

AMINO ACID SEQUENCE OF THE SIGNAL PEPTIDE OF OmpF, A MAJOR OUTER MEMBRANE PROTEIN OF *ESCHERICHIA COLI*

Norihiro MUTOH, Kaoru INOKUCHI and Shoji MIZUSHIMA

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

Received 2 December 1981

1. Introduction

Studies on a wide variety of secretory proteins have substantiated the signal hypothesis [1]. The signal hypothesis states that a protein exported through a membrane is synthesized as a precursor form with a peptide extension at the NH₂-terminal end (signal peptide). The cell envelope of Gram-negative bacteria consists of 2 membranous organelles; the outer membrane and the cytoplasmic membrane. Therefore, outer membrane proteins that are synthesized in the cytoplasm have to be exported through the cytoplasmic membrane. Indeed, almost all of the outer membrane proteins so far studied were found to be first synthesized as a precursor form, and in 3 cases (OmpA [2], major lipoprotein [3] and LamB [4]) the amino acid sequence of the signal peptide has been determined. The OmpF protein is another major outer membrane protein to be studied. To study the chemical structure of OmpF and its precursor as well as the mechanism controlling the synthesis and subsequent localization in the outer membrane of this protein, we isolated a specialized transducing λ phage that carries the *ompF* gene [5]. Here, we have subcloned the *ompF* gene and determined the DNA sequence that covers the NH₂-terminal region of OmpF. From the DNA sequence the amino acid sequence of the signal peptide was deduced. The general features of the signal peptides and those of the cleavage site are discussed in relation to protein localization in the cell surface.

2. Materials and methods

2.1. Materials

[γ -³²P]ATP (9000 Ci/mmol) was prepared from carrier-free [³²P]orthophosphate (Amersham Intl.)

and ADP (Sigma Chemicals) as in [6], and used for DNA sequencing. Restriction endonucleases *Eco*RI, *Hae*III, *Hap*II, *Hind*III, *Hha*I and *Sal*I were products of Takara Shuzo Co. and *Pst*I, *Pvu*II and *Taq*I were from New England Biolabs. Bacterial alkaline phosphatase was from Worthington Biochemical Corp., bacteriophage T4 polynucleotide kinase from P. L. Biochemicals, and T4 ligase from Takara Shuzo Co.

2.2. Bacterial strains and bacteriophages

The following *E. coli* strains and phages were used: HO202*mal*^r (F⁻, *thi rel asnS*^{ts}), a *mal*^r derivative of HO202 [7], KY2562 (*thi tsx malA ompB101*) [8], JF703 (*ompF4 ilv-277 metB65 his-53 purE41 proC24 cyc-1 xyl-14 lacY29 rpsL77 tsx-63*) [9], λ *ompF1* [5], and λ C1857Sam7 (from H. Yamagata).

2.3. Preparation of DNA

The λ *ompF1* DNA was prepared as in [5]. *E. coli* HO202*mal*^r grown in L-broth at 30°C was used as the host strain and λ C1857Sam7 as the helper phage. After digestion with endonucleases, DNA fragments FR1, FR2 and FR3, shown in fig.1B, were prepared using 0.8% agarose gel as in [5]. Ligation of these fragments with cloning vector pBR322 was carried out as in [10]. The plasmids thus prepared were subjected to phenol extraction, recovered by ethanol precipitation, and used for transformation which was done as in [11]. Plasmids were prepared from the transformed cells as in [12] and digested with restriction endonucleases. The DNA fragments shown in fig.1C were prepared on 5% polyacrylamide gel as in [13].

2.4. DNA sequencing

All sequencing methods were according to [13]. The cleaved products were separated by using the

thin sequencing gel system (0.04 × 20 × 40 cm) with 20% and 10% polyacrylamide in 7 M urea.

3. Results

3.1. Cloning of the *ompF* gene on pBR322

In order to determine localization of the *ompF* gene on λ *ompF1*, DNA fragments FR1, FR2 and FR3, shown in fig.1B, were prepared and cloned in a

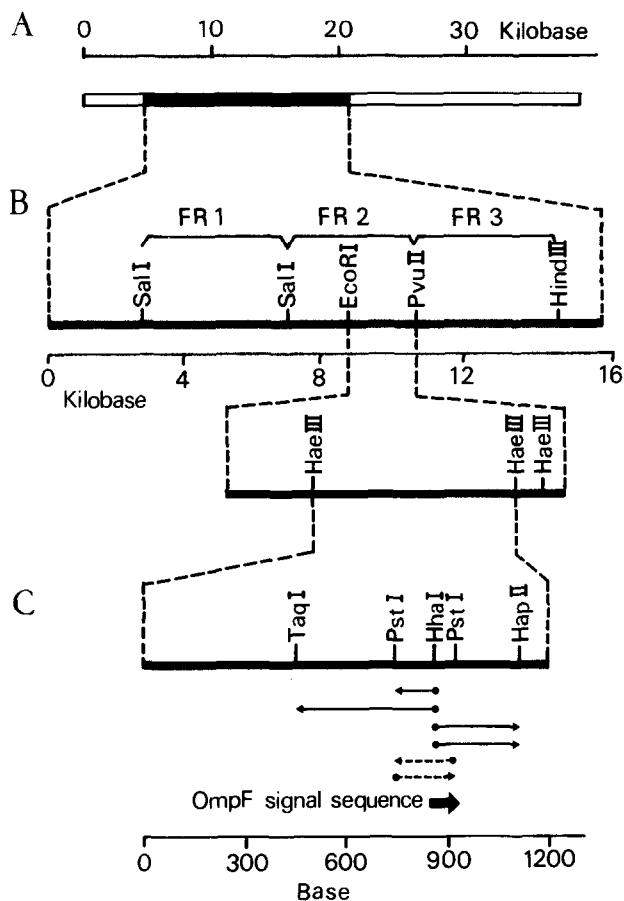


Fig.1. Restriction endonuclease cleavage sites and sequencing strategy around the *ompF* gene: (A) map of λ *ompF1* – the closed fragment represents the *E. coli* DNA carrying the *ompF* gene [5]; (B) fragments FR1 (4.3 kb), FR2 (3.8 kb) and FR3 (4.0 kb) prepared from the λ *ompF1* DNA are shown; (C) sequence strategy around the *ompF* gene. Only the restriction endonuclease cleavage sites used for the DNA sequencing analysis are shown. The singly labeled fragments were obtained either by secondary restriction enzyme cleavage indicated by arrows with a solid line or by strand separation indicated by arrows with a broken line. (●) The position of the ³²P-label at the 5'-end. The position of the OmpF signal sequence is also indicated.

high copy number plasmid, pBR322. *E. coli* JF703 was transformed with a recombinant plasmid carrying FR1 or FR3. However, the transformation did not result in production of OmpF or its precursor in the host cell. No transformant was obtained when JF703 was transformed with a recombinant plasmid carrying FR2. However, transformation with this plasmid was achieved with KY2562 which carried the *ompB* mutation that suppresses the OmpF synthesis. Since evidence has accumulated that overproduction of outer membrane proteins is most likely lethal to *E. coli* cells [14,15], we assumed that FR2 carries the *ompF* gene.

3.2. Restriction enzyme mapping and DNA sequencing

To find the *ompF* gene, FR2 was further fragmented with restriction endonucleases and a restriction map was made (fig.1C), and a DNA fragment which contained the DNA sequence corresponding to the amino acid sequence of the NH₂-terminal end of OmpF was searched for. The amino acid sequence of the NH₂-terminal end of OmpF has been determined [16]. We found the NH₂-terminal sequence of OmpF in the *HhaI*–*HapII* fragment (230 basepairs) shown in fig.1C. Then we determined the DNA sequence upstream of the NH₂-terminal sequence. The restriction enzyme fragments used for the sequence analysis are shown in fig.1C. The nucleotide sequence of 220 base pairs thus determined is shown in fig.2 together with the amino acid sequence deduced from it.

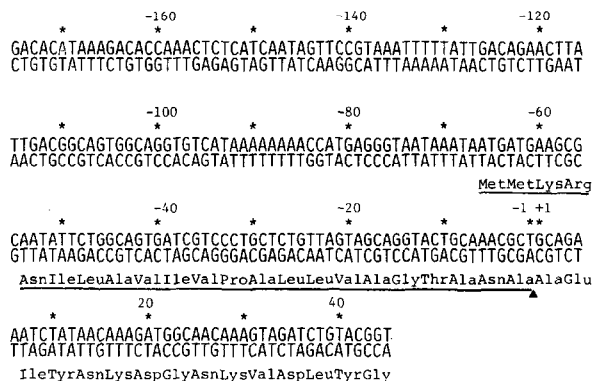


Fig.2. DNA sequence corresponding to the NH₂-terminal region of the OmpF precursor. The amino acid sequence deduced from the DNA sequence is also shown. The signal peptide sequence is underlined. The first nucleotide of the codon for the NH₂-terminal alanine residue of OmpF was numbered as +1. The cleavage site of the signal peptide is shown by an triangle.

4. Discussion

The DNA sequence from +1 to +27 (fig.2) can code for the amino acid sequence which is exactly the same as the NH₂-terminal sequence of OmpF of *E. coli* K-12 [16]. Furthermore, the amino acid sequence deduced from the DNA sequence from +1 to +45 was the same as the NH₂-terminal sequence of OmpF of *E. coli* B [17]. These facts indicate that FR2 contains a DNA fragment of ~580 base pairs that codes for the NH₂-terminal sequence of OmpF, and confirm that *ompF* is the structural gene. We would assume, therefore, that even the overproduction of the incomplete OmpF is most probably lethal to *E. coli* cells. The DNA sequence reveals the existence of a peptide extension at the NH₂-terminal end of OmpF, since no initiation codon was found around the NH₂-terminal alanine residue. The in-phase initiation codon can be found at the -66 to -64 position and the -63 to -61 position. This indicates that OmpF possesses a signal sequence of:

Met-Met-Lys-Arg-Asn-Ile-Leu-Ala-Val-
Ile-Val-Pro-Ala-Leu-Leu-Val-Ala-Gly-
Thr-Ala-Asn-Ala

as shown in fig.2. The view that ATG at the -66 to -64 position is the initiation codon was supported by the finding that this codon is preceded by ⁻⁸²ATGAGGGTAAT⁻⁷² and ⁻¹⁰³GCAGGTG⁻⁹⁷, both of which have a good homology with the Shine-Dalgarno sequence [18]. Furthermore, upstream of the sequences there were candidates for the Pribnow box (⁻⁹⁸TGTCATA⁻⁹² and ⁻¹³²TTTATTG⁻¹²⁶) [19] and the RNA polymerase recognition site (⁻¹³⁰TATTGACA⁻¹²³ and ⁻¹⁵⁵TCTCATCAAT⁻¹⁴⁶) [20]. The signal peptides so far studied have the following general characteristics:

- (i) Presence of a basic amino acid(s) near the NH₂-terminus;

| Outer Membrane Proteins | | | | | | | | Reference |
|----------------------------------|------|------|------|------|------|------|-------|------------|
| | -4 | -3 | -2 | -1 | ↓ | +1 | +2 | |
| Lipoprotein | -Leu | -Leu | -Ala | -Gly | -Cys | -Ser | -Ser- | [3] |
| Omp A | -Val | -Ala | -Gln | -Ala | -Ala | -Pro | -Lys- | [2] |
| Lam B | -Gln | -Ala | -Met | -Ala | -Val | -Asp | -Phe- | [4] |
| Omp F | -Thr | -Ala | -Asn | -Ala | -Ala | -Glu | -Ile- | This study |
| Periplasmic Proteins | | | | | | | | |
| Alkaline phosphatase | -Val | -Thr | -Lys | -Ala | -Arg | -Thr | -Pro- | [23] |
| Arabinose binding protein | -Gly | -Ala | -X | -Ala | -Glu | -Asn | -Leu- | [24] |
| Maltose binding protein | -Ser | -Ala | -Leu | -Ala | -Lys | -Ile | -Glu- | [25] |
| Leucine-specific binding protein | -Thr | -Ala | -Met | -Ala | -Asp | -Asp | -Ile- | [26] |
| β-lactamase | -Pro | -Val | -Phe | -Ala | -His | -Pro | -Glu- | [27] |

Fig.3. Amino acid sequences around the cleavage sites of precursors of outer membrane and periplasmic proteins of *E. coli*. The cleavage sites are indicated by an arrow. The NH₂-terminal residue of mature forms was numbered as +1.

- (ii) Presence of a long stretch of hydrophobic amino acid residues at the central region;
- (iii) Presence of a non-polar small amino acid residue at the cleavage site (-1 position) [21].

The amino acid sequence of the signal peptide of the OmpF precursor had all of these characteristics.

In addition to outer membrane proteins, periplasmic proteins are also proteins that have to be exported through the cytoplasmic membrane. However, no significant difference in the amino acid sequence of the signal peptide has been found between them, indicating that signal peptides do not determine localization of these proteins on the cell surface. A significant difference was found in the amino acid residue at the +1 position (fig.3). The residues at the +1 position of all of the outer membrane proteins are hydrophobic, while those of the periplasmic proteins are charged, that is, very hydrophilic. Although the amino acid residue of the lipoprotein at this position is cysteine, this residue has been modified with fatty acids before cleavage of the signal peptide [22], and hence should be very hydrophobic. The NH₂-terminus of OmpC, another major outer membrane protein, is also alanine [16]. However, the nature of the amino acid residues at the +2 and +3 positions is ambiguous; for example, some outer membrane proteins have a charged amino acid at the +2 position, while maltose-binding protein, a periplasmic protein, has isoleucine at this position. These facts strongly suggest that the property (hydrophobic/hydrophilic) of the +1 amino acid at the cleavage site plays an important role in the protein localization. At this moment, the role of the +1 amino acid residue is unclear. However, such a significant difference in the cleavage site would imply that the machinery for the signal peptide cleavage or its localization in the cell surface is different, which in turn determines the final destination of these proteins.

Acknowledgements

We thank T. Silhavy for informing us some of restriction cleavage sites around the promoter region of the *ompF* gene, and K. Nakamura and H. Yamagata for DNA sequencing technique and helpful discussion. This work was supported by grants from the Ministry of Education, Science and Culture of Japan and the Nisshin-Seifun Foundation.

References

- [1] Blobel, G. and Dobberstein, B. (1975) *J. Cell Biol.* 67, 835-851.
- [2] Movva, N. R., Nakamura, K. and Inouye, M. (1980) *J. Biol. Chem.* 255, 27-29.
- [3] Inouye, S., Wang, S. S., Sekizawa, J., Halegoua, S. and Inouye, M. (1977) *Proc. Natl. Acad. Sci. USA* 74, 1004-1008.
- [4] Hedgpeth, J., Clément, J.-M., Marchal, C., Perrin, D. and Hofnung, M. (1980) *Proc. Natl. Acad. Sci. USA* 77, 2621-2625.
- [5] Mutoh, N., Nagasawa, T. and Mizushima, S. (1981) *J. Bacteriol.* 145, 1085-1090.
- [6] Johnson, R. A. and Walseth, T. F. (1979) *Adv. Cyclic Nucl. Res.* 10, 135-167.
- [7] Yamamoto, M., Nomura, M., Ohsawa, H. and Maruo, B. (1977) *J. Bacteriol.* 132, 127-131.
- [8] Sato, T. and Yura, T. (1981) *J. Bacteriol.* 145, 88-96.
- [9] Chai, T. and Foulds, J. (1977) *J. Bacteriol.* 130, 781-786.
- [10] Weiss, B. and Richardson, C. C. (1967) *Proc. Natl. Acad. Sci. USA* 57, 1021-1025.
- [11] Dagert, M. and Ehrlich, S. D. (1979) *Gene* 6, 23-28.
- [12] Matsubara, K. (1975) in: *Seikagaku Jikken Koza* (Japanese Biochem. Soc. edn) vol. 2, *Nucleic Acid Chemistry I*, pp. 67-70, Tokyo Kagaku Dojin, Tokyo.
- [13] Maxam, A. M. and Gilbert, W. (1980) *Methods Enzymol.* 65, 499-560.
- [14] Henning, U., Royer, H.-D., Teather, R. M., Hindennach, I. and Hollenberg, C. P. (1979) *Proc. Natl. Acad. Sci. USA* 76, 4360-4364.
- [15] Lee, N., Nakamura, K. and Inouye, M. (1981) *J. Bacteriol.* 146, 861-866.
- [16] Ichihara, S. and Mizushima, S. (1978) *J. Biochem. (Tokyo)* 83, 1095-1100.
- [17] Chen, R., Krämer, C., Schmidmayr, W. and Henning, U. (1979) *Proc. Natl. Acad. Sci. USA* 76, 5014-5017.
- [18] Shine, J. and Dalgarno, L. (1974) *Proc. Natl. Acad. Sci. USA* 71, 1342-1346.
- [19] Pribnow, D. (1975) *Proc. Natl. Acad. Sci. USA* 72, 784-788.
- [20] Dickson, R. C., Abelson, J., Barnes, W. N. and Reznikoff, W. S. (1975) *Science* 187, 27-35.
- [21] Halegoua, S. and Inouye, M. (1980) in: *Bacterial Outer Membrane* (Inouye, M. ed) pp. 67-114, Wiley Interscience, New York.
- [22] Hussain, M., Ichihara, S. and Mizushima, S. (1980) *J. Biol. Chem.* 255, 3707-3712.
- [23] Kikuchi, Y., Yoda, K., Yamasaki, M. and Tamura, G. (1981) *Agric. Biol. Chem.* 45, 2401-2402.
- [24] Wilson, V. G. and Hogg, R. W. (1980) *J. Biol. Chem.* 255, 6745-6750.
- [25] Bedouelle, H., Bassford, P. J., Fowler, A. V., Zabin, I., Beckwith, J. and Hofnung, M. (1980) *Nature* 285, 78-81.
- [26] Oxender, D., Anderson, J. J., Daniels, C. J., Landick, R., Gunsalus, R. P., Zurawski, G. and Yanofsky, C. (1980) *Proc. Natl. Acad. Sci. USA* 77, 2005-2009.
- [27] Sutcliffe, J. G. (1978) *Proc. Natl. Acad. Sci. USA* 75, 3737-3741.