The primary structures of two arginine tRNAs (anticodons C-C-U and mcm⁵a²U-C-ψ) and of glutamine tRNA (anticodon C-U-G) from bovine liver

G.Keith

Institut de Biologie Moléculaire et Cellulaire du CNRS, 15 rue Descartes, 67084 Strasbourg Cedex, France

Received 16 November 1983; Accepted 1 February 1984

ABSTRACT

The primary structure of two arginine tRNAs (anticodons : C-C-U and mcm⁵a²U-C-ψ) and of a glutamine tRNA (anticodon : C-U-G) of bovine liver have been determined using [³²P] postlabeling methods.

They are tRNA^{Arg} : pG-C-C-C-C-A-G-U-m¹G-m²G-C-C-U-A-A-D-G-G-A-D-A-A-G-G-C-A-ψ-U-G-G-C-m³C-U-C-C-U-m¹G-A-G-C-C-A-G-G-G-A-D-U-G-ψ-G-G-G-T-ψ-C-G-m¹A-G-U-C-C-A-ψ-U-C-U-G-G-G-ψ-C-C-C-A.

tRNA^{Arg} mcm⁵a²U-C-ψ : pG-G-C-U-C-C-G-U-m¹G-G-C-G-C-A-A-D-G-G-A-D-A-G-C-G-C-A-ψ-ψ-G-G-A-m³C-U-mcm⁵a²U-C-ψ-m¹G-A-U-ψ-C-A-A-A-G-m⁷G-D-U-m⁵C-C-G-G-G-T-(ψ)-ψ-C-G-m¹A-G-U-C-C-C-G-G-C-G-G-A-G-U-C-G-C-C-A

tRNA^{Gln} : pG-G-U-U-C-C-A-U-m⁴G-G-U-G-ψ-A-A-D-Gm-G-D-D-A-G-C-A-C-U-C-ψ-G-G-A-Cm-U-C-U-G-A-A-ψ-C-C-A-G-C-G-A-U-m⁵C-C-G-A-G-ψ-ψ-C-A-m¹A-A-U-C-U-C-G-G-U-G-G-A-A-C-C-U-C-C-A.

INTRODUCTION

To study the mechanisms of post-transcriptional modification of the nucleotides in tRNAs, the processing of the precursor tRNA transcripts or the anticodon-codon recognition between tRNA and mRNAs, it is important to know the primary structures of the mature tRNAs. We studied therefore the primary structure of several bovine liver tRNAs.

Two mammalian arginine tRNAs have been sequenced : tRNA^{Arg}_{ICG} from mouse leukemia (1) and tRNA^{Arg}_{CCG} from bovine liver. These tRNAs are expected to recognize the codons of the quartet : C-G-A, C-G-U, C-G-C (for tRNA^{Arg}_{ICG}) and C-G-G (for tRNA^{Arg}_{CCG}). No other sequence of mammalian arginine tRNA is yet known. We describe here the nucleotide sequence of two arginine tRNAs which would be expected to recognize the two remaining arginine codons : U-C-U and U-C-C, which contain only purine nucleotides (doublet). In addition we describe the primary structure of glutamine tRNA^{Gln}_{CUG} which would be expected to recognize the C-A-G codon but not the second glutamine codon C-A-A.

MATERIALS AND METHODS

The two bovine liver arginine tRNAs and the glutamine tRNA were purified by a three steps procedure involving two column chromatography separations : the first being on BD-cellulose (this step was performed by M. Fournier in Bordeaux (3)) and the second on Sepharose 4B (4). The third separation was an electrophoresis performed on two-dimensional polyacrylamide gel slab (4,5). The tRNAs were extracted from the gel (5) and the nucleotide sequences were determined using [^{32}P] postlabeling methods.

To determine the 5' and 3' end sequences (6), the tRNAs were [^{32}P] postlabeled by T_4 polynucleotide kinase and γ [^{32}P] ATP at their 5' end or by T_4 RNA ligase and 5' [^{32}P] pCp at their 3' end (6). After separation of the [^{32}P] labeled tRNAs and of fragments on polyacrylamide gel and extraction from the gel matrix the labeled tRNAs were partially hydrolysed and the end sequences were determined by two dimensional electrophoresis-homochromatography on DEAE-cellulose thin layers (4).

To determine the internal sequences and the minor nucleotides, the tRNAs were partially hydrolysed in deionized dimethylsulfoxide or formamide, and the 5'-OH ends of the fragments were [^{32}P] labeled by T_4 polynucleotide kinase and γ [^{32}P] ATP as previously described (4,6). In these conditions neither intact tRNA nor fragments carrying the 5' end phosphate were labeled. The labeled fragments were purified by two dimensional polyacrylamide slab gel electrophoresis and eluted from the gel matrix. The 5' end [^{32}P] labeled terminal nucleotides were further characterized by enzymic digestion (P_1 RNase or venom phosphodiesterase) followed by thin layer chromatography (in one dimension for the major nucleotides : G, A, C and U and in two dimensions for the minor nucleotides). Most of the nucleotide sequence of the tRNAs could be established this way. However the regions between the extraloop and T- ψ and between nucleotide 25 and the anticodon loop, which contain G and C stretches, are able of refolding together and give very tight "hairpin structures" which could not be labeled by T_4 polynucleotide kinase. These fragments could even not be separated on polyacrylamide gel electrophoresis and appeared there as "compression zones". The nucleotide sequence of these "compression zones" were finally determined by two dimensional electrophoresis-homochromatography on DEAE-cellulose thin layers (4).

RESULTS AND DISCUSSION

The primary structure of arginine tRNA_{CCU} and tRNA_{mcm}^{5s2}UC ψ , and glutamine tRNA_{CUG} from bovine liver are shown in the cloverleaf model in

Fig. 1. Nucleotide sequence of tRNA^{Arg}_{CCU} from bovine liver.

Fig. 2. Nucleotide sequence of tRNA^{Arg}_{mcm⁵s²UCψ} from bovine liver.

Fig. 3. Nucleotide sequence of tRNA^{Gln}_{CUG} from bovine liver.

Figure 1, Figure 2 and Figure 3 respectively.

The two arginine tRNAs have the same length : 76 nucleotides. They differ by 25 nucleotides, thus having only 70% homology in their nucleotide sequences, the posttranscriptional modifications being not taken into account. They differ also by the number of modified nucleotides : 16 for tRNA^{Arg}_{mcm⁵s²UCψ} and only 13 for tRNA^{Arg}_{CCU}. The anticodon of the first of them is highly modified as is the whole loop : 4 nucleotides (m³C, mcm⁵s²U, ψ and m¹G) out of 7 are modified nucleotides for only 2 minor nucleotides in the anticodon loop and none in the anticodon of tRNA^{Arg}_{CCU}. The hypermodified mcm⁵s²U has been found before in that position in several other mammalian tRNAs (7), but ψ has never been found before in the 3' position of the anticodon in any other known tRNA. This tRNA is the second one sequenced up to date, carrying a minor nucleotide in that position, the other being tRNA^{Leu}₃ from bean chloroplast which has m¹G in that position (8). There are two subspecies of this arginine tRNA : one having ψ-ψ instead of the more commonly found T-ψ dinucleotide.

These two bovine arginine tRNAs are identical in the dihydrouridine loop and in the T-ψ loop including the last 3 pairs of the corresponding stem. The yeast arginyl- tRNA synthetase aminoacylates these two bovine tRNAs^{Arg} despite very large differences with their yeast homologues : 60% up to 65% of homology only. Here again most homologies between bovine and yeast tRNAs are

localized in the two loops (D and T-ψ) and in the T-ψ stem. These regions could therefore correspond to the target of the protein-tRNA recognition in the arginyl-tRNA synthetase-tRNA^{Arg} system.

The anticodons of these two arginine tRNAs contain only pyrimidic nucleotides. They therefore correspond to the tRNAs which would be expected to recognize the doublet of puric containing codons : A-G-A and A-G-G. It has been already suggested by others (9,10) that minor nucleotides such as s⁴U, mcm⁵U and mcm⁵s²U in the wobble position restrict the recognition to only A, it is therefore probable that these two tRNAs^{Arg} recognize each one of the two arginine codons : tRNA^{Arg}_{mcm⁵s²UCψ} recognizing A-G-A and tRNA^{Arg}_{CCU} recognizing A-G-G. It would be interesting to study the coding properties of these two tRNAs in order to confirm the above suggestion. A similar situation exists in yeast where the tRNA^{Arg}₃ has a mcm⁵U-C-U anticodon which also recognizes only A-G-A, involving the existence of another tRNA^{Arg} (in yeast) for translating A-G-G, which would correspond to the bovine tRNA^{Arg}_{CCU}.

As already mentioned, the nucleotide of two other arginine tRNAs from mammals are known (1,2). They are tRNA^{Arg}_{CCG} and tRNA^{Arg}_{ICG} which would be expected to recognize respectively the remaining arginine codon C-G-G, and the codons C-G-U, C-G-A and C-G-C. The tRNA^{Arg}_{CCG} could even misread the tryptophane codon U-G-G. It would be interesting to study the coding properties of this tRNA against the natural codon C-G-G and the tryptophane codon U-G-G. The presence of this tRNA^{Arg}_{CCG} in bovine liver suggests that such a tRNA exist, and is probably used in translation in mammals cytoplasm, despite the risk of tryptophane codon reading. This is different from yeast, where Weissenbach (9) explained the "absence" of tRNA^{Arg}_{CCG} because of the tryptophane codon misreading risk. Indeed at this date none G-G-C codon usage has been found in yeast, thus explaining the absence of tRNA^{Arg}_{CCG} in that organism. For A-G-G codon, only two usages have been found yet in yeast for a total of 83 arginine codon usages (11,12).

The overall sequence homologies of these two bovine liver arginine tRNAs to their other known arginine tRNAs from bovine are between 65% to 73%. The conserved residues in all sequenced arginine tRNAs whatever their origin, which are not universally present in all tRNAs are : G-G-C at nucleotides 9-11, the sequence G-C-Pu-ψ (or G-C-Pu-U) at nucleotides 24-27, G30, C32 or modified Cm32 or m³C32, A38, G45, D-U or U-U at nucleotides 47-48, G52, C62 and C66.

The glutamine tRNA^{Gln}_{CUG} is 75 nucleotides long including 13 minor ones (Figure 3). This tRNA is the first eucaryotic cytoplasmic tRNA^{Gln} sequenced up

to date. The T- ψ dinucleotide is replaced by ψ - ψ as it is in mammalian and avian tRNA^{Trp} and tRNA^{Pro} (7). The extraloop is only four nucleotides long.

Bovine liver glutamine tRNA^{CUG} is very different from the known tRNAs^{Gln} which are all from procaryotic origin (bacteria and phages). They have all together only 15 identical nucleotides including the invariant and semi invariant ones. All 16 known tRNAs^{Gln} and tRNA's^{Gln} genes (8) have only two conserved nucleotides in position G10 and U35 (in the anticodon), the 13 others are invariant or semi-invariant ones.

There is another major difference between this eucaryotic tRNA^{Gln} and all other known tRNAs^{Gln} of procaryotic origin except tRNA^{Gln} from T5 : this mammalian tRNA does not have a modified nucleotide next to the 3' side of the anticodon.

The anticodon C-U-G corresponds to the codon C-A-G. A second glutamine tRNA for decoding C-A-A is therefore probable in mammals.

ACKNOWLEDGEMENTS

We thank M. Fournier for providing us in BD-cellulose "0.8 M Choc" fractions of bovine liver tRNAs.

This work was supported by grants from INSERM (CRL n° 823.004) to Prof. G. Dirheimer in whose laboratory it was carried out. The skillfull technical assistance of Mrs C. Fix is gratefully acknowledged.

REFERENCES

1. Harada, F. and Nishimura, S. (1980) *Biochemistry international* 1, 539-546.
2. Miller, E.K., Pirtle, I.L., Dudock, B.S. and Pirtle, R.M. (1983) *Nucl. Acids Res.* 11, 2013-2016.
3. Fournier, M., Dorizzi, M., Sarger, C. and Labouesse, J. (1976) *Biochimie* 58, 1159-1167.
4. Keith, G., Pixa, G., Fix, C. and Dirheimer, G. (1983) *Biochimie*, in press.
5. Fradin, A., Gruhl, H. and Feldmann, H. (1975) *FEBS Lett.* 50, 185-189.
6. Keith, G. (1983) *Biochimie* 65, 367-370.
7. Gauss, D.H. and Sprinzl, M. (1983) *Nucl. Acids Res.* 11, r1-r53.
8. Osorio-Almeida, M.L., Guillemaut, P., Keith, G., Canaday, J. and Weil, J.H. (1980) *Biochem. Biophys. Res. Commun.* 92, 102-108.
9. Weissenbach, J. and Dirheimer, G. (1978) *Biochim. Biophys. Acta* 518, 350-354.
10. Dirheimer, G. (1983) in *Recent Results in Cancer Research*. Springer-Verlag Berlin, Heidelberg, 15-46.
11. Ikemura, T. (1982) *J. Mol. Biol.* 158, 573-597.
12. Bennetzen, J.L. and Hall, B.D. (1982) *J. Biol. Chem.* 257, 3026-3031.