Nucleotide sequence of the *lspA* gene, the structural gene for lipoprotein signal peptidase of *Escherichia coli*

Fujio Yu, Hisami Yamada, Kyoko Daishima and Shoji Mizushima*

Laboratory of Microbiology, Faculty of Agriculture, Nagoya University, Chikusa-ku, Nagoya 464, Japan

Received 13 June 1984

The nucleotide sequence of the *lspA* gene coding for lipoprotein signal peptidase of *Escherichia coli* was determined and the amino acid sequence of the peptidase was deduced from it. The molecular mass and amino acid composition of the predicted lipoprotein signal peptidase were consistent with those of the signal peptidase purified from cells harboring the *lspA* gene-carrying plasmid. The peptidase most probably has no cleavable signal peptide. The *lspA* gene was preceded by the *ileS* gene coding for isoleucyl-tRNA synthetase and the tandem termination codons of the *ileS* gene overlapped with the initiation codon of the *lspA* gene.

Lipoprotein signal peptidase *DNA sequence* *Isoleucyl-tRNA synthetase* ileS-lspA *operon*
Lipoprotein *Protein secretion*

1. INTRODUCTION

Signal peptidase plays an essential role in secretion of proteins across membranes. Two signal peptidases have been identified, leader peptidase and lipoprotein signal peptidase of *Escherichia coli*. The former was purified [1] and the nucleotide sequence of the structural gene (*lep*) determined [2]. The latter was also characterized biochemically [3] and genetically [4,5]. Recently, it was found that the *lspA* gene coding for lipoprotein signal peptidase and the *ileS* gene coding for isoleucyl-tRNA synthetase constitute a co-transcriptional unit with the order of promoter-*ileS-lspA*, and they localized the *lspA* gene within the 1.1 kb *Stu*I-*Eco*RI fragment [6] (see fig.1).

Here, we sequenced the entire *Stu*I-*Eco*RI region to determine the DNA sequence of the *lspA* gene and to deduce the amino acid sequence of the signal peptidase.

* To whom correspondence should be addressed

Abbreviations: LP, major outer membrane lipoprotein; kb, kilobase pairs

2. MATERIALS AND METHODS

2.1. *Materials*

Restriction endonucleases and an M13 sequencing kit were obtained from Takara Shuzo and deoxycytidine 5′-[α-^{32}P]triphosphate from Amersham International.

2.2. *Plasmids and bacteria*

Plasmids pYK160T [6] and pKD15 were used. pKD15 was constructed by insertion of the *lspA* gene-carrying 4.3 kb *Eco*RI fragment of pHY11 [7] into the *Eco*RI site of pSY343, a runaway replication plasmid vector [8]. The *E. coli* strain used was JE5506 [9], which was grown on L broth. For enzyme purification, cells grown in M9-glucose as described in [6] were used.

2.3. *DNA sequencing*

The DNA sequence was determined by the chain termination method in [10]. Restriction endonuclease fragments were subcloned into M13 phage vectors mp8 and mp9 [11]. JM103 cells were used as host cells.

Published by Elsevier Science Publishers B.V.

2.4. *Purification of lipoprotein signal peptidase*

The *Eco*RI fragment carrying the *trp* promoter-controlled *lspA* gene was prepared from pYK160T [6] and inserted into the *Eco*RI site of pSY343 to construct pYKR160T. JE5506 was then transformed with pYKR160T. After induction of runaway replication of pYKR160T at 37°C and that of lipoprotein signal peptidase by β-indoleacrylic acid (20 μg/ml), cells were harvested and the peptidase was purified based on the method in [3]. DEAE–cellulose column chromatography was carried out in the presence of 2 mM Na-EDTA. Details of the purification procedure will be published elsewhere.

2.5. *Other methods*

The in vitro assay for lipoprotein signal peptidase was described in [3]. The amino acids were determined with a Hitachi 835 amino acid analyzer as in [12].

3. RESULTS AND DISCUSSION

3.1. *DNA sequencing*

The 1.1 kb *Stu*I-*Eco*RI fragment was isolated from pKD15. This fragment carries the entire *lspA* gene cloned in pYK160T [6]. A restriction endonuclease cleavage map of the 1.1 kb fragment was constructed (fig.1) and the nucleotide sequence covering the entire region was determined (fig.2). In the upstream region, there is an open reading frame of 435 bp (nucleotides 2–436). The deduced amino acid sequence of this region (nucleotides 323–424) is consistent with the sum of those of the peptides derived from isoleucyl-tRNA synthetase (*ileS* product) [13], indicating that the region is the 3′-terminus of the *ileS* gene. The reading frame is terminated by contiguous termination codons (nucleotides 437–442). Another long open reading frame (nucleotides 439–930) is initiated at the termination codons and terminated at nucleotides 931–936, thus making a protein of about 18 kDa. Since this is the only long open reading frame initiated with the ATG codon, we tentatively concluded that the DNA sequence represents the *lspA* gene. The sequence is followed by the possible ρ-independent transcription termination signal (nucleotides 964–1012) [14]. The amino acid sequence deduced is also shown in fig.2. Consistent with [6], the proposed *lspA* gene was found to follow the *ileS* gene immediately without its own promoter being in between.

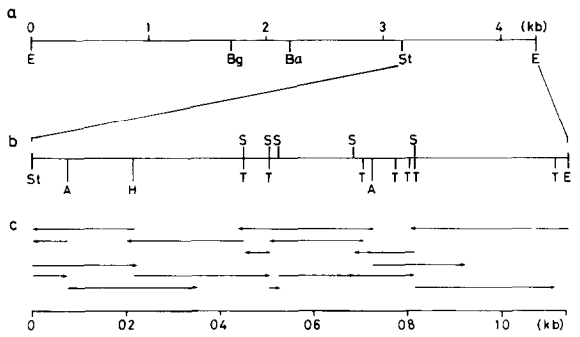

Fig.1. Restriction map and sequencing strategy for the *lspA*-carrying 1.1 kb *Stu*I-*Eco*RI fragment. (a) The *lspA*-carrying 4.3 kb fragment derived from the *E. coli* chromosome [7]. (b) The *lspA*-carrying 1.1 kb *Stu*I-*Eco*RI fragment. (c) The sequencing strategy. Arrows indicate the length and direction of DNA sequences determined from each sequencing reaction. Restriction sites: A, *Alu*I; Ba, *Bam*HI; Bg, *Bgl*II; E, *Eco*RI; H, *Hin*cII; S, *Sau*3AI; St, *Stu*I; T, *Tth*HB8I (*Taq*I).

3.2. *Purification of lipoprotein signal peptidase*

To prove that this open reading frame codes for lipoprotein signal peptidase, the peptidase was purified as described in section 2. The purified preparation was rich in one protein band, the molecular mass of which was about 18 kDa (fig.3, lane D). The band was prominent when the cells harbored pYKR160T (fig.3, lanes A–C), and was always cofractionated with the signal peptidase activity throughout the purification. These results indicate that the band represents lipoprotein signal peptidase, which in turn supports the view that the open reading frame codes for the signal peptidase. The amino acid sequence shown in fig.2 does not support the presence of a cleavable signal peptide. The coincidence of the predicted molecular mass (fig.2) with that determined on gel (fig.3) also indicates the absence of a signal peptide.

3.3. *Amino acid composition of the purified lipoprotein signal peptidase*

The signal peptidase preparation was further purified in 0.2% SDS–0.2 M Na-phosphate buffer (pH 7.1) solution on a Toyo G-3000SW column with a Hitachi 638 high-performance liquid chromatograph (fig.3, lane E). After recycling of

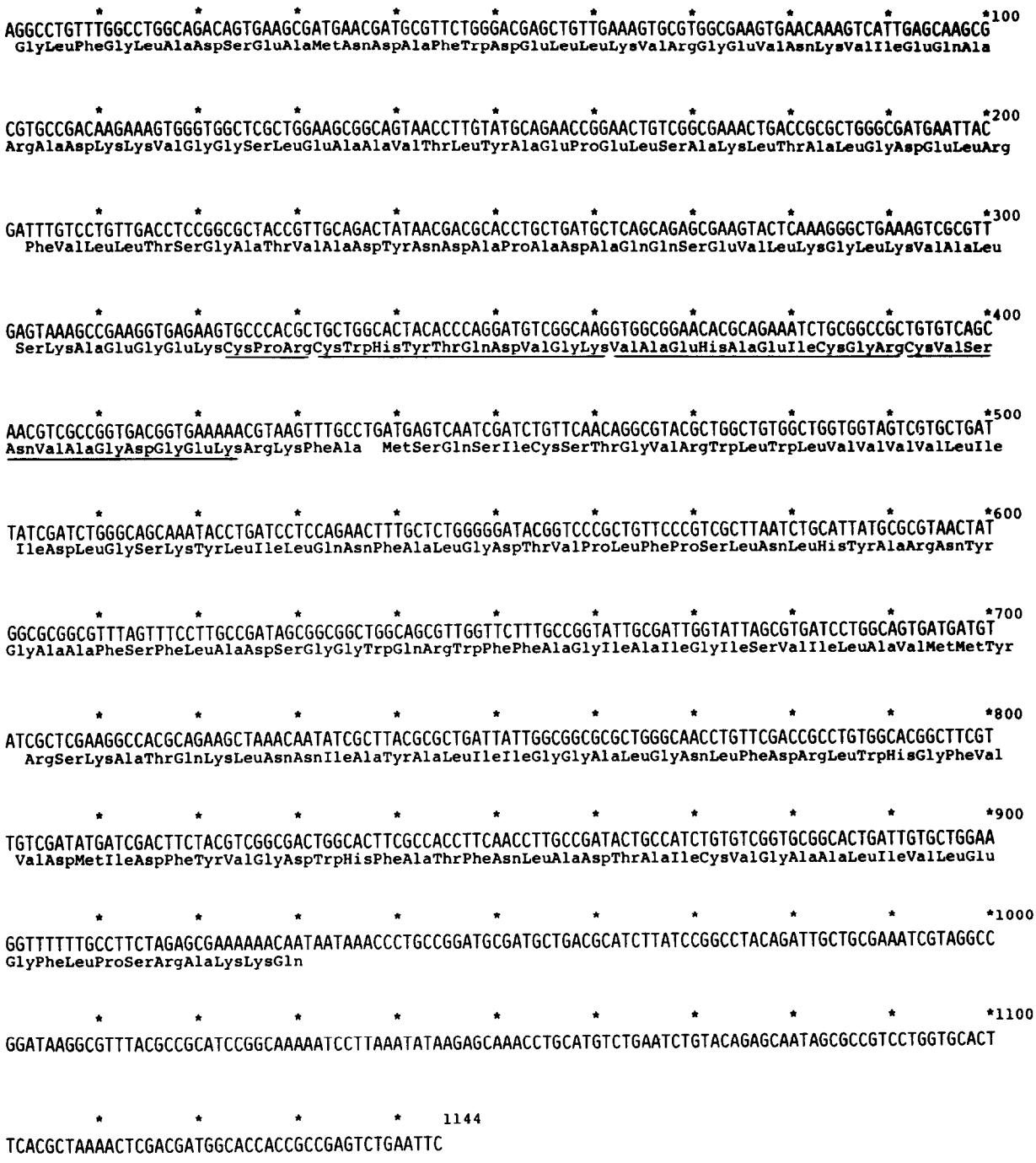

Fig.2. The nucleotide sequence encompassing the *lspA* gene and the deduced amino acid sequence of lipoprotein signal peptidase. The downstream region of the *ileS* gene and the deduced amino acid sequence of isoleucyl-tRNA synthetase are also shown. The amino acid sequences within the synthetase that had been determined previously [13] are underlined.

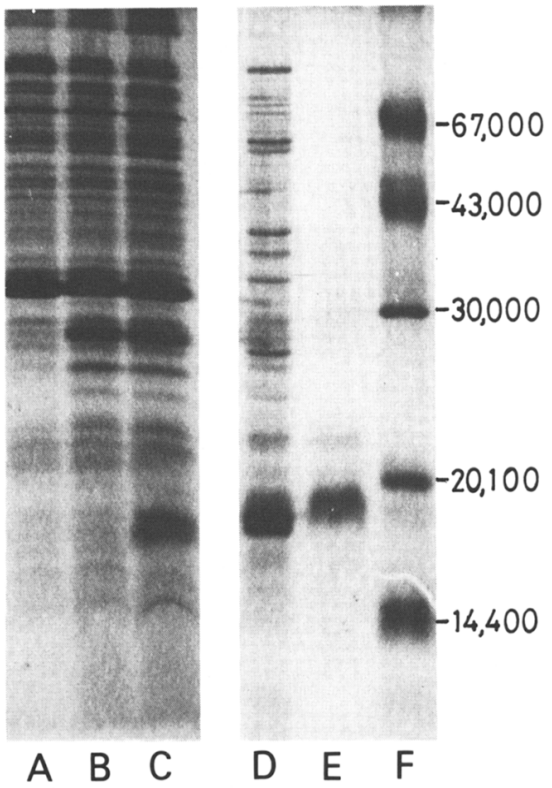

Fig.3. Gel electrophoretic profiles of lipoprotein signal peptidase preparations. Lanes A,B,C: cell envelope fractions of JE5506, JE5506/pSY343 and JE5506/pYKR160T, respectively. Lipoprotein signal peptidase was purified by ammonium sulfate fractionation and DEAE–cellulose chromatography as in [3] (lane D) and further purified on a G-3000SW column (lane E). Lane F, M_r standards; bovine serum albumin (67000), ovalbumin (43000), carbonic anhydrase (30000), trypsin inhibitor (20100), α-lactalbumin (14400). Gel electrophoresis was carried out as in [16].

the G-3000SW column chromatography, the amino acid composition was determined. The result essentially coincided with that predicted from the deduced amino acid sequence (table 1). The hydropathy (hydrophobicity and hydrophilicity) of the signal peptidase was calculated along the amino acid sequence as in [15] (fig.4). The peptidase is very hydrophobic as a whole and contains several long hydrophobic segments that may be responsible for spanning the cytoplasmic membrane of this protein. The amino acid sequence of the amino terminus of the purified enzyme could not be determined.

Table 1

Amino acid composition of lipoprotein signal peptidase

Amino acid	Residues in polypeptide (mol/mol)	
	Predicted	Obtained[a]
Asp + Asn	15	14.7
Thr[b]	5	5.8
Ser[b]	10	8.0
Glu + Gln	6	11.9
Pro	3	2.2
Gly	15	16.7
Ala	18	18.1
Val	13	12.0
Met	4	1.5
Ile	14	11.3
Leu	21	21.0
Tyr	6	–
Phe	12	9.7
His	3	1.6
Lys	5	6.1
Arg	6	6.1
Trp	6	–
Half-Cys	2	–

[a] The number of leucine residues was taken as 21
[b] Extrapolated to zero time

(–) Not determined

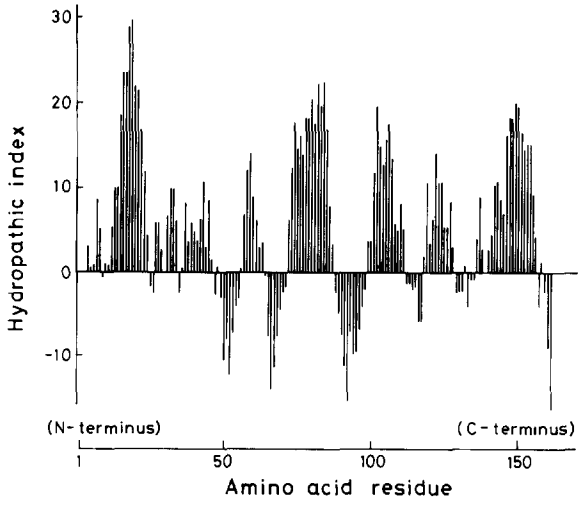

Fig.4. Hydropathy patterns of lipoprotein signal peptidase. Hydropathy was calculated with a span of 7 residues. The portion above the midpoint line represents the hydrophobic regions.

ACKNOWLEDGEMENTS

We thank Drs K. Takahashi and T. Kageyama for amino acid analysis, Dr T. Tsuchiya for hydropathy analysis and S. Teranishi for secretarial support. This work was supported by grants from the Ministry of Education, Science and Culture of Japan and Nisshin Seifun Foundation.

REFERENCES

[1] Wolfe, P.B., Silver, P. and Wickner, W. (1982) J. Biol. Chem. 257, 7898–7902.
[2] Wolfe, P.B., Wickner, W. and Goodman, J.M. (1983) J. Biol. Chem. 258, 12073–12080.
[3] Yamada, H., Yamagata, H. and Mizushima, S. (1984) FEBS Lett. 166, 179–182.
[4] Yamagata, H., Taguchi, N., Daishima, K. and Mizushima, S. (1983) Mol. Gen. Genet. 192, 10–14.
[5] Regue, M., Remenick, J., Tokunaga, M., Mackie, G.A. and Wu, H.C. (1984) J. Bacteriol. 158, 632–635.
[6] Yamada, H., Kitagawa, M., Kawakami, M. and Mizushima, S. (1984) FEBS Lett., in press.
[7] Yamagata, H., Daishima, K. and Mizushima, S. (1983) FEBS Lett. 158, 301–304.
[8] Yasuda, S. and Takagi, T. (1983) J. Bacteriol. 154, 1153–1161.
[9] Yamagata, H., Ippolito, C., Inukai, M. and Inouye, M. (1982) J. Bacteriol. 152, 1163–1168.
[10] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
[11] Messing, J. and Vieira, J. (1982) Gene 19, 269–276.
[12] Spackman, D.H., Stein, W.H. and Moore, S. (1958) Anal. Chem. 30, 1190–1206.
[13] Kula, M.-R. (1973) FEBS Lett. 35, 299–302.
[14] Rosenberg, M. and Court, D. (1979) Annu. Rev. Genet. 13, 319–353.
[15] Kyte, J. and Doolittle, R.F. (1982) J. Mol. Biol. 157, 105–132.
[16] Laemmli, U.K. (1970) Nature 227, 680–685.