Volume 250, number 2, 416–418 FEB 07306 July 1989

A small basic ribosomal protein in *Sulfolobus solfataricus* equivalent to L46 in yeast: structure of the protein and its gene

Celia Ramirez, K. Andrea Louie and Alastair T. Matheson

Department of Biochemistry and Microbiology, University of Victoria, Victoria V8W 2Y2, BC, Canada

The structure of the gene for a small, very basic ribosomal protein in *Sulfolobus solfataricus* has been determined and the structure of the protein coded by this gene (L46e) has been confirmed by partial amino acid sequencing. The protein shows substantial sequence homology to the eukaryotic ribosomal proteins L39 in rat and L46 in yeast. There is no sequence homology to any of the eubacterial ribosomal proteins suggesting that this protein is absent in the eubacterial ribosome.

Ribosomal protein; Gene structure; Archaebacteria; Evolution; (*Sulfolobus solfataricus*)

1. INTRODUCTION

It is now evident from the studies by Cammarano and his co-workers [1] that there are two distinct classes of ribosomes within the archaebacterial kingdom. The ribosomes from the extreme halophiles and most of the methanogens are similar in size and composition to those of the eubacteria while those from the thermophilic archaebacteria and several methanogens [2] are larger in mass and contain significantly more protein than do the eubacterial ribosomes. One would, therefore, expect to find ribosomal proteins in this latter group of archaebacteria that do not have a counterpart in eubacteria.

We have found an ORF upstream from the L11 operon in *Sulfolobus solfataricus* which codes for a basic ribosomal protein homologous to ribosomal protein L46 in yeast [3] and L39 in rat liver [4] but having no counterpart in eubacteria. We have isolated and characterized the ribosomal protein coded by this gene. Here, we report on the structure of the L46e protein and its gene.

Correspondence address: A. Matheson, Department of Biochemistry and Microbiology, University of Victoria, Victoria V8W 2Y2, BC, Canada

Abbreviations: ORF, open reading frame; L46e, equivalent ribosomal protein to L46 in yeast; bp, base pairs; Sce, *Saccharomyces cerevisiae;* Sso, *Sulfolobus solfataricus;* Rra, *Ratus ratus*

2. MATERIALS AND METHODS

2.1. *Sequencing of the L46e gene*

A 6.2 kb *Eco*RI-*Bam*HI fragment from an *S. solfataricus* DNA library, containing the L46e gene as well as genes for L11, L1, L10 and L12 (Ramirez and Matheson, in preparation), was cloned into pUC18 using standard methods [5]. The nucleotide sequence of the subcloned DNA fragment was determined using the dideoxy chain termination method [6]. A series of deletion plasmids [7] were constructed and used as templates.

2.2. *Identification of the L46e protein*

S. solfataricus P1 was grown at 85°C in the medium described by Zillig et al. [8]. The ribosomal subunits were isolated as described by Matheson et al. [9]. The 50 S ribosomal subunits were extracted with acetic acid [10] and the ribosomal proteins fractionated on Sephadex G-75. The low molecular mass proteins were further fractionated by HPLC on a C-8 reverse phase column (not shown). Proteins of the predicted size and relative charge were partially sequenced on an Applied Biosystems 470A gas-phase sequencer in a search for the L46e protein. A protein with the predicted amino acid sequence was identified and the sequence of the first 46 amino acid residues was determined on the intact protein.

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/89/$3.50

2.3. *Computer analysis*

The amino acid sequence of the L46e protein was compared to other proteins in the National Biomedical Research Foundation (NBRF) protein data bank using a FASTP computer program [11].

3. RESULTS AND DISCUSSION

Fig.1 shows the nucleotide sequence of the L46e gene and the predicted amino acid sequence of the protein. The gene codes for a small (50 residues), very basic (1 acidic residue and 16 basic residues) protein of molecular mass 5784 Da. Partial sequencing of the isolated protein confirmed the first 46 amino acid residues.

When the amino acid sequence of the *Sulfolobus* L46e protein was compared to the structures of the eubacterial ribosomal proteins, no sequence similarity was evident. This would suggest that this protein has no equivalent in the eubacterial ribosome. However, a search of the structure of known eukaryotic ribosomal proteins showed a substantial homology between L46e and yeast L46 [3] (40% identical residues) and rat liver L39 [4] (46% identical residues) as shown in fig.2. All 3 proteins are small (50 amino acids) and very basic. It should also be noted that the two tryptophan residues are conserved in all 3 proteins.

```
AGAAAAGATTATAAGATTACGATTAAAGAGAGAGG ATG GAA
                                     (M) (E)
1                                          33
ATG AGC AAG CAT AAG TCC TTA GGC AAA AAA TTG
 M   S   K   H   K   S   L   G   K   K   L
                                           66
AGA CTA GGT AAA GCG TTA AAA AGA AAC TCT CCT
 R   L   G   K   A   L   K   R   N   S   P
                                           99
ATT CCT GCT TGG GTC ATA ATA AAA ACT CAA GCT
 I   P   A   W   V   I   I   K   T   Q   A
                                          132
GAG ATA AGG TTT AAT CCA CTA AGA AGA AAT TGG
 E   I   R   F   N   P   L   R   R   N   W
                                          167
AGA AGA AAT AAT TTA AAG GTG TGA ATAAATGAGCC
 R   R   N   N   L   K   V   *
```

Fig.1. The nucleotide sequence of *S. solfataricus* L46e gene and the predicted amino acid sequence of the encoded protein. The solid line indicates the amino acid sequence obtained from the partial sequencing of the isolated protein. Two possible ATG initiator codons are present. Since the N-terminal amino acid sequence of the mature protein is Ser-Lys-His-, it is likely that the ATG closest to the AGC serine codon is the initiator.

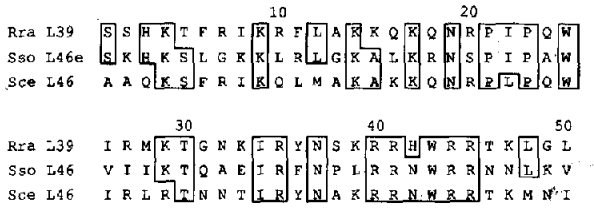

Fig.2. A comparison of the amino acid sequence of *Sulfolobus* L46e and two equivalent eukaryotic proteins, SceL46 from yeast [3] and RraL39 from rat liver [4]. Identical residues in the archaebacterial and eukaryotic proteins are indicated by the boxes.

Unlike the yeast L46 gene, which has a 385 bp intron at the 5′-end of the gene [3], no intron is present in the *Sulfolobus* L46e gene (fig.3).

Several other examples of archaebacterial ribosomal proteins, which appear to be absent in the eubacterial ribosome but present in eukaryotes, have been reported in *Methanococcus vannielii* [12] and *Halobacterium marismortui* [13]. However, studies by Cammarano et al. [1] indicate that the total number of ribosomal proteins in the extreme halophiles is very similar to that found in eubacteria. If this is so, the presence, in all branches of the archaebacteria, of ribosomal proteins which are not present in the eubacteria, raises the possibility that ribosomal proteins may be found that are common to eubacteria and eukaryotes but missing in the archaebacteria. If, as suggested by Auer et al. [12], the 70 S eubacterial ribosome resulted in part from a reduction in the number of ribosomal proteins present in the ancestral ribosome; a similar evolution of the archaebacterial ribosome might result in the loss of ribosomal proteins different from those removed during the evolution of the eubacterial ribosome. The sequencing of all the ribosomal proteins in a eukaryotic ribosome will be required to evaluate

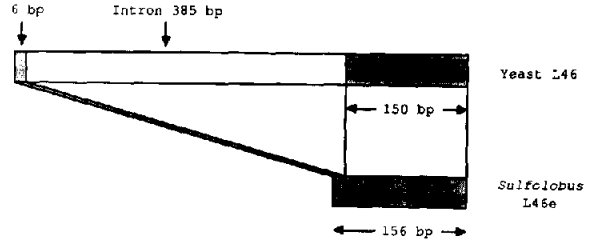

Fig.3. A comparison of the gene structure of yeast L46 and *Sulfolobus* L46e, indicating the location of the 385 bp intron in the yeast gene.

the above suggestion. It will also be of great interest to determine the function of the 'extra' ribosomal proteins present in archaebacteria.

Acknowledgements: This work was supported by NSERC Grant A6546 to A.T.M. C.R. was supported in part by a World University Service of Canada Scholarship and a Fellowship from the University of Victoria. We would also like to thank S. Kielland and the Tripartite Microsequencing Centre for help with the protein sequencing.

REFERENCES

[1] Cammarano, P., Teichner, A. and Londei, P. (1986) System. Appl. Microbiol. 7, 137–146.
[2] Schimid, G. and Böck, A. (1982) Zbl. Bakt. Hyg. I. Abt. Orig. C3, 347–353.
[3] Leer, R.J., Van Raamsdonk-Duin, M.M.C., Hagendoorn, M.J.M., Mager, W.H. and Planta, R.J. (1985) Nucleic Acids Res. 13, 701–709.
[4] Lin, A., McNally, J. and Wool, I.G. (1984) J. Biol. Chem. 259, 487–490.
[5] Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
[6] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
[7] Lin, H-C., Lei, S.-P. and Wilcox, G. (1985) Anal. Biochem. 147, 114–119.
[8] Zillig, W., Stetter, K.O., Wunderl, S., Schulz, W., Priess, H. and Scholz, I. (1980) Arch. Microbiol. 125, 259–269.
[9] Matheson, A.T., Louie, K.A. and Böck, A. (1988) FEBS Lett. 231, 331–335.
[10] Hardy, S.J.S., Kurland, C.G., Voynow, P. and Mora, G. (1969) Biochemistry 8, 2897–2905.
[11] Lipmann, D.J. and Pearson, W.R. (1985) Science 227, 1435–1441.
[12] Auer, J., Lechner, K. and Böck, A. (1989) Can. J. Microbiol. 35, 200–204.
[13] Kimura, M., Arndt, E., Hatakeyama, T. and Kimura, J. (1989) Can. J. Microbiol. 35, 195–199.