA^{f\text{-}Met}$ and $tRNA^{Trp}$, and the 5′ terminal 136 ntp of the ATPase6 gene. Transcription of all genes is from left to right, as indicated. The predicted amino acid sequence of each of the protein genes is shown below the sequence and termination codons are indicated by ter. TGA codons are shown to specify tryptophan. tRNA genes are overlined and the anticodons are indicated by brackets. **B** Summary gene map of the 3240-ntp *H. attenuata* mtDNA sequence shown in **A,** and conserved gene arrangements relative to *H. attenuata* mtDNA, found in the mtDNAs of *Renilla kollikeri* and *Sarcophyton glaucum* (Anthozoa, Octocorallia), and *Metridium senile* (Anthozoa, Hexacorallia) (Beagley et al. 1995, 1998; Beaton et al. 1998). The number of apparently noncoding nucleotides between adjacent genes are shown (numbers of *S. glaucun* noncoding nucleotides are in parentheses). The single nucleotide pair overlap between the *H. attenuata* ATPase8 and ATPase6 genes is indicated by -1. The precise 5′ and 3′ termini of the *H. attenuata* mt-l-rRNA gene have not been determined.

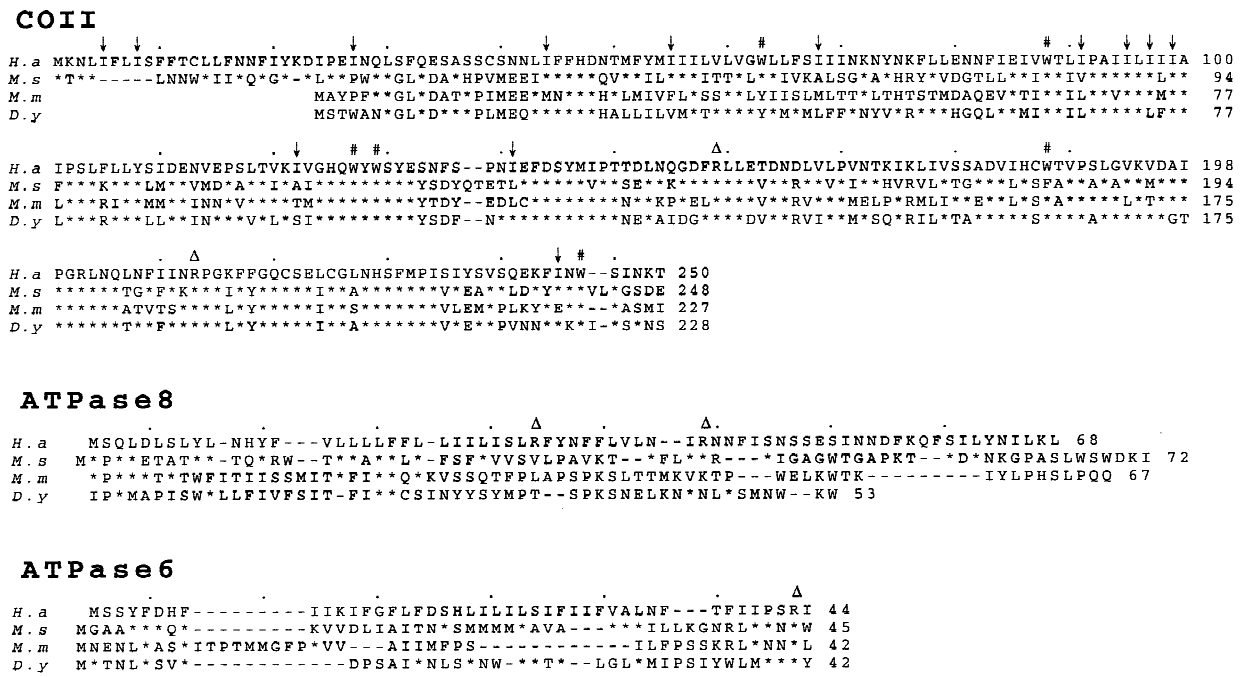

Fig. 2. Comparisons of the predicted amino acid sequences of the two completely sequenced mt-protein genes, COII and ATPase8, and the partially sequenced ATPase6 gene (see gene designations in the Fig. 1 legend) of *H. attenuata* mtDNA with the amino acid sequences of the corresponding proteins of *Metridium senile* (*M.s.*, Beagley et al. 1998), *Drosophila yakuba* (*D.y.*, Clary and Wolstenholme 1985), and mouse (*M.m.*, Bibb et al. 1981). All amino acid sequences are inferred from nucleotide sequences. The amino acids identical to those of *H. attenuata* are indicated by aterisks. A dash indicates the absence in one sequence of an amino acid that occurs in one or more of the other, corresponding sequences. Genetic code modifications used for translation of mtDNA nucleotide sequences are as follows. TGA specifies tryptophan in all species; ATA specifies methionine in *D. yakuba* and mouse sequences, but isoleucene in the *H. attenuata* and *M. senile* sequences. AGA specifies serine in the *D. yakuba* sequence (AGA and AGG do not specify an amino acid in mouse mt-protein genes, and AGG does not specify an amino acid in *D. yakuba* mt-protein genes [Bibb et al. 1981; Clary and Wolstenholme 1985]), but arginine in the *H. atenuata* and *M. senile* sequences. Above the *H. attenuata* sequences # identifies tryptophans all specified by TGA codons, Δ identifies arginines specified by AGA codons, and ↓ identifies isoleucines specified by ATA codons.

Results and Discussion

The sequence shown in Fig. 1A is a molecule end containing 3240-ntp *EcoRI* fragment of the 8.1-kb mtDNA molecule of *H. attenuatta* (including the added 9 ntp *EcoRI* site–containing linker, ntp 1–9, underlined, Fig. 1A). Nucleotides 225–3240 were obtained by Sanger sequencing of deletion clones. As this 3016-ntp sequence lacked an *Eco*RI site at one end, an attempt was made to determine additional sequence from PCR-generated segments of the fragment end lacking an *Eco*RI site. An additional 100 ntp (105–224 ntp) were obtained. However, Southern blotting of ^{32}P-end labeled oligonucleotide 720-4 (complementary to ntp 192–214, Fig. 1A) to electrophoretically separated natural *H. attenuatta* mtDNA digested with *Sca*I (which cleaves at nt 544 [AGTACT], Fig. 1A) revealed a band of about 525 ntp (data not shown). This is about 100 ntp larger than that expected from the sequences obtained by the Sanger method, suggesting that extension synthesis had been blocked. Therefore, to determine the remaining end-proximal region of the molecule we relied on Maxam and Gilbert chemical sequencing reactions of the relevant segment of the cloned pRW-E3.2 fragment. The sequence obtained (195 ntp: ntp 1–195; Fig. 1A) included a terminal *Eco*RI site and overlapped and added 104 ntp to the previously determined sequence. Successful addition of an *Eco*RI linker to an end of a natural *H. attenuata* mtDNA molecule by the T4 DNA polymerase–T4 DNA ligase procedure employed, indicates that the ends of the complementary strands of the molecule are not linked to each other and is consistent with them terminating with a 3′ hydroxyl and a 5′ phosphate.

From nucleotide and predicted amino acid sequence comparisons to mtDNAs of *Metridium senile*, *Drosophila yakuba*, and mouse (Fig. 2) (Bibb et al. 1981; Clary and Wolstenholme 1985; Beagley et al. 1998), it was determined that the *H. attenuata* mtDNA sequence contains, in the order given, complete genes for tRNA$^{f\text{-}Met}$, l-rRNA, tRNA^{Trp}, and COII. All of these genes would be transcribed in the same direction, left to right, as shown in Fig. 1A. The open reading frame following the COII gene is interpreted as the ATPase8 gene based on a minimal similarity of the predicted amino acid sequence to that of the *M. senile* ATPase8 gene (Fig. 2) and on hydropathic profile similarities to those of other metazoan ATPase8 genes (data not shown). The 136 ntp following the *H. attenuata* ATPase8 gene are

Table 1. Comparisons of predicted amino acid sequences of protein genes encoded in mtDNAs of *Hydra attenuata, Metridium senile,* mouse, and *Drosophila yakuba*

	Number of amino acids				Percentage amino acid sequence similarity[a,b]		
	H. attenuata	*M. senile*[a]	mouse[a]	*D. yakuba*[a]	*H. attenuata/M. senile*	*H. attenuata*/mouse	*H. attenuata/D. yakuba*
COII	250	248	227	228	52.8	44.0	51.8
ATPase8	68	72	67	53	21.5	11.4	9.3

[a] Data for *M. senile,* mouse, and *D. yakuba* are from Pont-Kingdon et al. (1994) and Beagley et al. (1998); Bibb et al. (1981); and Clary and Wolstenholme (1985), respectively

[b] These values are calculated from Fig. 2. the values are the minimum similarities that include all insertion/deletions

tentatively identified as the 3′ end proximal region of the ATPase6 gene, based on the occurrence of three short amino acid sequence motifs found in corresponding regions of the *M. senile* ATPase6 gene (Fig. 2). The ATPase8 and ATPase6 genes would be transcribed in the same direction as the other, upstream genes (Fig. 1A).

It remains unclear as to which portion of the 260-ntp sequence upstream from the $tRNA^{f\text{-}Met}$ gene might be concerned with transcription initiation and as to whether a region of the sequence contains a replication origin. A notable feature in this region is the run of 31 Gs (ntp 76–106, Fig. 1A), which may have blocked the extension synthesis sequencing reactions. Also, the DNA sequence upstream from the run of Gs, the sequence between the run of Gs and the $tRNA^{f\text{-}Met}$ gene, and the entire 260-ntp sequence have considerable stable secondary structure potential (ΔG values are −13.3, −12.3, and −27.0 kcal/mol, respectively). It is of interest to note that within an intergenic region of mtDNA of the sea urchin *Strongylocentrotus purpuratus,* there occurs a run of 20 Gs immediately upstream from the sequence that acts as template for synthesis of a displacement loop (D-loop) (Jacobs et al. 1989). However, we could not detect further sequence similarities upstream or downstream from the runs of Gs in the *H. attenuata* and *S. purpuratus* mtDNA molecules. Batch sequence comparisons of the 260-ntp *H. attenuata* sequence revealed a high degree of similarity between an A+T-rich 31-ntp sequence (ntp 220–247; Fig. 1A) and sequences found in gene and nongene regions of genomic DNAs of organism that include *Saccharomyces cerevisiae, Caenorhabditis elegans,* and *Homo sapiens.* However, none of these latter sequences have been identified as, or are presumed to be transcription or replication initiation sequences. Further, short tandem repeated sequences that are characteristic of the telomeres of eukaryotic linear chromosomal DNA molecules (Henderson 1995) and the linear mtDNAs of the protozoan ciliates *Tetrahymena* and *Paramecium* (Morin and Cech 1988a, 1988b; Cummings 1992) are not apparent in the end regions upstream from the *H. attenuata* mt-$tRNA^{f\text{-}Met}$ gene.

The relative arrangement of the genes in the *H. attenuata* mtDNA molecule has some similarities to mtDNA molecules of both the octocorallian anthozoans *Renilla kollikeri* and *Sarcophyton glaucum* and the hexacorallian anthozoan *M. senile* (Beagley et al. 1995, 1998; Beaton et al. 1998). In *R. kollikeri* and *S. glaucum* mtDNA the COII-ATPase8-ATPase6 gene order is conserved in a segment of the molecule that also includes the COIII gene followed by a $tRNA^{f\text{-}Met}$ gene (Fig. 1B), which is transcribed in the opposite direction to all other genes (Beagley et al. 1995; Beaton et al. 1998). In *M. senile* mtDNA the ATPase8-ATPase6 gene order is also found, a situation common to mtDNAs of vertebrates, echinoderms, arthropods, and the polyplacophoran mollusc *K. tunicata* (Wolstenholme 1992a; Boore and Brown 1994). However, although the overlap of the 3′ end of the ATPase8 gene and the 5′ end of the ATPase6 gene seen in *H. attenuata* mtDNA is also found in vertebrates, echinoderms, and arthropods, the two genes are separated by 20, 18, and 29 ntp in *R. kollikeri, S. glaucum,* and *M. senile* mtDNAs, respectively (Fig. 1B). The *H. attenuata* gene order $tRNA^{f\text{-}Met}$-l-rRNA is conserved among cnidarians only in *M. senile* mtDNA (Fig. 1B).

The *H. attenuata* COII gene is 23 and 22 codons larger than the COII genes of *D. yakuba* and mouse, respectively, but only two codons larger than the *M. senile* COII gene (Fig. 2, Table 1). The major size difference between these COII genes occurs at their 5′ ends. The *H. attenuata* ATPase8 gene is of a size (68 codons) similar to those of *M. senile* (72 codons) and mouse (67 codons), but considerably larger than that of *D. yakuba* (54 codons) (Fig. 2, Table 1).

Codon Usage and the Genetic Code

Codon usage among the complete COII and ATPase8 genes and partial ATPase6 gene is summarized in Table 2. Although all 20 amino acids are collectively encoded by the three *H. attenuata* mt-protein genes, only 38 of the expected 62 amino acid–specifying codons (assuming that TGA specifies tryptophan, see below) are used. The 24 codons not used have a bias reflecting the low average G (9.2%) and C (10.4%) content of the sense strands of the COII, ATPase8, and ATPase6 protein genes and the low average frequencies with which codons of these genes end in G (3.0%) or C (4.1%): of the 8 codons comprising exclusively G and/or C, only 2 are used; of the 24 codons in which 2 of the nucleotides are G and/or

Table 2. Codon usage in the COII and ATPase8 genes and partial ATPase6 gene of *Hydra attenuata*

Phe	TTT	34	Ser	TCT	18	Tyr	TAT	14	Cys	TGT	5
	TTC	5		TCC	2		TAC	0		TGC	0
Leu	TTA	39		TCA	10	Ter	TAA	2	Trp	TGA	6
	TTG	3		TCG	0		TAG	0		TGG	0
Leu	CTT	5	Pro	CCT	9	His	CAT	6	Arg	CGT	1
	CTC	0		CCC	0		CAC	1		CGC	0
	CTA	5		CCA	3	Gln	CAA	9		CGA	0
	CTG	0		CCG	0		CAG	0		CGG	0
Ile	ATT	29	Thr	ACT	6	Asn	AAT	30	Ser	AGT	6
	ATC	2		ACC	0		AAC	4		AGC	0
	ATA	23		ACA	5	Lys	AAA	14	Arg	AGA	5
Met	ATG	7		ACG	0		AAG	0		AGG	0
Val	GTT	9	Ala	GCT	6	Asp	GAT	14	Gly	GGT	6
	GTC	0		GCC	0		GAC	0		GGC	1
	GTA	8		GCA	0	Glu	GAA	12		GGA	1
	GTG	0		GCG	0		GAG	0		GGG	1

The total number of occurrences are given for each codon. The only deviation from the standard genetic code appears to be the use of TGA codons to specify tryptophan

C, only 8 are used. In the COII gene the only codon ending in G is ATG, the single codon used to specify methionine (see below). The average G + C content (19.6%) of *H. attenuata* mt-protein genes is much less than the G + C contents of mt-protein genes of other Cnidaria: *M. senile*, 37.5%; *S. glaucum*, 35.5% (Beagley et al. 1998; Beaton et al. 1998). As is the case for all other cnidarian protein genes so far reported, the three *H. attenuata* mt-protein genes all begin with an ATG codon.

In the mt-genetic codes of the hexacorallian anthozoan *M. senile* and the octocorallian anthozoan *S. glaucum* the use of TGA to specify tryptophan is the only modification found, relative to the standard genetic code (Pont-Kingdon et al. 1994, 1998; Beagley et al. 1998). This contrasts with the mt-genetic codes of other Metazoa in which four or five codon-specificity modifications occur (Wolstenholme 1992a, 1992b; Wolstenholme and Fauron 1995). The most common of these involve the codons ATA, AGA, and AGG. In most invertebrate and all vertebrate mtDNAs ATA specifies methionine rather than isoleucine. In invertebrate mtDNAs (other than those of Cnidaria) AGA and AGG specify serine except that in urochordates they specify glycine (Yokobori et al. 1993). However, in vertebrate mtDNAs these codons do not specify an amino acid, although in rare cases each may act as a translation termination codon (see Wolstenholme 1992a).

Among the two complete and one partial *H. attenuata* mt-protein gene sequences obtained, six TGA codons occur, all in the COII gene. The interpretation that these TGA codons specify tryptophan is supported by the observation that there is a five out of six correspondence in position of the *H. attenuata* COII TGA codons to codons specifying tryptophan in each of the *M. senile*, mouse, and *D. yakuba* COII genes (Fig. 2). Also, *H. attenuata* mtDNA encodes a tRNA with a 5′ UCA anticodon (Figs. 1A and 3) expected to recognize 5′ UGA, and 5′ UGG codons, although TGG codons do not occur in the *H. attenuata* mt-protein genes so far sequenced (Fig. 1A). A greater use of TGA than TGG codons to specify tryptophan has been noted for all non-cnidarian mt-protein genes and correlates with a universal occurrence among these mtDNAs of more codons ending in A than in G. However, in contrast, in both *M. senile* and *S. glaucum* mtDNAs although there is again a greater overall frequency of codons ending in A than in G (1.7:1 and 2.2:1, respectively) the frequency of TGG codons is in considerable excess to TGA codons (1.6:1 in *M. senile* and 30.0:1 in *S. glaucum*) (Pont-Kingdon et al. 1994, 1998; Beagley et al. 1998; Beaton et al. 1998).

A total of five AGA codons (but no AGG codons) occur in the *H. attenuata* COII gene (2), ATPase8 gene (2), and partial ATPase6 gene (1). Three of these AGA codons (two in the COII gene and one in the ATPase6 gene) are similarly located to arginine-specifying codons in all of the corresponding *M. senile*, mouse, and *D. yakuba* mt-protein gene sequences. Also, a fourth AGA codon in the *H. attenuata* ATPase8 gene is similarly located to an arginine-specifying codon in the *M. senile* ATPase8 gene. These observations strongly support the interpretation that at least AGA has the standard genetic code specification of arginine in the *H. attenuata* mt-genetic code, as is the case in the other two cnidarians examined to date.

From the following observations, ATA also appears to have the standard genetic code specificity of isoleucine in *H. attenuata* mtDNA. Of the 13 ATA codons that occur in the *H. attenuata* COII gene, six, four, and five correspond in position to isoleucine-specifying codons in the *M. senile*, mouse, and *D. yakuba* COII genes, respectively. Among the *M. senile*, mouse, and *D. yakuba* COII genes only two locations (both in the mouse COII gene)

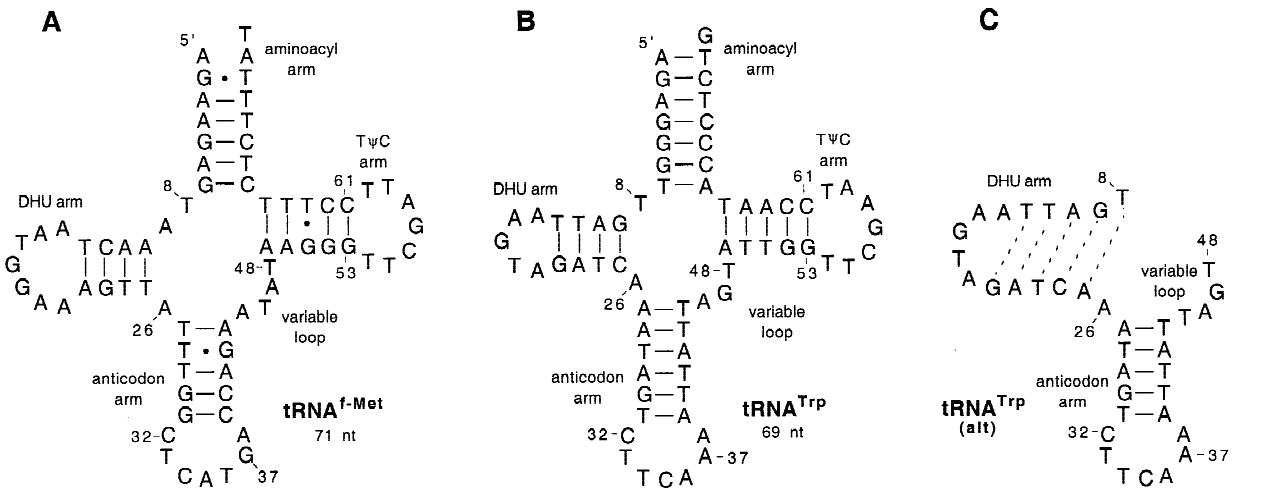

Fig. 3. Potential secondary structures of the genes for tRNA$^{f\text{-}Met}$ (**A**) and tRNA^{Trp} (**B** and **C,** alternative forms) encoded in *H. attenuata* mtDNA. The nucleotide numbers follow the numbering system used for yeast tRNA^{Phe} (Sprinzl et al. 1989). As the *H. attenuata* tRNA^{Trp} cannot be completely folded into the usual secondary structure, two alternative, nonorthodox secondary structural forms are shown for this tRNA gene. In **C,** dashed lines indicate possible base pairing between complementary sequences of five nucleotides.

that correspond to ATA codons in the *H. attenuata* COII gene are occupied by methionine-specifying codons (Fig. 2). Further, of the five ATG codons in the *H. attenuata* COII gene, four, three, and three correspond to methionine-specifying codons in the *M. senile*, mouse, and *D. yakuba* COII genes, respectively. None of the *H. attenuata* ATG codons correspond to isoleucine-specifying codons in the *M. senile*, mouse, and *D. yakuba* COII genes.

The above findings add further strength to the conclusion drawn from our *M. senile* and *S. glaucum* studies (Pont-Kingdon et al. 1994, 1998; Beagley et al. 1998) that the tryptophan specificity of TGA predates the emergence of Metazoa, but that the switch in specificities of AGR codons from arginine to serine and ATA codons from isoleucine to methionine postdated divergence of the cnidarian line from the ancestral lines of all other extant Metazoa.

The H. attenuata *mt-tRNA$^{f\text{-}Met}$ and mt-tRNA^{Trp} Genes*

The *H. attenuata* mt-tRNA$^{f\text{-}Met}$ gene (Fig. 3A) predicts a tRNA that is identical in size (71 nt) and secondary structure potential to the mt-tRNA$^{f\text{-}Met}$ of both *M. senile* and *S. glaucum* (Pont-Kingdon et al. 1994, 1998). Overall sequence similarities of the mt-tRNA$^{f\text{-}Met}$ gene of *H. attenuata* to those of *M. senile* and *S. glaucum* are 66.2% and 63.4%, respectively. Also as reported for the two previously described cnidarian mt-tRNA$^{f\text{-}Met}$ genes, the *H. attenuata* mt-tRNA$^{f\text{-}Met}$ gene contains two major features of standard tRNAs (prokaryotic tRNAs and tRNAs encoded in eukaryotic nuclear and chloroplast DNAs) that are absent in most metazoan mt-tRNAs: both G_{18} and G_{19} in the DHU loop, and the complete 5′ T_{54}, T_{55}, C_{56}, Pu_{57}, A_{58}, N_{59}, Py_{60} TψC loop. Further, almost all of the other nucleotides that are highly conserved (T_8, Pu_9, A_{14}, Pu_{21}, Py_{32}, T_{33}, Pu_{37}, Py_{48}, G_{53}, C_{61}) or semi-conserved (Pu_{10}, Pu_{13}, Pu_{15}, Py_{25}, Py_{27}, and Pu_{43}) among standard tRNAs (Rich and RajBhandary 1976; Dirheimer et al. 1979; Singhal and Fallis 1979; Sprinzl et al. 1989) are found in the *H. attenuata* mt-tRNA$^{f\text{-}Met}$ gene (Fig. 3A). Also as in the other cnidarian mt-tRNA$^{f\text{-}Met}$ genes, the *H. attenuata* mt-tRNA$^{f\text{-}Met}$ gene has an A_{11}-T_{24} pair (rather than a Py_{11}-Pu_{24} pair found in almost all standard tRNA genes) in the DHU stem, and a mismatched pair (AA) at the top of the aminoacyl stem (C A and T C in the *M. senile* and *S. glaucum* tRNA$^{f\text{-}Met}$ genes, respectively), two features found in the *E. coli* initiator tRNA$^{f\text{-}Met}$. The three cnidarian mt-tRNA$^{f\text{-}Met}$ genes so far described lack one (*H. attenuata* and *S. glaucum*) or two (*M. senile*) of the three G-C pairs found at the base of the anticodon stem of prokaryotic initiator tRNA$^{f\text{-}Met}$s and eukaryotic nuclear DNA-encoded initator tRNA^{Met}s, that in *E. coli* have been shown to be important for targeting tRNA$^{f\text{-}Met}$ to the ribosome P site (RajBhandari and Chow 1995). All three of these G-C pairs are present in mt-tRNA$^{f\text{-}Met}$ genes of vertebrates, insects, most echinoderms, an annelid, a bivalve mollusc, and some gastropod molluscs. However, in nematodes only two anticodon-proximal G-C pairs occur, in *Asterina pectinifera* (echinoderm) the central pair is G·T, and in both *K. tunicata* (polyplacophoran mollusc) and *Cepaea normalis* (gastropod mollusc) the anticodon-distal pair is a G A mismatch (Sprinzl et al. 1998).

In the *H. attenuata* mtDNA sequence (Fig. 1A), a 69-ntp segment located between the 3′ end of the l-rRNA gene and the COII gene has most of the primary and potential secondary structural characteristics expected

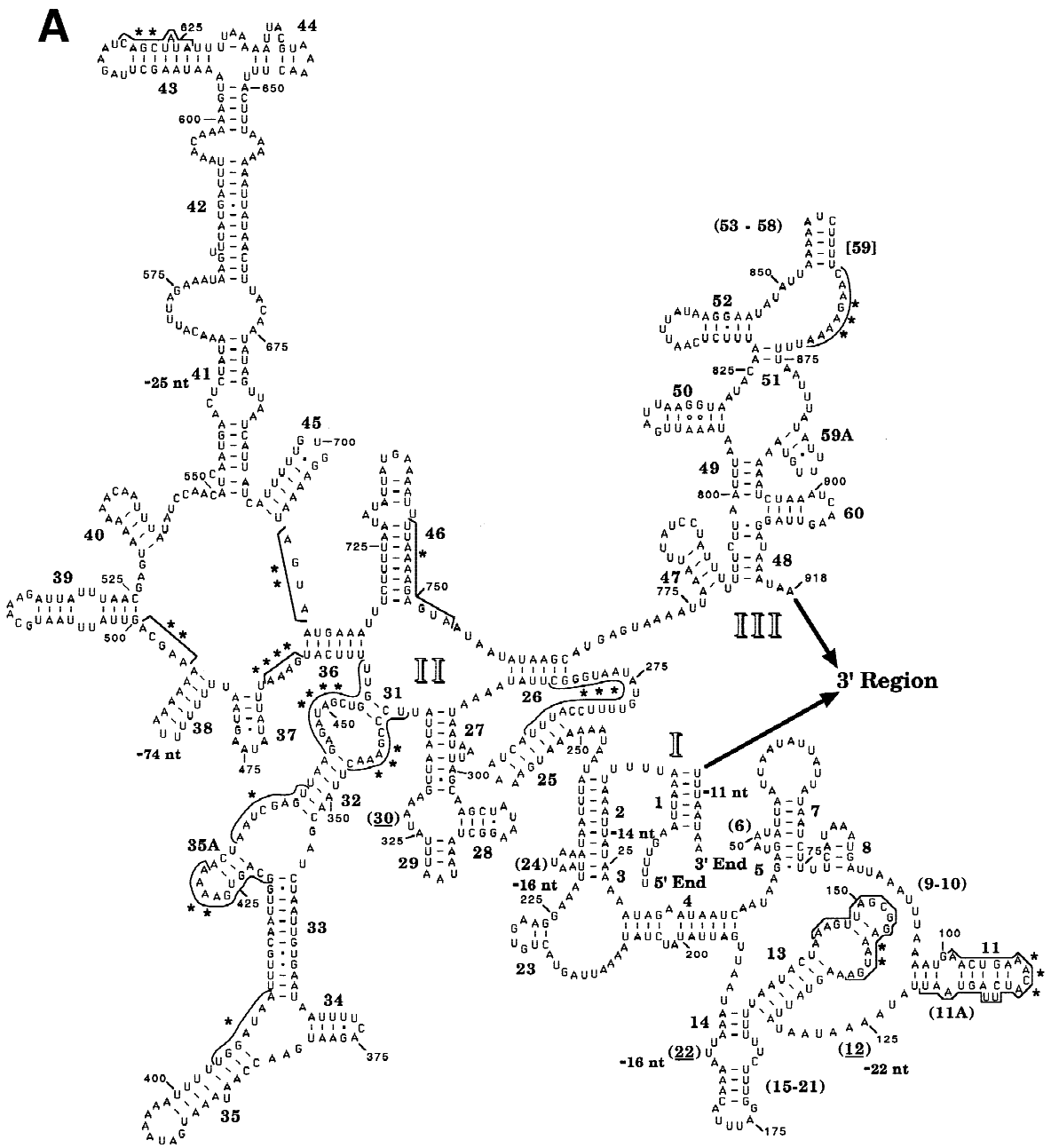

Fig. 4. Secondary structure model of the *H. attenuata* mt-l-rRNA. The 5′ and 3′ regions are contained in **A** and **B,** respectively. The 5′ and 3′ ends are defined as corresponding to the nucleotides that follow the 3′ terminal nucleotide of the $tRNA^{f\text{-}Met}$ gene and preceed the 5′ terminal nucleotide of the $tRNA^{Trp}$ gene (Fig. 1A). The sequence is numbered every 25 nt from the 5′ end. Watson-Crick base pairings are indicated by a dash between nucleotides. Pairing of G and U is indicated by G·U and presumed pairing of G and A is indicated by G○A. The solid outlines indicate sequences that, in the *Metridium senile* mt-l-rRNA and *Escherichia coli* 23S rRNA secondary structure models are at identical locations and are 100% conserved (Gutell et al. 1993; Beagley et al. 1998). The similarities of these latter sequences to *H. attenuata* mt-l-rRNA are indicated: **** 100%; *** 86–91%; ** 70–80%; * 46–63%. Roman numerals identify six major domains corresponding to the six domains of the *E. coli* 23S secondary structure model (Gutell et al. 1993). Bold numbers identify helices that appear to have been conserved relative to the *E. coli* model (1–101, as defined by Brimacombe et al. 1990). Bold numbers in parentheses indicate the equivalent locations of *E. coli* 23S rRNA helices that clearly cannot be formed from the *H. attenuata* sequence. Those numbers in parenthesis that are underlined indicate helices that are absent in *H. attenuata* but present in *M. senile.* The large bold negative numbers shown indicate differences of more than 10 nucleotides in either one or more helical elements in the *H. attenuata* mt-l-rRNA model relative to the *M. senile* mt-l-rRNA model.

for a $tRNA^{Trp}$ gene (anticodon 5′ TCA). This sequence is one nucleotide shorter than and has 57.8% similarity to the *M. senile* mt-$tRNA^{Trp}$ gene (Beagley et al. 1998). As is the case in the *H. attenuata* mt-$tRNA^{f\text{-}Met}$ gene, the mt-$tRNA^{Trp}$ gene includes a complete TψC loop sequence (T_{54}, T_{55}, C_{56}, Pu_{57}, A_{58}, N_{59}, Py_{60}) and other nucleotides that are highly conserved in standard tRNAs (T_8, Py_{32}, T_{33}, Pu_{37}, Py_{48}, G_{53}, C_{61}). However, in the

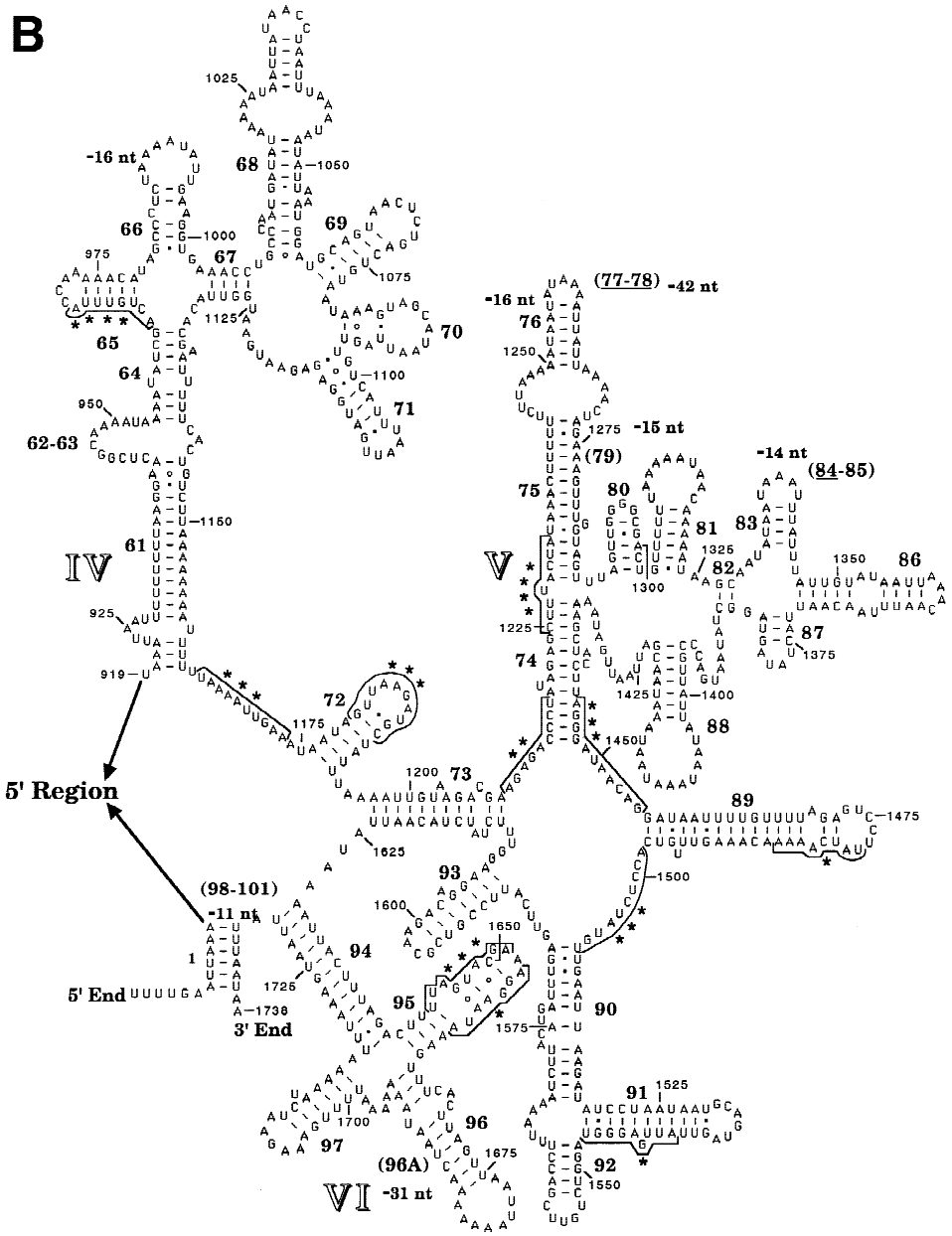

Fig. 4. Continued.

predicted *H. attenuata* mt-tRNA^{Trp} gene sequence only one G is found in the location expected for the highly conserved G_{18}-G_{19}, and the sequence cannot be folded into a totally orthodox mt-tRNA secondary structure. There are, in fact, two ways in which this sequence can be folded, as diagrammed in Figs. 3B and 3C. In the secondary structure shown in Fig. 3B, in contrast to what is usually found in mt-tRNAs, there is one nt (a T) rather than 2 nt (TA) between the amino-acyl stem and the DHU stem, 6 ntp rather than 5 ntp in the anticodon stem, and 3 nt rather than 4 nt or 5 nt in the variable loop. This is similar to the secondary structure proposed for mt-tRNA^{Ser} (UCN) genes of mammals (Yokogawa et al. 1991; Watanabe et al. 1994). In the alternative secondary structure shown in Fig. 3C, a DHU stem is formed by pairing between two complementary sets of five nucleotides with elimination of both nucleotides usually found between the aminoacyl and DHU stems. Interestingly, the mt-tRNA$^{f\text{-}Met}$ gene of five species (*pyriformis, thermophila, pigmentosa, hyperangularis,* and *hegewischi*) of the ciliated protozoan *Tetrahymena* also contain a DHU arm sequence in which three consecutive nucleotide pairs could form with elimination of nucleotides between the aminoacyl and DHU stems (Suyama et al. 1987; Morin and Cech 1988a).

DHU arm sequence modifications have been described previously for metazoan mt-tRNAs that recognize serine-specifying codons. All known mtDNA-encoded tRNA^{Ser} (AGN) genes of invertebrates, tRNA^{Ser} (AGR) genes of vertebrates, and tRNA^{Ser} (UCN) genes

of nematodes, some molluscs, and the annelid worm *Lumbricus terrestris*, have a DHU arm replacement loop of between 6 and 13 nt that either lacks or has unusual internal base pairing potential (Wolstenholme 1992a; Boore and Brown 1994, 1995; Terrett et al. 1996; Yamazaki et al. 1997; Beagley et al. 1999). A similar secondary structure has been proposed for other mt-tRNAs of the gastropod *Albinaria coerulea* (Hatzoglou et al. 1995). Genes for other tRNAs with different modifications of their secondary structural form have been found among metazoan mtDNAs, including genes for 20 of the 22 mt-tRNAs of nematodes in which the TψC arm is replaced with a single loop of between 6 and 12 nt that lacks intra-strand pairing potential (Wolstenholme et al. 1987; Okimoto et al. 1992). From computer modeling studies further unusual secondary structural features have been suggested for a variety of metazoan mt-tRNAs (Steinberg et al. 1997).

As is the case for all other metazoan mt-tRNA genes the trinucleotide CCA, that occurs at the 3′ end of prokaryotic tRNA genes is absent from this position in each of the two *H. attenuata* mt-tRNA genes.

The H. attenuata *mt-l-rRNA Gene*

If the entire 1738-ntp sequence between the $tRNA^{f\text{-}Met}$ and $tRNA^{Trp}$ genes (Fig. 1A) encodes the mt-l-rRNA, then this mt-l-rRNA is 451 nt shorter than that of *M. senile*, the largest metazoan mt-l-rRNA so far reported (Beagley et al. 1998).

A secondary structure model for the *H. attenuata* mt-l-rRNA based on the secondary structure model of *E. coli* 23S rRNA (Brimacombe et al. 1990; Gutell et al. 1993) is shown in Fig. 4. This model, in accordance with modeling of the *E. coli* 23S rRNA, is divided into 5′ and 3′ regions (A and B, Fig. 4) that collectively comprise helical elements numbered as in Brimacombe et al. (1990). All of the 25 helical elements absent from the *M. senile* mt-l-rRNA model (Beagley et al. 1998) are also absent from the *H. attenuata* mt-l-rRNA model. Relative to the *M. senile* mt-l-rRNA model there is loss of a further six helical elements (12, 22, 30, 77, 78, 84) in the *H. attenuata* model. Also, neither of the two helical elements present in the *M. senile* model but not in the *E. coli* 23S model (11A and 96A) are present in the *H. attenuata* model.

The largest single nucleotide difference between the *H. attenuata* and *M. senile* mt-l-rRNA secondary structure models results from a severe shortening of helical element 38 by 74 nt. From other distinct locations (helical elements and loops) in the *H. attenuata* model there are 13 losses (totaling 323 nt) of between 11 and 42 nt (shown in Fig. 4) and 32 losses (totaling 121 nt) of between one and 10 nt. Also, there are 13 locations at which between one and 5 nt (totaling 31 nt) have been added. The sum of the nucleotides predicted to be lost or gained at specific locations is 423 which accounts for 91.6% of the difference in size of 451 nt between the *H. attenuata* and *M. senile* mt-l-rRNAs.

As is the case for the secondary structure models proposed for other metazoan mt-l-rRNAs (Gutell et al. 1993; Okimoto et al. 1994), including that of *M. senile* (Beagley et al. 1998), the secondary structure model for *H. attenuata* shown in Fig. 4 includes all of the core nucleotide elements given for the *E. coli* 23S rRNA by Cedergren et al. (1988). This observation, together with conservation of the overall secondary structure of the *H. attenuata* mt-l-rRNA relative to the ml-l-rRNA of *M. senile* and the 23S rRNA of *E. coli* mt-l-rRNAs, argue in favor of the *H. attenuata* mt-l-rRNA being functional in regard to mitochondrial protein synthesis.

Acknowledgments. We are grateful to Dr. Hans Bode for providing us with *Hydra attenuata.* The *H. attenuata* mitochondrial DNA fragment used in this work was cloned and partially sequenced in the laboratory of Dr. J.G. Gall. RW would like to gratefully acknowledge Dr. Gall's support. The remainder of the work reported was supported by National Institutes of Health grant number GM18375.

References

Asakawa S, Himeno H, Miura K-I, Watanabe K (1995) Nucleotide sequence and gene organization of the starfish *Asterina pectinifera* mitochondrial genome. Genetics 140:1047–1060

Beagley CT, Macfarlane JL, Pont-Kingdon GA, Okimoto R, Okada NA, Wolstenholme DR (1995) In: Palmieri F, Papa S, Saccone C, Gadaleta N (eds) Mitochondrial genomes of Anthozoa (Cnidaria). Progress in Cell Research Symposium on Thirty Years of Progress in Mitochondrial Bioenergetics and Molecular Biology. Amsterdam: Elsevier Science BV, pp 149–153

Beagley CT, Okada NA, Wolstenholme DR (1996) Two mitochondrial group I introns in a metazoan, the sea anemone *Metridium senile*: one intron contains genes for subunits 1 and 3 of NADH dehydrogenase. Proc Natl Acad Sci USA 93:5619–5623

Beagley CT, Okimoto R, Wolstenholme DR (1998) The mitochondrial genome of the sea anemone *Metridium senile* (Cnidaria): introns, a paucity of tRNA genes, and a near standard genetic code. Genetics 148:1091–1108

Beagley CT, Okimoto R, Wolstenholme DR (1999) Mytilus mitochondrial DNA contains a functional gene for a $tRNA^{Ser}$ (UCN) with a dihydrouridine arm-replacement loop and a pseudo-$tRNA^{Ser}$ (UCN) gene. Genetics 152:641–652

Beaton MJ, Roger AJ, Cavalier-Smith T (1998) Sequence analysis of the mitochondrial genome of *Sarcophyton glaucum:* conserved gene order among octocorals. J Mol Evol 47:697–708

Bibb MJ, Van Etten RA, Wright CT, Walberg MW, Clayton DA (1981) Sequence and gene organization of mouse mitochondrial DNA. Cell 26:167–180

Boore JL (1999) Animal mitochondrial genomes. Nucleic Acids Res 27:1767–1780

Boore JL, Brown WM (1994) Complete DNA sequence of the mitochondrial genome of the black chiton *Katharina tunicata.* Genetics 138:423–443

Boore JL, Brown WM (1995) Complete sequence of the mitochondrial DNA of the annelid worm *Lumbricus terrestris.* Genetics 141:305–319

Bridge D, Cuningham CW, Schierwater B, Desalle R, Buss LW (1992) Class level relationships in the phylum Cnidaria: evidence from

mitochondrial genome structure. Proc Natl Acad Sci USA 89:8750–8753

Brimacombe R, Greuer B, Mitchell P, Obwald M, Rinkeappel J, et al. (1990) Three dimensional structure and function of *Escherichia coli* 16S and 23S rRNA as studied by crosslinking techniques. In: Hill WE, Dahlberg A, Garrett RA, Moore PB, Schlissinger D, Warner JR (eds) The ribosome: structure, function and evolution. Washington, DC: ASM Press, pp 93–106

Cedergren R, Gray MW, Abel Y, Sederoff D (1988) The evolutionary relationships among known life forms. J Mol Evol 28:98–112

Clary DO, Wolstenholme DR (1985) The mitochondrial DNA molecule of *Drosophila yakuba*: nucleotide sequence, gene organization, and genetic code. J Mol Evol 22:252–271

Clayton DA (1992) Transcription and replication of animal mitochondrial DNAs. In: Wolstenholme DR, Jeon KW (eds) Mitochondrial genomes. International Reveiw of Cytology, vol. 141. New York, NY: Academic Press, pp 217–232

Cummings DJ (1992) Mitochondrial genomes of ciliates. In: Wolstenholme DR, Jeon KW (eds) Mitochondrial genomes. International Review of Cytology, vol. 141. New York, NY: Academic Press, pp 1–64

Dale RMK, McClure B, Houghins JP (1985) A rapid single-stranded cloning strategy for producing overlapping clones for use in DNA sequencing: application to sequencing the corn mitochondrial 18S rDNA. Plasmid 13:31–40

Dirheimer G, Keith G, Sibler A-P, Martin RP (1979) The primary structure of tRNAs and their rare nucleosides. In: Schimmel PR, Soll D, Abelson JN (eds) Transfer RNA: structure, properties and recognition. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory, pp 19–41

Gutell RR, Gray MW, Schnare MN (1993) A compilation of large subunit (23S and 23S-like) ribosomal RNA structures. Nucleic Acids Res 21:3055–3074

Hatzoglou E, Rodakis GC, Lecanidou R (1995) Complete sequence and gene organization of the mitochondrial genome of the land snail *Albinaria coerulea.* Genetics 140:1353–1366

Henderson E (1995) Telomere DNA structure. In: Blackburn EH, Greider CW (eds) Telomeres. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, pp 11–34

Jacobs HT, Herberts ER, Rankine J (1989) Sea urchin mitochondrial DNA contains a short displacement loop (D-loop) in the replication origin region. Nucleic Acids Res 17:8949–8965

Krettek A, Gullberg A, Arnason U (1995) Sequence analysis of the complete mitochondrial DNA molecule of the hedgehog, *Erinaceus europaeus,* and the phylogenetic position of the Lipotyphla. J Mol Evol 441:952–957

Loomis WR, Lenhoff H (1956) Growth and sexual differentiation of *Hydra* in mass culture. J Exp Zool 132:555–574

Maxam AM, Gilbert W (1980) Sequencing end labeled DNA with base-specific chemical cleavages. Meth Enzymol 65:499–560

Morin GB, Cech TR (1988a) Phylogenetic relationships and altered genome structures among *Tetrahymena* mitochondrial DNAs. Nucleic Acids Res 16:327–346

Morin GB, Cech TR (1988b) Mitochondrial telomeres: suprising diversity of repeated telomere DNA sequences among six species of *Tetrahymena.* Cell 52:367–374

Okimoto R, Macfarlane JL, Clary DO, Wolstenholme DR (1992) The mitochondrial genomes of two nematodes, *Caenorhabdidis elegans* and *Ascaris suum.* Genetics 130:471–498

Okimoto R, Macfarlane JL, Wolstenholme DR (1994) The mitochondrial ribosomal RNA genes of the nematodes, *Caenorhabditis elegans* and *Ascaris suum*: consensus secondary structure models and conserved nucleotide sets for phylogenetic analysis. J Mol Evol 39:598–613

Pont-Kingdon GA, Beagley CT, Okimoto R, Wolstenholme DR (1994) Mitochondrial DNA of the sea anemone, *Metridium senile* (Cnidaria): prokaryote-like genes for $tRNA^{f\text{-}Met}$ and small subunit ribosomal RNA, and standard genetic code specificities for AGR and ATA codons. J Mol Evol 39:387–399

Pont-Kingdon GA, Okada NA, Macfarlane JL, Beagley CT, Wolstenholme DR, Cavalier-Smith T, Clark-Walker GD (1995) A coral mitochondrial MutS gene. Nature 375:109–111

Pont-Kingdon GA, Okada NA, Macfarlane JL, Beagley CT, Watkins-Sims CD, Cavalier-Smith T, Clark-Walker GD, Wolstenholme DR (1998) Mitochondrial DNA of the coral, *Sarcophyton glaucum* contains a gene for a homologoue of bacterial MutS: a possible case of gene transfer from the nucleus to the mitochondria. J Mol Evol 46:419–431

RajBhandary UL, Chow CM (1995) Initiator tRNAs and initiation of protein synthesis. In: Söll D, Raj Bhandary UL (eds) tRNA Structure, biosynthesis and function. American Society for Microbiology, pp 511–528

Rich A, RajBhandary UL (1976) Transfer RNA: molecular structure, sequence and properties. Ann Rev Biochem 45:805–860

Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA 74:5463–5467

SantaLucia J (1998) A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. Proc Natl Acad Sci USA 95:1460–1465

Shadel GS, Clayton DA (1997) Mitochondrial DNA maintenance in vertebrates. Annu Rev Biochem 66:409–435

Singhal RP, Fallis PAM (1979) Structure, function, and evolution of transfer RNAs. Prog Nucleic Acids Res Mol Biol 23:227–290

Sprinzl M, Hartmann T, Weber J, Blank J, Zeidler R (1989) Compilation of tRNA sequences and sequences of tRNA genes. Nucleic Acids Res 17(suppl):r1–r72

Sprinzl M, Horn C, Brown M, Ioudovitch A, Steinberg S (1998) Compilation of tRNA sequences and sequences of tRNA genes. Nucleic Acids Res 26(L):148–153.

Steinberg S, Leclerc F, Cedergren R (1997) Structural rules and conformational compensation in the tRNA L-form. J Mol Biol 266: 269–282

Suyama Y, Jenney F, Okawa N (1987) Two transfer RNA sequences abut the large ribosomal RNA gene in *Tetrahymena* mitochondrial DNA: $tRNA^{leu}$ (anticodon UAA) and $tRNA^{met}$ (anticodon CAU). Curr Genet 11:327–330

Terrett JA, Miles S, Thomas RH (1996) Complete DNA sequence of the mitochondrial genome of *Cepaea nemoralis* (Gastropoda: Pulmonata). J Mol Evol 42:160–168

Warrior R, Gall J (1985) The mitochondrial DNA of *Hydra attenuata* and *Hydra littoralis* consists of two linear molecules. Arch Sci Geneva 38:439–445

Watanabe Y, Kawai G, Yokogawa T, Hayashi N, Kumazawa Y, Ueda T, Nishikawa K, Hirao I, Miura K, Watanabe K (1994) Higher order structure of bovine mitochondrial $tRNA^{ser}$ UGA: chemical modification and computer modeling. Nucleic Acids Res 22:5378–5384

Wolstenholme DR (1992a) Animal mitochondrial DNA: structure and evolution. In: Wolstenholme DR, Jeon KW (eds) Mitochondrial genomes. International Review of Cytology, vol. 141. New York, NY: Academic Press, pp 173–216

Wolstenholme DR (1992b) Genetic novelties in mitochondrial genomes of multicellular animals. Curr Opin Genet Dev 2:918–925

Wolstenholme DR, Fauron CM-R (1995) Mitochondrial genome organization. In: Levings CS III, Vasil IK (eds) Advances in cellular and molecular biology of plants. Vol. 3: Molecular biology of the mitochondria. Dordrecht: Kluwer Academic Publishers, pp 1–59

Wolstenholme DR, Macfarlane JL, Okimoto R, Clary DO, Wahleithner JA (1987) Bizarre tRNAs inferred from DNA sequences of mitochondrial genomes of nematode worms. Proc Natl Acad Sci USA 84:1324–1328

Yamazaki N, Ueshima R, Terrett JA, Yokobori S, Kaifu M, Segawa R, Kobayashi T, Numachi K, Ueda T, Nishikawa K, Watanabe K, Thomas RH (1997) Evolution of pulmonate gastropod mitochondrial genomes: comparisons of gene organizations of Euhadra, Ce-

paea and Albinaria and implications of unusual tRNA secondary structures. Genetics 145:749–758

Yokobori S, Ueda T, Watanabe K (1993) Codons AGA and AGG are read as glycine in ascidian mitochondria. J Mol Evol 36:1–8

Yokogawa T, Watanabe Y, Kumazawa Y, Ueda T, Hirao I, Miura K, Watanabe K (1991) A novel cloverleaf structure found in mammalian mitochondrial $tRNA^{ser}$ (UCN). Nucleic Acids Res 19:6101–6105

Zardoya R, Meyer A (1997) The complete DNA sequence of the mitochondrial genome of a "living fossil", the Coelacanth (*Latimeria chalumnae*). Genetics 146:995–1010

Zucker M, Mathews DH, Turner DH (1999) Algorithms and thermodynamics for RNA secondary structure predictions: a practical guide. In: Barciszewski J, Clark BFC (eds) RNA biochemistry and biotechnology. NATO ASI Series. Dordrecht: Kluwer Academic Publishers, pp 11–43