Xenopus laevis ribosomal protein S1a cDNA sequence

M.Di Cristina, R.Menard and P.Pierandrei-Amaldi
Istituto di Biologia Cellulare CNR, via Marx 43, 00137 Rome, Italy

Submitted February 1, 1991

EMBL accession no. X57322

In *Xenopus laevis* the coordinated synthesis of the over 80 ribosomal proteins (r-proteins) is attained by both post-transcriptional and translational level regulation (1, 2). Also several r-protein genes analyzed present structural similarities probably involved in the common regulatory mechanisms. Our understanding, however, of the structure and expression of these genes is based on the analysis of very few of them and it is thus important to extend these studies to other r-protein genes.

We have analyzed the nucleotide sequence of a full length cDNA for *Xenopus* r-protein S1 by using the Maxam and Gilbert technique. This sequence has been named S1a; in fact most genes in *Xenopus laevis* are present in two gene copies per haploid genome. The cDNA was isolated from a tadpole cDNA library (kindly provided by George Spohr) by screening with the originally cloned S1 cDNA small fragment (3) corresponding to the other copy S1b.

The deduced amino acid sequence indicates that S1a is a basic protein of 246 amino acids and has a molecular weight of 26,999. Analysis of new partial cDNA clones obtained for S1b indicates that a sequencing error in the previously published partial sequence (3) had put the code out of frame. Comparison of the full length S1a sequence with the newly determined partial S1b sequence shows that in these 144 aa (carboxy terminal sequence) only two conservative aminoacid substitutions have occurred whereas 24 differences at the nucleotide level are found. This indicates a strong evolutive pressure for protein sequence conservation. Computer search on sequence data banks has revealed that the Xenopus S1 is homologous to the rat r-protein S3 (4), with six aminoacid substitutions, mostly conservative, and a small insertion of three aminoacids close to the carboxy-terminal end of the *Xenopus* protein.

Of particular interest are the structures of the 5′ and 3′ untranslated regions (UTRs) shown in Figure 1. In fact the cDNA sequence shows that the 5′ and 3′ of r-protein S1 mRNA are both very short, 28 and 75 nt respectively, resembling the other *Xenopus* r-protein mRNAs so far characterized. Moreover the 5′ UTR starts with the run of pyrimidines, typical of all *Xenopus* r-protein mRNAs, and also of other vertebrates. It is possible that these typical structural features of the 5′ UTR and 3′ UTR are involved in the regulations common to the class of r-protein mRNAs.

ACKNOWLEDGEMENTS

Supported by 'P.F.Biotecnologie e Biostrumentazione' and 'P.F.Ingegneria Genetica', CNR.

REFERENCES

1. Pierandrei-Amaldi,P. *et al.* (1985) *Cell* **42**, 317–323.
2. Amaldi,F. *et al.* (1989) *Trends Biochem. Sci.* **14**, 175–178.
3. Amaldi,F. *et al.* (1982) *Gene* **17**, 311–316.
4. Chan,Y.L. *et al.* (1990) *Arch. Biochem. Biophys.* **283**, 546–550.

```
5'UTR

S1a     CCTTTCCTGCCAGCGGCGCTTAGCAAAGATG...

S8a     CCTCTTCTAGGCCGTCACTGAGAAAGGACGAAAGGCCATG...
S19a    CCTTTCCTTCGTCACCGTGAGAGATAGCCGGCAAGATG...
L1a     CCTTTTCTCTTCGTGGCCGCTGTGGAGAAGCAGCGAGGAGATG...
L14a    CCTTTCCTCCCCGAGAAGCAGCTGCTGCTACAGCCGCCATCATG...
L32a    CCTTTTCCTCCATCTTGGATACCAGTGCGGTCCTGCCTGCATCTGGACATTAGGAGAGCAAGATG...


3'UTR

S1a   ...TAATGGCCTGCTGAAGACCTGGATTCAAATATTTTGGATGCAGCATCAGAAAATCTAAAATAAAAAATAAAATCCCA(A)n

S8a   ...TAATGTGCAGAATTGAATAATAAAAAAAAAATACTCCTTGTTA(A)n
L1a   ...TAATCCCAGAGCGTTATCTCATGTTCAGCACTTTGGATTTACTCAATAAATTCTGTTTAATACTTA(A)n
L14a  ...TAAATTCACTCCTCATGATTTGTACACAATAAAGTTCTTTGGACTTGA(A)n
L32a  ...TAAATGTTAAGAAAAAATAAATAGAAACATTCTGG(A)n
```

Figure 1. Comparison of the 5′ UTR and 3′ UTR of *Xenopus* S1a cDNA with those of the other *Xenopus* r-proteins previously analyzed. Only one copy for each gene is shown.