J Mol Evol (1995) 41:1016–1021

JOURNAL OF MOLECULAR EVOLUTION

Structural Constraints in Expansion Segments from a Midge 26S rDNA

Eduardo Gorab, Mario Garcia de Lacoba, Luísa Maria Botella

Centro de Investigaciones Biológicas (C.S.I.C.), Velázquez 144, 28006 Madrid, Spain

Received: 10 March 1995 / Accepted: 31 July 1995

Abstract. DNA sequences representing approximately 40% of the large-subunit rRNA gene from the lower dipteran *Chironomus thummi* were analyzed. Once aligned with their *Drosophila* counterparts, sequence and base content comparisons were carried out. Sequence identity was found to be high overall, except for six regions that displayed a local bias in nucleotide composition toward AT. These regions were identified as expansion segments D3, D4, D5, D6, D7a, and D12. Besides base sequence divergence, differences in length were observed between the respective variable domains of the two species, particularly for D7a. Prediction of secondary structure showed that the folding of the *Chironomus* expansion segments analyzed is in agreement with the general patterns proposed for eukaryotic LSU rRNA. The comparison with *Drosophila* revealed also that the *Chironomus* secondary structures of the variable domains are supported by multiple compensatory substitutions or even compensatory insertions. *Chironomus* D7a displayed an unusual structural feature with respect to the insect D7a models that have been inferred up to now. The structural constraint observed in the expansion segments of Diptera so distantly related as midges and *Drosophila* suggests that these regions contribute to some functional role. Concerning the D7a of insects so far analyzed, there can be, in addition to a conserved secondary structure, a nucleotide composition constraint that might be important for the process giving rise to the alpha and beta halves of the 26S rRNA.

Key words: *Chironomus* — LSU rRNA evolution — Expansion segments — Compensatory mutations — rRNA processing

Correspondence to: E. Gorab

Introduction

The genes of the eukaryotic large-subunit rRNA are structured as a mosaic of conserved and variable domains (Ware et al. 1983). The "core" segments have primary and secondary structures conserved in prokaryotes and eukaryotes. The variable regions are responsible for the difference in size between eukaryotic and prokaryotic rRNA. They are named divergent (Hassouna et al. 1984) or expansion segments (Clark et al. 1984).

Despite the variation in length and sequence among the expansion segments, their position and secondary structure are rather conserved in different species (Michot et al. 1984; Michot and Bachellerie 1987; Engberg et al. 1990; Linares et al. 1991). A process entitled "compensatory slippage" (Hancock and Dover 1990) has been assigned for maintaining the secondary structure of the expansion segments, which argues for a constraint possibly related to the functionality of these regions. On the other hand, it has been suggested that expansion segments provide no contribution to ribosome function (Gerbi et al. 1985).

Although the role of these variable regions remains unknown, a particular expansion segment is involved in the rRNA processing of insects. Visualized as a gap in R-loop images (Wellauer and Dawid 1977; Renkawitz-Pohl et al. 1981; Schmidt et al. 1982), removal of a few bases gives rise to the alpha and beta halves of the 26S rRNA. This process occurs in the expansion segment D7a (Delanversin and Jacq 1983; Ware et al. 1985; Fu-

jiwara and Ishikawa 1986); rRNA motifs have been proposed as specific processing signals taking part in this event.

This report deals with the expansion segments contained in two sequences of the dipteran *Chironomus thummi thummi* 26S rDNA. Concerning the insects whose rDNA has been studied with an evolutionary focus, Chironomidae (midges) and Culicidae (mosquitoes) are among the most primitive. The divergence of higher and lower flies is supposed to be 150 million years (Woodley 1991). Expansion segments provide an opportunity to compare some structural marks in the large subunit of the rDNA with those found in chorionated insects in order to verify whether and how the evolutionary divergence in Diptera could be reflected in the divergent domains of the 26S rRNA gene.

Materials and Methods

A 1-kb rDNA fragment was recloned from an rDNA complete repeat unit from *Chironomus thummi thummi* (Schmidt et al. 1982). It was chosen since a local AT enrichment has been detected in this region after hybridization experiments (data not shown). This prompted us to analyze its nucleotide composition as no other region in the *Chironomus* rDNA was found to present AT-rich sequences, except for the IGS (Schmidt and Godwin 1983).

Hybridization, subcloning, and nucleic acid extraction procedures were done according to usual protocols (Sambrook et al. 1989). The initial subclone constructed in pUC-18 contains an *Hpa*II fragment from the 3′-end of 26S alpha rDNA, according to the map of Schmidt et al. (1982). It hybridized to the rRNA in Northern blots as well as to the nucleolar organizer region of *Chironomus* and *Drosophila* polytene nuclei. The fragmentation of *Chironomus thummi* 26S rRNA is clearly seen in denaturing Northern blots as a band around 1.8 kb hybridizes to 26S rDNA probes.

Southern blot experiments with *C. thummi* genomic DNA cut with *Hpa*II showed that the subclone used as a probe strongly hybridized to a 1.1-kb band as expected. Two new subclones derived from the initial construction were done to facilitate the sequencing by the dideoxynucleotide method (Sanger et al. 1977) using the TaqTrack kit (Promega). The sequence is available in the EMBL data library with the accession number X80912. Another sequence analyzed, by comparison with *Drosophila* data, was found to contain the 3′-end of the 26S beta rDNA from *Chironomus thummi*, besides the IGS (Schmidt and Godwin 1983). Further data on insect sequences are referred in the text.

Secondary-structure models were obtained by comparing *Drosophila* and *Saccharomyces* structures (Hancock et al. 1988; Gutell et al. 1993), using also Zuker's Fold program (Zuker and Stiegler 1981), which is included in the University of Wisconsin GCG sequence analysis software package.

Results

Nucleotide Composition in Core and Expansion Segments

The first *Chironomus thummi* sequence (1,086 nucleotides) encompasses the 3′-end of the 26S alpha rRNA gene plus 63 bases from the putative beginning of the 26S beta rDNA (Ware et al. 1985). The alignment was found starting with base 4202 of *Drosophila melanogaster* rDNA (Tautz et al. 1988) taken as a reference. The first base of the second, previously sequenced stretch (440 nucleotides), was aligned with *Drosophila* base 6691, which mapped in the 3′-end of the *Drosophila* 26S beta rDNA. Sequence similarity with *Drosophila* rDNA is 80% overall. In the core regions it is usually higher than 90% and the differences were basically due to point mutations. The base composition was very similar to *Drosophila* throughout both stretches and in the core regions as well (Table 1). Six segments, where the identity with *Drosophila* significantly decreases (38% on average; 53% maximum), show a bias in base composition toward AT base pairs. Their total length represents around 15% of the *Chironomus thummi* 26S rRNA gene (Schmidt et al. 1982) and they coincide in position with those defined as expansion segments D3, D4, D5, D6, D7a, and D12. The AT composition in the analyzed expansion segments of *Chironomus* is not consistently higher than the AT content in the regions of the core. In *Drosophila*, there is a marked difference between the AT composition of the expansion segments with respect to both core regions and the overall stretches (Table 1).

Table 1. AT content in *Chironomus thummi*[a]

AT content (%)	Overall	"Core" regions	Exp. segments
Chironomus	57 (1527)	54 (991)	62 (536)
Drosophila	60 (1441)	55 (980)	73 (461)

[a] Percentage of AT content in the two sequences of *Chironomus thummi* 26S rRNA gene analyzed (accession number X80912, positions 1–1086; positions 1–440 in the sequence published by Schmidt and Godwin 1983) and their respective *Drosophila* counterparts (positions 4202–5193; 6691–7139). In parenthesis, the lengths (in base pairs) of the core regions and the expansion segments in the sequences analyzed. *Drosophila* data are taken from Tautz et al. (1988) and Hancock et al. (1988)

Expansion Segment Primary Structure

An alignment with regions containing the expansion segments of *Chironomus* and *Drosophila* was carried out (Fig. 1). Differences in length and base sequence between the respective expansion segments were evident even though isolated bases and certain motifs, ranging from dinucleotides to octanucleotides, are common to both sequences. The *Drosophila* expansion segments are in general shorter than those of *Chironomus*. D5 and D7a of *Chironomus* are clearly larger, due to insertions of 28 and 47 base pairs, respectively, at the limits of each expansion segment (positions 387–414; 974–1020).

In some situations, the limits for core regions and expansion segments were not clear. For instance, the *Drosophila* core sequence 5′-ATATATGCTGT-3′ (positions 4251–4261) is aligned with D3 of *Chironomus* (po-

D3 *C* : 1-75; *D* : 4202-4273

```
C-.CGGGGTATTTTCATCAACA-TATGCTTGCATATGGAAAATATACCATGAGCACACACTGGATACGACCCGAAAGAT..
D-.TAATGTATATTTATATTATTTATGCCTCTAACTGGAACG--TACCTTGAGCATATATGCTGT--GACCCGAAAGAT..
```

D4 *C* : 229-273; *D* : 4428-4472

```
C-.TAGCTGGAGCATAGAAAAGTTGTATTGCTAACTCATACCTGGTAA...
D-.TAGCTGGTGCATTTTAATATTATATAAAATAATCTTATCTGGTAA...
```

D5 *C* : 332-428; *D* : 4530-4600

```
C-.ATGGGTATGTAAGTCAACATGCTTGAT-TGATGTTGACCATTAGCTTACGTTCTTA*CAGTGGCCAAGTTT...
D-.ATGGGTAAGAACCTTAACTTTCTTGATATGAAGATCAAGGTTATGATATAAATGTCCCAGTGGGCCACTTT...
     *: TGAACGTAGCACGTAAGATTGTATCGCG
```

D6 *C* : 616-703; *D* : 4786-4869

```
C-.TGGCGCTCAAGTCGTTTGCCGATACATGTCGTTAAGATAAAATCAGTGTGTGTTCATTGTTGGACAT*ATTTTGAAA
D-.TGGCGCTTAAGTTGTATACCTATACATTACCGCTAAAGTAGATGATTTATATTAC-TTGTGATATAA ATTTTGAAA

C- TCTTAACGA...
D- CTTTAGTGA...
     *: TTC
```

D7a *C* : 888-1031; *D* : 5035-5138

```
C-.TCCTAAGCTCTAGGGGAAACCTGAAACGATGAGTAGATATAAAACTATATTTTTCAGTTTTGTGATGTATCAAAAATTCT
D-.TCCTAAGTTCAAGGCGAAAGCGAAAATTTTCAAGTAAAACAAAAATGCCTAACTATATAAACAAAGCGAATTATAATACA

C- TTATCA*GAGCGAAAGGG...
D- CTTGAATGAACGAAAGGG...
     *: TGTTGCGAAGTGTTGTTGCGTTCACGCAGTGCACTTCCTCTAGCAAA
```

D12 *C* : 216-440; *D* : 6909-7139

```
C-.AACGCCTCTAAGGTCGTAGCCGTGCCAGAACAACAGTTCATCTAAAGGCTCTTATACACATTGTATTGGTTATAAT
D-.AACGCCTCTAAGGTCGTATCCGTGC-TGGACTGCAATGATAAATAAGGGGC-AATTTGCATTGTATGGCTTCTAAA

C-.GGACGGGCACATATCTATTCTCGCAATATTTTTGTACTCACCTAAATCCATTCTCATATTCTCCCATGTGTCT-GTACGTA
D-.CCATTTAAAGTTTATAATTTACTTTATAAACGACAATGGATGTGATGCCAATGTAATTTGTAACATAGTAAATTGGGAGGA

C-.TCATATCGA-AGCTAGTAATTACGATTATGTTTTTCGAACGC-TCCGATATTG--TTTACA--GTC---AATCAATGG.
D-.TCTT--CGATCACCTGATGCCGCGCTAGTTACATATAAAAGCATTATTTAATACAATGACAAAGCCTAGAATCAATTG.
```

Fig. 1. Sequences within and flanking the expansion segments of *C. thummi thummi* (*C*) and *D. melanogaster* (*D*). *Numbering* refers to the positions of the respective stretches. *D3* is partially represented as the entire expansion segment of *Chironomus* was not subcloned. Bases and motifs *underlined* match the *Drosophila* sequence. *Chironomus* insertions represented with *asterisks* did not match the expansion segments of *Drosophila*. *Bolded* motifs of both rDNA sequences point to the limits of the expansion segments. Criteria for fixing the limits are similarity between the flanking regions in both rDNA stretches and size of *Drosophila* expansion segments. The alignment was done manually and the gaps (-) were inserted to allow maximum base matching. *Drosophila* data are taken from Tautz et al. (1988) and Hancock et al. (1988). *Chironomus* D12 sequence and numbering were taken from Schmidt and Godwin (1983).

sitions 51–61) as part of the expansion segment; the *Drosophila* motif 5′-ATTTTGAAA-3′ (positions 4852–4860) mapped in the expansion segment D6 (Hancock and Dover 1988; Hancock et al. 1988) has a perfect counterpart in *Chironomus* (positions 686–694), which makes doubtful its inclusion as a core or expansion segment. The limits also became less clear when the sequences of *Chironomus, Drosophila,* and another lower dipteran, *Aedes* (Kjer et al. 1994), were simultaneously compared. Concerning the latter species, a low number of base matches was observed in a tentative alignment with the respective expansion segments of *Chironomus.* One could at least expect a better matching than that found for *Drosophila* and *Chironomus,* as midges and mosquitoes belong to the same subgroup (Culicoidea). Instead, considerable differences in nucleotide sequence and expansion segments length as well were seen between both lower dipterans (data not shown).

Expansion Segment Secondary Structure

To examine the folding pattern of the variable segments, prediction of secondary structure for D4, D5, D6, D7a and D12 was performed. *Drosophila* and *Saccharomyces* rRNA secondary structures were used as models (Hancock et al. 1988; Gutell et al. 1993). While there was divergence concerning their length and base composi-

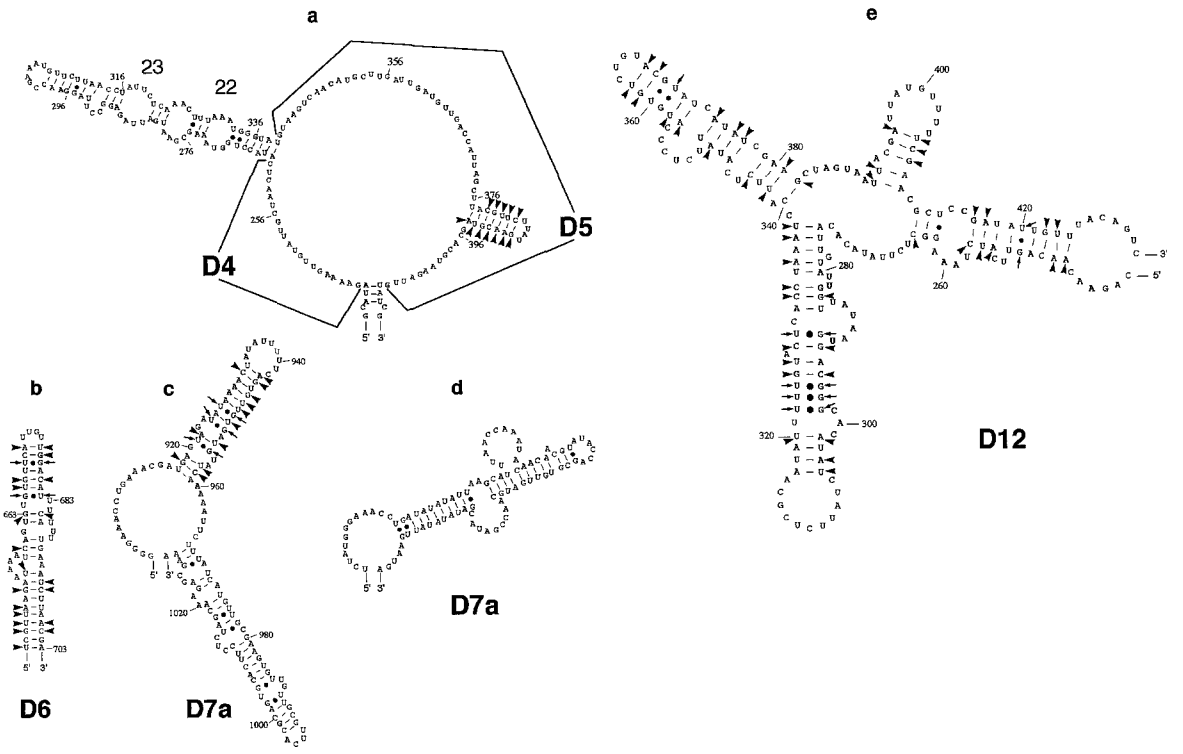

Fig. 2. Secondary structure models for the expansion segments *D4* and *D5* (*boxed*) together with core stems *22* and *23* **(a)**, *D6* **(b)**, *D7a* **(c)**, *D12* **(e)** of *Chironomus thummi*, and *D7a* of *Aedes* **(d)**. D3 was omitted as its entire sequence is not represented in the subclone. As the limits for each expansion segment are not generally clearcut, they do not exactly coincide with those imposed in Fig. 1. Every 20th nucleotide is *numbered*. Core stems 22 and 23 were *numbered* according to Michot et al. (1984). Watson-Crick base pairing is represented by a *solid line* (-); noncanonical GU base pairing by *dots*. Compensatory changes were taken from the alignment between the expansion segments of *Chironomus* and *Drosophila* (Fig. 1). Total compensatory substitutions are indicated by *arrowheads* and partial compensatory changes by *arrows*. *Aedes* D7a sequence was taken from Kjer et al. (1994) and the numbering of D12 is according to that published by Schmidt and Godwin (1983).

tion, similarities with the models were found (Fig. 2), besides a particular feature in *Chironomus* with regard to expansion segment D7a of insects.

The stem of D5 is supported by 12 total compensatory changes (Fig. 2a). An insertion motif relative to the *Drosophila* expansion segment sequence (5′-GAA-CGUA-3′ at positions 388–394) supports part of the stem. A second potential stem-loop in D5 is localized at positions 346–367. It was omitted since it was incompatible with the general structure for this divergent domain. D6 secondary structure for *Chironomus* (Fig. 2b) displays in relation to *Drosophila* 20 base changes which were found to be compensatory, besides 4 partial compensatory substitutions. Curiously, bulges and the internal loop common to the secondary structures of both dipterans also contain the same motifs (5′-UUUGAAA-3′ and 5′-UUGU-3′, respectively).

Expansion segment D7a of *C. thummi* could be split into two parts (Fig. 2c). The first one (positions 917–959) is comparable in general terms to the other insects—namely, a AT-rich region (77% AT of 43 nucleotides) encoding a potential stem-loop structure. It is supported by 12 total plus 6 partial compensatory changes. In the second part (positions 967–1027), the AT composition decreases (around 57% AT) and the RNA encoded in this region is apt to form an additional stem. Some motifs in this second segment (5′-TTGTTGCGTT-3′, 5′-CTTCCTCT-3′), by their repetition pattern, could be interpreted as examples of compensatory slippage (Hancock and Dover 1990). This novel structure, with respect to D7a of insects, is composed of an "internally compensated" insertion regarding the *Drosophila* sequence (Fig. 1) which results in the largest D7a that has been observed up to now. A comparable structure to that predicted for D7a of *Chironomus*, although with shorter stems, has been also inferred for yeast (positions 1524–1592 in the sequence with accession numbers J01355, K01048; see Gutell et al. 1993).

The secondary structure for *Aedes* D7a was also included in order to verify if the unusual two-stem pattern of *Chironomus* D7a could be present in another primitive dipteran more related to midges. The nucleotide composition in *Aedes* D7a is different from the rest of its expansion segments. It is interestingly biased to AT, which seems to be a characteristic of this variable domain in insects (table 2). Even though its base sequence (Kjer et al. 1994) could not satisfactorily match both the *Drosophila* and *Chironomus* D7a sequences, a folding pattern with one stem was also observed for this expansion segment (Fig. 2d).

Table 2. AT content in base pairs[a]

Genus	D3	D4	D5	D6	D7a	D12
Chironomus	60 (63)	67 (21)	62 (68)	70 (60)	62 (131)	64 (191)
Aedes	35 (101)	53 (19)	45 (88)	30 (33)	67 (93)	50 (182)
Drosophila	71 (160)	100 (19)	73 (41)	76 (53)	75 (92)	68 (182)
Bombyx	–	–	–	62 (68)	69 (77)	–
Sciara	–	–	–	–	72 (70)	–

[a] Percentage of AT content and respective lengths in base pairs (in parenthesis) of the expansion segments from insects named in the table. D3 values for *Chironomus* correspond only to the 3′-end of the expansion segment (see Fig. 1). Data are taken from Kjer et al. (1994) (*Aedes*); Hancock and Dover (1988), Hancock et al. (1988) (*Drosophila*); Fujiwara and Ishikawa (1986) (*Bombyx*); Ware et al. (1985) (*Sciara*); Schmidt and Godwin (1983) (*Chironomus* D12)

Chironomus D12 is the largest expansion segment analyzed in the present work. Its folding pattern resembles the *Drosophila* D12, while some variation in its overall structure was found especially with respect to D12 from such distantly related organisms as yeast and mouse (Michot et al. 1984). *Chironomus* D12 stems (Fig. 2e) are supported by 45 total compensatory mutations plus 11 partial compensatory substitutions.

Discussion

Our observations, although restricted to a limited number of expansion segments, showed that their base sequence, are considerably divergent in *Chironomus* and *Drosophila,* examples of lower and higher dipterans, respectively. On the other hand, bias in nucleotide composition toward AT in the variable domains of *Drosophila* rDNA also occurs in *Chironomus* and it differs from that found in the expansion segments of vertebrates, which are GC-rich (Hancock and Dover 1988). However, bias in base composition toward AT in expansion segments cannot be taken as a feature of dipterans. An example is given by *Aedes,* whose base content in its 26S rDNA is biased toward GC, particularly in most of the variable domains (Kjer et al. 1994).

It is worth noting that the expansion segments of *Drosophila* are generally more AT-rich than those of lower Diptera such as *Chironomus* and *Aedes* (Tables 2 and 3). The expansion segments D2 and D10 of *Glossina* (tsetse fly) are also AT-rich differently from D2 and D10 of *Aedes* (40% of 622 base pairs). In addition, we were able to localize the *C. thummi* D12 previously sequenced, whose AT content is 64%; AT composition of *Aedes* D12 is 50% of 182 nucleotides. These values are lower than those found for *Drosophila* D12 (Hancock and Dover 1988; Hancock et al. 1988). Although the 26S rDNA of *Chironomus* has not been completely sequenced, Southern blot results using an AT-rich repetitive element as a probe (Schmidt 1984) suggest that other variable regions of *Chironomus* 26S rDNA do not have AT content higher than the known divergent domains of chorionated dipterans such as *Drosophila* and *Glossina.* The data argue for an increase of AT in the 26S rDNA expansion segments of higher Diptera. This hypothesis, however, has to be reinforced with more data on the nucleotide composition in expansion segments of dipterans, which are lacking so far.

Variation in nucleotide sequence and length between expansion segments from distantly related species as *Drosophila* and *Chironomus* is expected. On the other hand, a relatively high percentage of base matches was observed for D2 and D10 of two higher dipterans that belong to different families as *Drosophila* and *Glossina* (>82% of 580 nucleotides; Linares et al. 1991). This might reflect a closer relationship for both representatives of the Schizophora subgroup. The same did not occur in the case of the two lower dipterans *Chironomus* and *Aedes,* at least for the expansion segments studied in the present work. This suggests that, in the Nematocera suborder, families such as Chironomidae and Culicidae have considerably diverged in relation to Glossinidae and Drosophilidae (for a systematics of Diptera, see Woodley 1991). The folding pattern observed for *Aedes* D7a, indicating that the two-stem model for this variable region is not conserved among the primitive dipterans, might constitute another sign of genetic distance between midges and mosquitoes.

Despite the sequence divergence between the expansion segments of *Drosophila* and *Chironomus,* prediction of secondary structure revealed certain similarities for both dipterans. In addition, they are in good agreement with those proposed as general ones for eukaryotic LSU rRNA (Michot et al. 1984). The structural constraints implied by these observations suggest a functional role for these regions. This could be especially important for D7a, where the cleavage of insect 26S rRNA takes place. In the case of insects, there seems to be not only a secondary structure constraint, but also a nucleotide composition constraint for D7a. A remarkable example supporting this assumption lies in the *Aedes* LSU rRNA, whose base composition is clearly biased toward GC in the expansion segments, except for D7a (Kjer et al. 1994). It is intriguing in this species how this variable

Table 3. AT content from dipterans[a]

Genus	D1	D2	D8	D9	D10	D11
Aedes	46 (110)	38 (345)	23 (154)	37 (27)	41 (287)	43 (14)
Drosophila	64 (148)	75 (341)	68 (215)	63 (38)	71 (214)	63 (8)
Glossina	–	75 (348)	–	–	70 (222)	–

[a] Percentage of AT content and respective lengths in base pairs (in parenthesis) of the expansion segments from dipterans named in the table. Data are taken from Kjer et al. (1994) (*Aedes*); Hancock et al. (1989) (*Drosophila*); Linares et al. (1991) (*Glossina*).

domain has maintained a conserved secondary structure and a nucleotide composition different from the rest of its 26S rRNA molecule but comparable with that of any insect D7a.

A comparison of sequences in two dipteran and one lepidopteran species has led to the possibility that the motif 5′-UAAU-3′ found in the loop of D7a acts as a signal involved in the 26S rRNA break (Fujiwara and Ishikawa 1986). While they are AU-rich overall, neither the expansion segment D7a of *C. thummi* (Fig. 1) nor the *Aedes* D7a (Kjer et al. 1994) encodes this motif. It therefore does not participate in the 26S rRNA processing of insects, as previously inferred by statistical criteria (Hancock et al. 1988).

Acknowledgments. The authors are greatly indebted to Drs. John M. Hancock and Gloria Morcillo for their interest, helpful suggestions, and critical reading of the manuscript. We also thank Drs. José Luís Díez and Gabriel A. Dover for their interest and Dr. Erwin R. Schmidt for providing us with the pCtt-1505 DNA. We are grateful to Mrs. A. Partearroyo and J. Fernández-Cabrera for technical assistance. This work was supported by grants from DGICYT, PB87-0256, Spain. E.G. is a predoctoral fellow of the National Research Council of Brazil (CNPq).

References

Clark CG, Tague BW, Ware VC, Gerbi SA (1984) *Xenopus laevis* 28S ribosomal RNA: a secondary structure model and its evolutionary and functional implications. Nucleic Acids Res 12:6197–6220

Delanversin G, Jacq B (1983) Sequence of the central break region of *Drosophila* 26S precursor ribosomal RNA. C R Acad Sci Paris 296(III):1041–1044

Engberg J, Nielsen H, Lenaers G, Murayama O, Fujitani H, Higashinakagawa T (1990) Comparison of primary and secondary 26S rRNA structures in two *Tetrahymena* species: evidence for a strong evolutionary structural constraint in expansion segments. J Mol Evol 30:514–521

Fujiwara H, Ishikawa H (1986) Molecular mechanism of introduction of the hidden break into the 28S rRNA of insects: implication based on structural studies. Nucleic Acids Res 14:6393–6401

Gerbi SA, Jeppesen C, Stebbins-Boaz B, Ares Jr M (1985) Evolution of eukaryotic rRNA: constraints imposed by RNA interactions. Cold Spring Harbor Symp Quant Biol 52:709–719

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

Hancock JM, Dover GA (1990) "Compensatory slippage" in the evolution of ribosomal RNA genes. Nucleic Acids Res 18:5949–5954

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

Hancock JM, Dover GA (1988) Molecular coevolution among cryptically simple expansion segments. Mol Biol Evol 5:377–391

Hassouna N, Michot B, Bachellerie JP (1984) The complete nucleotide sequence of mouse 28S rRNA gene. Implications for the process of size increase of the large subunit rRNA in high eukaryotes. Nucleic Acids Res 12:3563–3583

Kjer KM, Baldridge AM, Fallon AM (1994) Mosquito large subunit ribosomal RNA: simultaneous alignment of primary and secondary structure. Biochim Biophys Acta 1217:147–155

Linares AR, Hancock JM, Dover GA (1991) Secondary structure constraints on the evolution of *Drosophila* 28S ribosomal RNA expansion segments. J Mol Biol 219:381–390

Michot B, Bachellerie JP (1987) Comparisons of large subunit rRNAs reveal some eukaryotic-specific elements of secondary structure. Biochimie 69:11–23

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

Renkawitz-Pohl R, Matsumoto L, Gerbi SA (1981) Two distinct intervening sequences in different ribosomal DNA repeat subunits. Nucleic Acids Res 9:3747–3764

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory Manual. Cold Spring Harbor University Press, Cold Spring Harbor, NY

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

Schmidt ER, Godwin EA, Keyl HG, Israelewski N (1982) Cloning and analysis of ribosomal DNA of *Chironomus thummi thummi* and *Chironomus thummi piger*. Chromosoma 87:389–407

Schmidt ER, Godwin EA (1983) The nucleotide sequence of an unusual non-transcribed spacer and its ancestor in the rDNA in *Chironomus thummi*. EMBO J 2:1177–1183

Schmidt ER (1984) Clustered and interspersed repetitive DNA family of *Chironomus thummi*. J Mol Biol 178:1–15

Tautz D, Hancock JM, Webb DA, Tautz C, Dover GA (1988) Complete sequences of the rRNA genes of *Drosophila melanogaster*. Mol Biol Evol 5:366–376

Ware VC, Tague BW, Clark CG, Gourse RL, Brand RC, Gerbi SA (1983) Sequence analysis of 28S ribosomal DNA from the amphibian *Xenopus laevis*. Nucleic Acids Res 11:7795–7817

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

Wellauer PK, Dawid IB (1977) The structural organization of ribosomal DNA of *Drosophila melanogaster*. Cell 10:193–212

Woodley NE (1991) In: McAlpine JF (ed) Manual of nearctic diptera, vol 3. Biosystematics Research Centre, Ottawa, Canada

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