The complete amino acid sequence of the ribosomal A protein (L12) from the archaebacterium *Sulfolobus acidocaldarius*

A.T. Matheson, K.A. Louie and August Böck[+]

Department of Biochemistry and Microbiology, University of Victoria, Victoria, BC V8W 2Y2, Canada and [+]Institut für Genetik und Mikrobiologie, Universität München, Maria-Ward-Strasse 1a, 8000 München 19, FRG

The ribosomal A protein (SacL12) from the archaebacterium *Sulfolobus acidocaldarius* has been sequenced. The protein contains 105 amio acids, has a composition of Asp_2, Asn_3, Thr_4, Ser_6, Glu_{17}, Gln_4, Pro_3, Gly_7, Ala_{18}, Val_7, Met_2, Ile_7, Leu_9, Tyr_2, Phe_1, His_1, Lys_{11}, Arg_1 and a molecular mass of 11 126 Da. The *Sulfolobus* protein shows many features in common with the equivalent proteins in the eukaryotes such as 35–40% sequence homology and similar hydrophilicity profiles, features much less evident when this protein is compared to eubacterial L12 proteins. SacL12 contains an unusual sequence of alternating clusters of lysine and glutamic acid (–EKKEEKKEEEKK–) in the C-terminal region. Similar sequences are found in some eukaryotic L12 proteins.

Ribosomal A protein; Amino acid sequence; (*Sulfolobus acidocaldarius*)

1. INTRODUCTION

In eubacteria such as *Escherichia coli* [1,2] and *Bacillus stearothermophilus* [3,4] the ribosomal A protein L12 (The ribosomal proteins have been designated by the equivalent protein in *E. coli* (Eco) as determined by amino acid sequence similarity. The ribosomal A protein from *S. acidocaldarius* (Sac) is designated SacL12 and is equivalent to EcoL12.) is present as a 4:1 complex with ribosomal protein L10 and forms a well defined domain in the 50 S ribosomal subunit [5]. The L12 protein is thought to be the binding site for several of the factors involved in protein synthesis [6] and appears to be essential for accurate translation [7].

Because of its ubiquitous nature and ease of purification, the L12 protein has been used as a phylogenetic probe to study molecular evolution (see [8]). A large number of L12 proteins have been completely sequenced from eubacterial and eukaryotic sources (see [9]) and partial sequences have also been reported from archaebacterial sources [8,10,11].

In this paper we report the complete amino acid sequence of the ribosomal A protein SacL12 from the archaebacterium *Sulfolobus acidocaldarius*.

2. MATERIALS AND METHODS

Sulfolobus acidocaldarius (DSM 639) was grown at 85°C in the medium described by Zillig et al. [12] and harvested in mid-late log phase growth. The cells were disrupted using a French press (16000 psi) in a buffer containing 20 mM Tris, pH 7.5, 100 mM KCl, 10 mM MgOAc, 6 mM βMe. The ribosomes were dissociated in the extraction buffer containing 1 mM Mg^{2+} and separated in a Type 15 zonal rotor using a 8–32% sucrose gradient.

The ribosomal protein SacL12 was selectively extracted from the 50S ribosomal subunit at 46°C with NH_4Cl/ethanol and purified on CM-cellulose at pH 4.6 [4]. The protein was further purified on HPLC (Beckman RPSC C-3 column, 0.1% trifluoroacetic acid/acetonitrile (TFA/ACN), 0–60% gradient over 60 min).

The amino acid composition of the protein and peptides was determined by hydrolysis in constantly boiling HCl for 24 h and the hydrolysate was analyzed on a Beckman 119CL amino acid analyzer.

The protein and peptides were sequenced using an Applied

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

Biosystems 470-A gas phase sequenator with an on-line PTH analyzer model 120A.

The peptides were obtained by treating the purified protein with either cyanogen bromide or proteolytic enzymes (see [13] for methods). The peptides were fractionated on HPLC (Beckman RPSC C-3 column, gradient 0.1% TFA/ACN) and further identified on Swank-Munkres gels [14]. The peptides used for sequencing were chosen in part by their size and amino acid composition and only those peptides required to complete the overall sequence were analyzed further. The proteolytic enzymes used were trypsin (Sigma) and carboxypeptidase Y (Boehringer, Mannheim). In some experiments using trypsin, the ribosomal protein was chemically modified by maleylation of the lysine residues [13].

The hydrophilicity of the SacL12 protein was determined using the SURFPLOT program of Parker et al. [15].

3. RESULTS AND DISCUSSION

The complete amino acid sequence of SacL12 is shown in fig.1. No amino acids were released when the intact protein was treated in the sequenator suggesting SacL12 contained a block N-terminal residue. When the intact protein was treated with cyanogen bromide, two peptides were obtained and these peptides contained all the amino acids present in the intact protein except for one methionine residue (table 1). Since both cyanogen bromide peptides contained free N-terminal residues, it suggested the blocked N-terminal amino acid in SacL12 might be this missing
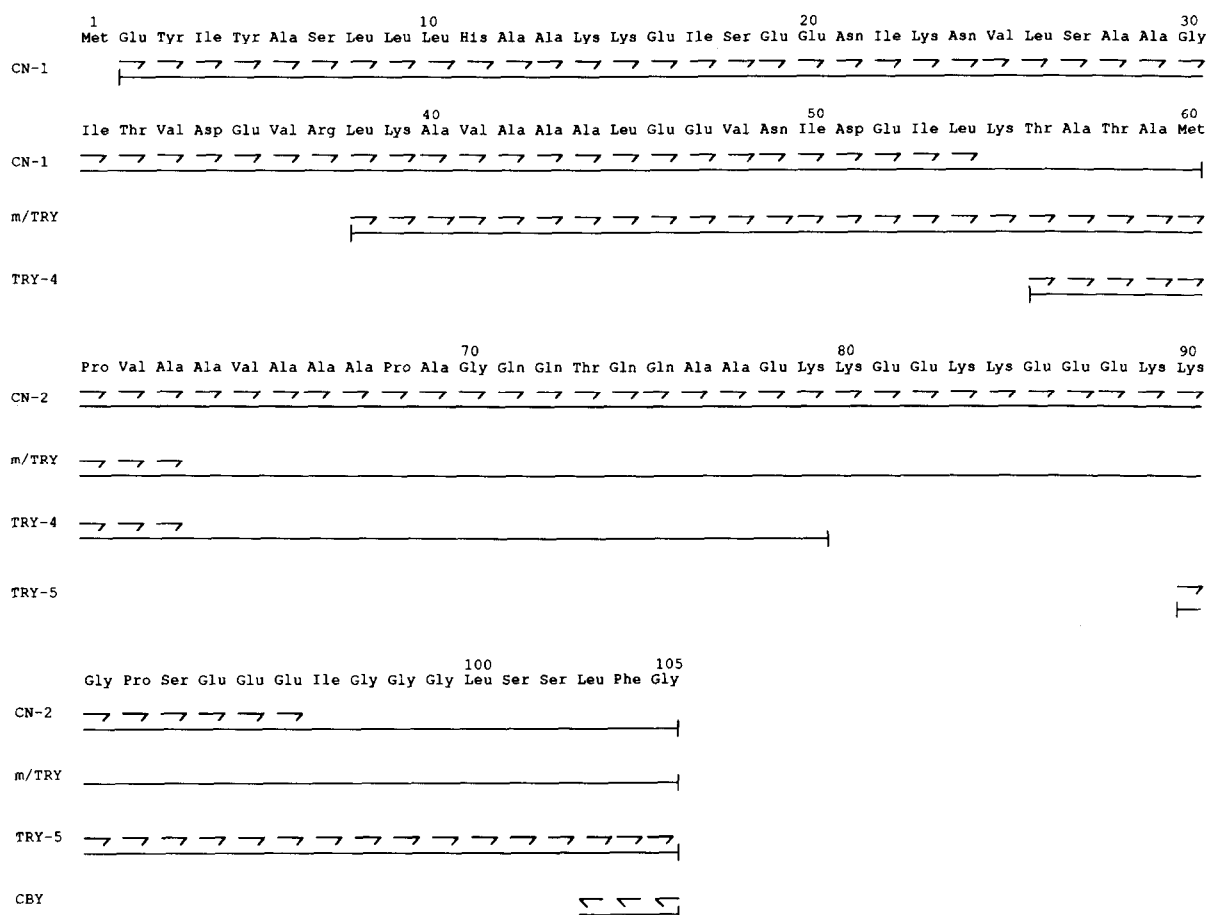

Fig.1. The amino acid sequence of the ribosomal A protein SacL12 from *S. acidocaldarius*. The methods used to determine the amino acid sequence of this protein are indicated as follows: ⟶, automatic Edman degradation of peptides; ⟵, carboxypeptidase Y hydrolysis of intact protein; CN, peptides obtained from cleavage with cyanogen bromide; TRY and m/TRY, peptides obtained by digestion with trypsin and trypsin of SacL12 and maleylate SacL12, respectively. The N-terminal Met residue appears to have a blocked αNH$_2$ group.

Table 1

The amino acid composition of the ribosomal protein SacL12 and the peptides obtained by treatment of the protein with cyanogen bromide

	Intact protein		CN1	CN2
	analysis	sequence	sequence	sequence
Asp	5.4	2	2	0
Asn		3	3	0
Thr	3.8	4	3	1
Ser	6.2	6	3	3
Glu	20.7	17	8	9
Gln		4	0	4
Pro	2.9	3	0	3
Gly	6.9	7	1	6
Ala	18.6	18	11	7
Val	6.7	7	5	2
Met	2.1	2	1	0
Ile	6.3	7	6	1
Leu	8.7	9	7	2
Tyr	1.8	2	2	0
Phe	0.9	1	0	1
His	0.8	1	1	0
Lys	11.6	11	5	6
Arg	0.9	1	1	0
Total		105	59	45

methionine. This was subsequently confirmed from the structure of the SacL12 gene (Ramirez, C., personal communication).

As shown in fig.1, residues 2–54 were obtained from peptide CN-1 while residues 61–96 were obtained from peptide CN-2. Two of the tryptic peptides, TRY-4 and TRY-5, provided residues 56–63 and 90–105, respectively. The C-terminal peptide, obtained when the maleylated protein was hydrolyzed with trypsin at the single arginine residue, provides the sequence of residues 38–63, confirming the sequence of this region of the protein. Digestion of SacL12 with carboxypeptidase Y confirmed that the C-terminal sequence was -Leu-Phe-Gly-COOH. From these results the complete amino acid sequence of SacL12 was obtained. It contains 105 amino acids and has an M_r of 11 126.

The hydrophilicity profile of SacL12 was determined [15] and showed more similarity to the equivalent eukaryotic protein (AsaL12) than to the corresponding eubacterial protein (EcoL12), as indicated in fig.2.

A comparison of the amino acid sequence of SacL12 with the equivalent proteins from eukaryotes and eubacteria indicated that the
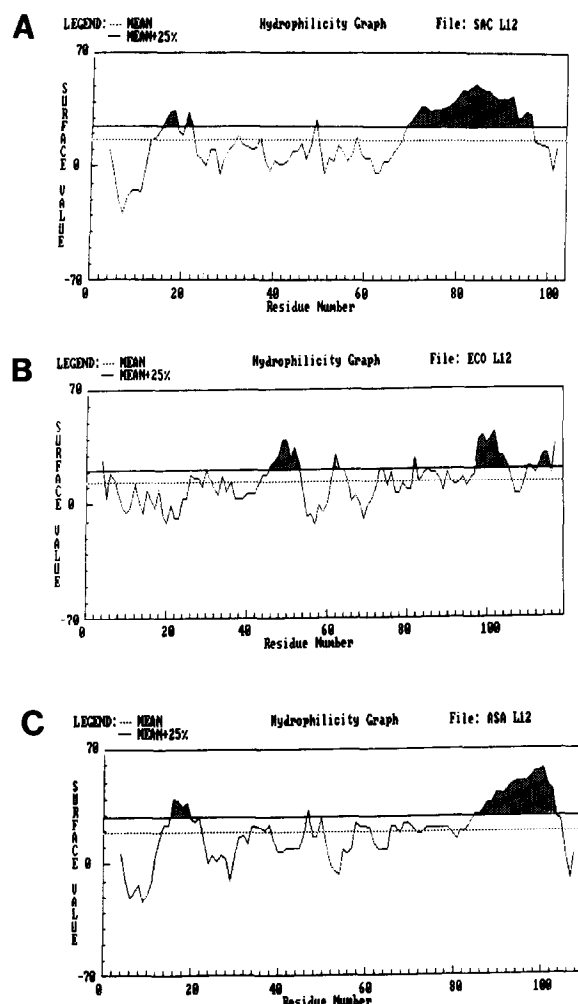

Fig.2. The hydrophilicity of ribosomal protein L12 from (A) S. acidocaldarius (SacL12), (B) E. coli [16] (EcoL12) and (C) Artemia salina [17] (AsaL12), using the SURFPLOT program of Parker et al. [15].

Sulfolobus protein showed much more sequence similarity to the equivalent eukaryotic proteins than to those from eubacteria [21]. When the sequence of the SacL12 is compared to the equivalent proteins in eukaryotes (fig.3), 35–40% of the 105 amino acids in the Sulfolobus protein are identical to those present in the eukaryotic L12 proteins and no rearrangement of the molecules is required.

However, when the archaebacterial L12 protein is compared to the equivalent proteins from eubacteria, rearrangement of the eubacterial L12 is required in order to obtain optimal structural
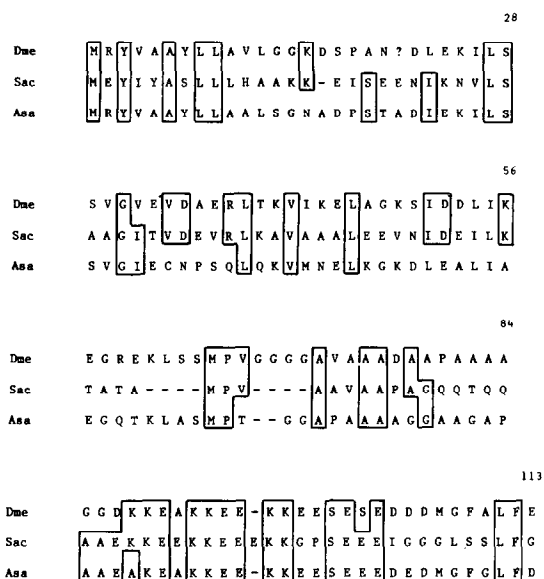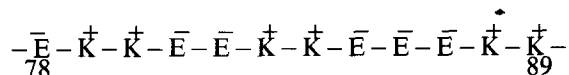
```
                                                  28
Dme   M R Y V A A Y L L A V L G G K D S P A N ? D L E K I L S
Sac   M E Y I Y A S L L L H A A K K - E I S E E N I K N V L S
Asa   M R Y V A A Y L L A A L S G N A D P S T A D I E K I L S

                                                  56
Dme   S V G V E V D A E R L T K V I K E L A G K S I D D L I K
Sac   A A G I T V D E V R L K A V A A A L E E V N I D E I L K
Asa   S V G I E C N P S Q L Q K V M N E L K G K D L E A L I A

                                                  84
Dme   E G R E K L S S M P V G G G G A V A A A D A A P A A A A
Sac   T A T A - - - - M P V - - - - A A V A A P A C Q Q T Q Q
Asa   E G Q T K L A S M P T - - G G A P A A A G G A A G A P

                                                  113
Dme   G G D K K E A K K E E - K K E E S E S E D D D M G F A L F E
Sac   A A E K K E E K K E E E K K C P S E E E I G G G L S S L F G
Asa   A A E A K E A K K E E - K K E E S E E E D E D M G F G L F D
```

Fig.3. Amino acid sequence homology between the archaebacterial ribosomal protein L12 (Sac) and the equivalent protein from two eukaryotic sources, *Drosophila melanogaster* (DmeL12) [18] and *Artemia salina* (AsaL12) [17]. Identical amino acids in the archaebacterial and eukaryotic proteins are indicated by boxes.

similarity [20,21]. A similar situation exists when eukaryotic and eubacterial L12 proteins are compared [19]. Several transposition models have been proposed [8,19,20] depending on the parameters that are compared. A detailed comparison of the eubacterial and archaebacterial L12 proteins will be reported elsewhere (in preparation).

Several specific structural features of the *Sulfolobus* L12 protein should be noted. SacL12, like all archaebacterial and eukaryotic L12 proteins, contains 2 tyrosine residues which are present in the N-terminal region of these proteins. The eubacterial L12 proteins contain no tyrosine.

The SacL12 protein contains a single histidine residue, an amino acid absent in all other L12 proteins studied thus far except for the L12 protein from the archaebacterium *Methanobacterium thermoautotrophicum* where the single histidine is also located in residue 11 [11].

The C-terminal region of SacL12 contains an unusual sequence of alternating clusters of lysines and glutamic acids

$$-\overset{-}{\text{E}}-\overset{+}{\text{K}}-\overset{+}{\text{K}}-\overset{-}{\text{E}}-\overset{-}{\text{E}}-\overset{+}{\text{K}}-\overset{+}{\text{K}}-\overset{-}{\text{E}}-\overset{-}{\text{E}}-\overset{-}{\text{E}}-\overset{\ast}{\overset{+}{\text{K}}}-\overset{+}{\text{K}}-$$
$$\underset{78}{} \qquad\qquad\qquad\qquad\qquad\qquad \underset{89}{}$$

as shown in fig.1. A survey of protein data banks (Bionet) indicates this sequence is unique to the L12 protein. Although the equivalent sequence is missing in eubacterial L12 proteins and in HcuL12 [21], a similar sequence is found in many of the eukaryotic L12 proteins (see fig.3) emphasizing again that the *Sulfolobus* L12 protein shows properties much closer to those of the eukaryotic L12 proteins than to the equivalent eubacterial proteins.

REFERENCES

[1] Osterberg, R., Sjöberg, B., Pettersson, I., Liljas, A. and Kurland, C.G. (1977) FEBS Lett. 73, 22–24.

[2] Gudkov, A.T., Tumanova, L.G., Venyaminov, S.Yu. and Khechinashvilli, N.N. (1978) FEBS Lett. 93, 215–218.

[3] Marquis, D. and Fahnestock, S.R. (1978) J. Mol. Biol. 119, 557–567.

[4] Marquis, D. and Fahnestock, S.R. (1980) J. Mol. Biol. 142, 161–179.

[5] Lake, J.A. (1980) in: Ribosomes: Structure, Function, and Genetics (Chambliss et al. eds) pp.207–236, University Park Press, Baltimore, MD.

[6] Möller, W. (1974) in: Ribosomes (Numura, M. et al. eds) pp.711–731, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

[7] Pettersson, I. and Kurland, C.G. (1980) Proc. Natl. Acad. Sci. USA 77, 4007–4010.

[8] Matheson, A.T. (1985) in: The Bacteria, vol.8, Archaebacteria (Woese, C.R. and Wolfe, R.S. eds) pp.345–377, Academic Press, Orlando, FL.

[9] Wittmann-Liebold, B. (1986) in: Ribosomes: Structure, Function, and Genetics (Hardesty, B. and Kramer, G. eds) pp.326–361, Springer-Verlag, New York.

[10] Yaguchi, M., Matheson, A.T., Visentin, L.P. and Zuker, M. (1980) in: Genetics and Evolution of RNA Polymerase, tRNA and Ribosomes (Osawa, S. et al. eds) pp.589–599, University of Tokyo Press, Tokyo.

[11] Matheson, A.T., Yaguchi, M., Balch, W.E. and Wolfe, R.S. (1980) Biochim. Biophys. Acta 626, 162–169.

[12] Zillig, W., Stetter, K.O., Wunderl, S., Schulz, W., Priess, H. and Scholz, I. (1980) Arch. Microbiol. 125, 259–269.

[13] Allen, G. (1981) Laboratory Techniques in Biochemistry and Molecular Biology, vol.9, Elsevier, Amsterdam, New York.

[14] Swank, R.T. and Munkres, K.D. (1971) Anal. Biochem. 39, 462–477.

[15] Parker, J.M.R., Guo, D. and Hodges, R.S. (1986) Biochemistry 25, 5425–5432.

[16] Terhorst, C., Möller, W., Laursen, R. and Wittmann-Liebold, B. (1973) Eur. J. Biochem. 34, 138–152.

[17] Maassen, J.A., Schop, E.N., Brands, J.H.G.M., Van Hemert, F.J., Linstra, J.A. and Möller, W. (1985) Eur. J. Biochem. 149, 609–616.

[18] Qian, S., Zhang, J.-Y., Kay, M.A. and Jacobs-Lorena, M. (1987) Nucleic Acids Res. 15, 987–1003.

[19] Lin, A., Wittmann-Liebold, B., McNally, J. and Wool, I.G. (1982) J. Biol. Chem. 257, 9189–9197.

[20] Liljas, A., Thirup, S. and Matheson, A.T. (1986) Chemica Scripta 26B, 109–119.

[21] Matheson, A.T., Louie, K.A. and Henderson, G.N. (1986) System. Appl. Microbiol. 7, 147–150.