Vol. 173, No. 1, 1990 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
November 30, 1990 Pages 53-59

PRIMARY STRUCTURE AND EXPRESSION OF THE Ssc-PROTEIN OF *SALMONELLA TYPHIMURIUM*¶

Laura Hirvas, Pertti Koski and Martti Vaara*

**Department of Bacteriology and Immunology,
University of Helsinki, 00290 Helsinki, Finland**

Received October 10, 1990

SUMMARY: A 1020-bp open reading frame (ORF) was found immediately downstream of the *ompH* gene of *Salmonella typhimurium*. This ORF (ORF-36) encodes a moderately hydrophobic protein with 341 amino acid residues (calculated molecular mass, 35,928 Da). The ORF-36 product was detected in minicells. Downstream of ORF-36, another ORF was found. It is highly homologous to the *E. coli* ORF (ORF-17.4) which precedes the *lpx*-genes involved in lipid A biosynthesis. ORF-36 is probably analogous to the *firA* gene of *E. coli*, the sequence of which has not yet been published. Thus it appears that the enterobacterial *ompH* and *lpx* genes are separated only by the ORF-36 and ORF-17.4 genes. We also discuss the data on the function of the ORF-36 protein. On this basis, we suggest that the protein could be called the Ssc protein.

Gram-negative enteric bacteria are known to be remarkably resistant to hydrophobic drugs because their outermost cell structure, the outer membrane, is an effective permeability barrier against them (1,2). In 1984, we described an antibiotic supersensitive mutant of *Salmonella typhimurium* (the mutant SH7622) which carries the SS-C mutation (SS for supersensitivity)(3). The supersensitivity to many unrelated hydrophobic antibiotics, drugs, and detergents indicated that the mutant has a defective permeability barrier (2,3). Despite strenuous research, no molecular basis for the defect was found (2-5).

Recently, when working with the novel *Salmonella* outer membrane protein OmpH and its gene, which we have isolated, cloned, and sequenced (6,7), we found that a plasmid (pUCHS16) harboring the *ompH* gene and its immediate 1 kb-long downstream sequence, completely reverted the antibiotic sensitive phenotype of SH7622 (8). Separate controls indicated that the plasmid carrier itself did not revert the phenotype, Furthermore, pUCHS16 did not revert the phenotype of supersensitive mutants unrelated to SS-C.

In this communication, we report the discovery of a protein-encoding region (ORF-36) immediately downstream of the *ompH* gene of *S. typhimurium*. This region is completely included in pUCHS16, is expressed in minicells, and is apparently analogous to the *firA* gene

¶ GenBank Accession Number M35193.
* Corresponding author.

0006-291X/90 $1.50

of *E. coli* (9), the sequence of which has not yet been published. Whether the ORF-36 product is responsible for the reversion of the phenotype of the SS-C mutant SH7622, will be discussed.

MATERIALS AND METHODS

Plasmids. The plasmids pUCHS14 (7) and pUCHS16 (6) have been described earlier.

DNA sequencing. Sequencing was done from denatured plasmids as in (10) by the dideoxyribonucleotide chain termination method (11) with Sequenase (United States Biochemical, Cleveland, OH) and [α-^{35}S] dATP. Both the pUC universal primers and specific synthetic oligodeoxyribonucleotides were used as primers. Occasional problems arising from compressions in the gel were overcome by substituting 7-deazaguanosine 5'-triphosphate (12) for dGTP.

Expression in minicells. The plasmid-encoded proteins were studied using *E. coli* PK251 minicells (7). Minicells were isolated by the method of Dougan and Kehoe (13). Proteins were labeled with [^{35}S]methionine (Amersham; specific activity, 1 Ci/ml; final concentration, 150 μCi/ml) in conditions described earlier (7) and boiled for 10 min in the presence of 36% sodium deoxycholate (SDS) and 10% mercaptoethanol. SDS polyacrylamide gel electrophoresis (15% polyacrylamide, ref. 14) was followed by soaking in Amplify solution (Amersham) for 30 min before drying and autoradiography.

Computer analyses. Analysis of the nucleotide and amino acid sequence data were performed using the HIBIO DNASIS and PROSIS computer program packages (Pharmacia, Bromma, Sweden). The hydrophobicity profile was calculated according to Kyte and Doolittle (15) with the window width of six amino acids. The secondary structure was predicted according to Chou, Fasman and Rose with the PROSIS program.

RESULTS AND DISCUSSION

The region downstream of *ompH* was sequenced by using plasmids pUCHS16 and pUCHS14. pUCHS16 carries the 2.3 kb *BamHI-SalI* fragment of *S. typhimurium* chromosome and contains, besides the entire *ompH* gene, also a 1550 bp-long downstream region (7). pUCHS14 carries the 1,6 kb *PstI-PstI* fragment and contains most of the structural *ompH* gene and the first 1250 bp of its downstream region (6). The sequencing strategy is shown in Fig. 1.

Two open reading frames were found (Fig. 2). The first, ORF-36 (also called as *ssc*, see below), contains 1020 bases from the initiation codon ATG to the stop codon TAA. This corresponds to a protein with 341 amino acid residues (calculated molecular mass, 35,928 Da). There are only three nucleotides between the stop codon of *ompH* and the initiation codon of ORF-36. ORF-36 is preceded 10 bases upstream by a typical Shine-Dalgarno sequence GAAACAGGT. A putative promoter region was also found. The sequences TTGAGA and TCCAAC have a spacing of 17 bp and share homology (5/6 and 3/6, respectively) to the canonical (16) -35 region (TTGACA) and -10 region (TATAAT) of *E. coli* promoters. This putative promoter overlaps with the protein coding region of *ompH*. Comparison of the nucleotide sequence of ORF-36 with the sequences deposited in EMBL databank or in GenBank revealed no significant homologies.

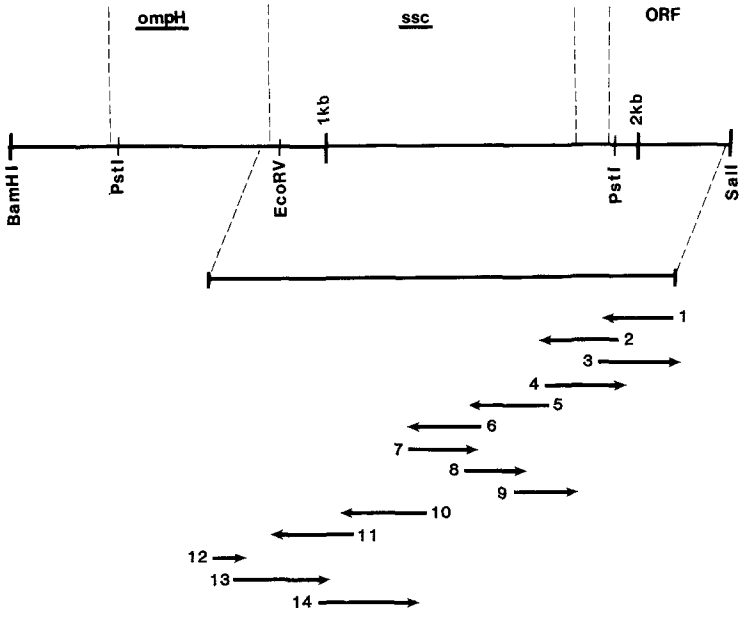

Figure 1. Sequencing the region downstream of *ompH* of *S. typhimurium*. Small arrows represent sequenced fragments. Fragments 1-4 were sequenced from pUCHS16 (7), fragments 5-7, 9, 10, 12, and 14 from pUCHS14 (6), and fragments 8, 11, and 13 from both plasmids. Fragments 1 and 5 were sequenced by using pUC universal primers and the others by using synthetic 17-mer oligonucleotide primers. Abbreviation: ORF, open reading frame. *ssc* stands for the found ORF-36 (see text).

Another reading frame lies downstream of ORF-36 and is separated from it by a 105 bp-long spacing region (Fig. 2). This ORF is 90.3% homologous with ORF-17.4 of *E. coli* which is known to be located immediately upstream of the *lpx* genes (*lpxA*, *lpxB*) involved in lipid A biosynthesis (9,17).

The expression of ORF-36 in plasmid pUCHS16 was studied in *E. coli* minicells (Fig. 3). The plasmid encoded, besides ß-lactamase (30 kDa; precursor, 32 kDa) and the OmpH protein (16 kDa; precursor, 18 kDa), a third protein with the apparent molecular mass of 36 kDa. The band for this protein was weaker than the bands corresponding to ß-lactamase and OmpH. No precursor form was detected. We believe that this 36 kDa protein is the ORF-36 gene product.

As deduced from the nucleotide sequence, the ORF-36 product lacks an amino-terminal signal sequence. The protein contains 45.2% hydrophobic and 16.7% charged residues (26 basic, 31 acidic residues). The charged residues are distributed throughout the protein. The hydrophobicity profile indicates that the protein is slightly hydrophobic (mean hydrophobicity index, 0.15) and has numerous (more than ten) short, approx. 10 amino acid-long hydrophobic stretches (Fig. 4). The secondary structure prediction reveals that the protein is rich in ß-sheet (51%, data not shown). Moreover, most of the hydrophobic stretches have the ß-sheet structure. Since ß-sheet can cross the thickness of a membrane

```
ompH      1  GCGCAGGCTTTTGAGAAAGATCGCGCTCGTCGTTCCAACGAAGAACGCAACAAACTGGTG
             AlaGlnAlaPheGluLysAspArgAlaArgArgSerAsnGluGluArgAsnLysLeuVal

            61  ACTCGTATCCAGACTGCGGTGAAAAAAGTGGCTAACGACCAGAGTATCGATCTGGTGGTA
             ThrArgIleGlnThrAlaValLysLysValAlaAsnAspGlnSerIleAspLeuValVal

           121  GACGCAAACACCGTTGCTTACAACAGCAGCGATGTGAAAGACATCACCGCTGACGTACTG
             AspAlaAsnThrValAlaTyrAsnSerSerAspValLysAspIlrThrAlaAspValLeu

ORF-36     181  AAACAGGTTAAATAAGTAATGCCTTCAATTCGACTGGCTGACTTAGCAGAACAGTTGGAT
             LysGlnValLys***   MetProSerIleArgLeuAlaAspLeuAlaGluGlnLeuAsp

           241  GCAGAATTACACGGTGATGGCGATATCGTCATCACCGGCGTTGCGTCCATGCAATCTGCA
             AlaGluLeuHisGlyAspGlyAspIleValIleThrGlyValAlaSerMetGlnSerAla

           301  ACAACAGGCCACATTACGTTTATGGTGAATCCTAAGTACCGTGAACACTTAGGTTTATGC
             ThrThrGlyHisIleThrPheMetValAsnProLysTyrArgGluHisLeuGlyLeuCys

           361  CAGGCTTCTGCGGTTGTCATGACGCAGGACGATCTTCCTTTTGCTAAGAGTGCGGCGCTG
             GlnAlaSerAlaValValMetThrGlnAspAspLeuProPheAlaLysSerAlaAlaLeu

           421  GTAGTTAAAAATCCCTACCTGACCTACGCGCGCATGGCGCAAATTTTAGATACTACGCCG
             ValValLysAsnProTyrLeuThrTyrAlaArgMetAlaGlnIleLeuAspThrThrPro

           481  CAGCCCGCGCAGAATATCGCGCCAAGCGCCGTGATTGATGCGACGGCAACGCTGGGTAGC
             GlnProAlaGlnAsnIleAlaProSerAlaValIleAspAlaThrAlaThrLeuGlySer

           541  AATGTTTCAGTCGGCGCGAATGCGGTGATTGAATCTGGCGTACAACTGGGCGATAACGTG
             AsnValSerValGlyAlaAsnAlaValIleGluSerGlyValGlnLeuGlyAspAsnVal

           601  GTTATCGGCGCAGGCTGTTTCGTCGGAAAAAATAGCAAAATCGGGGCGGGTTCACGCTTG
             ValIleGlyAlaGlyCysPheValGlyLysAsnSerLysIleGlyAlaGlySerArgLeu

           661  TGGGCGAACGTAACGATTTACCACGACATTCAGATCGGTGAGAATTGCCTGATCCAGTCC
             TrpAlaAsnValThrIleTyrHisAspIleGlnIleGlyGluAsnCysLeuIleGlnSer

           721  AGTACGGTGATCGGCGCGGACGGTTTTGGCTACGCTAACGATCGTGGCAACTGGGTGAAG
             SerThrValIleGlyAlaAspGlyPheGlyTyrAlaAsnAspArgGlyAsnTrpValLys

           781  ATCCCACAACTGGGCCGGGTCATTATTGGCGATCGTGTCGAGATCGGCGCTTGTACCACC
             IleProGlnLeuGlyArgValIleIleGlyAspArgValGluIleGlyAlaCysThrThr

           841  ATTGACCGTGGCGCGTTGGATGATACTGTTATTGGCAATGGCGTGATTATTGATAATCAG
             IleAspArgGlyAlaLeuAspAspThrValIleGlyAsnGlyValIleIleAspAsnGln

           901  TGCCAGATTGCACATAACGTCGTGATTGGCGACAATACGGCAGTTGCCGGTGGCGTCATT
             CysGlnIleAlaHisAsnValValIleGlyAspAsnThrAlaValAlaGlyGlyValIle

           961  ATGGCGGGTAGCCTGAAGATTGGCCGTTACTGCATGATTGGCGGCGCCAGCGTGATCAAT
             MetAlaGlySerLeuLysIleGlyArgTyrCysMetIleGlyGlyAlaSerValIleAsn

          1021  GGGCATATGGAAATATGCGACAAAGTCACGGTAACTGGCATGGGTATGGTGATGCGTCCC
             GlyHisMetGluIleCysAspLysValThrValThrGlyMetGlyMetValMetArgPro

          1081  ATCACGGAACCGGGCGTCTACTCCTCAGGCATTCCGCTGCAACCCAACAAAGTATGGCGT
             IleThrGluProGlyValTyrSerSerGlyIleProLeuGlnProAsnLysValTrpArg

          1141  AAAACTGCTGCACTGGTGATGAACATTGATGATATGAGCAAGCGTCTCAAAGCGATTGAG
             LysThrAlaAlaLeuValMetAsnIleAspAspMetSerLysArgLeuLysAlaIleGlu

          1201  CGCAAGGTTAATCAACAAGACTAACGTTCCGCCTTGTAGTTGCCATTCTTTTCCGGCCTG
             ArgLysValAsnGlnGlnAsp***

          1261  TCACATTCATACGATTGCGGCAGGCCGTGTTATTATTGCCTTTTTGTATATTTGGACAGG

ORF-II     1321  AAGAGTATTTTGACTACTAACACTCATACTCTGCAGATTGAAGAGATTTTAGAGCTTCTG
                       MetThrThrAsnThrHisThrLeuGlnIleGluGluIleLeuGluLeuLeu

          1381  CCGCACCGTTTTCCGTTTTTACTGGTCGATCGCGTGCTGGACTTTGAAGAAGGTCGTTTT
             ProHisArgPheProPheLeuLeuValAspArgValLeuAspPheGluGluGlyArgPhe

          1441  CTGCGTGCGGTGAAAAATGTCTCCGTCAACGAGCCGTTTTTCCAGGGGCATTTCCCGGGC
             LeuArgAlaValLysAsnValSerValAsnGluProPhePheGlnGlyHisPheProGly

          1501  AAACCGATTTTGCCAGGCGTGCTGATTCTGGAAGCGATGGCGCAGGCAACCGGTATTCTG
             LysProIleLeuProGlyValLeuIleLeuGluAlaMetAlaGlnAlaThrGlyIleLeu

          1561  GCGTTTAAAAGCGTTGGTAAACTGGAACCTGGCGAACTGTATTATTTCGCGGGTATTGAT
             AlaPheLysSerValGlyLysLeuGluProGlyGluLeuTyrTyrPheAlaGlyIleAsp

          1621  GAAGCGCGCTTTAAGCGTCCGGTGGTGCCAGGCGATCAGATGATCATGGAAGTCACTTTC
             GluAlaArgPheLysArgProValValProGlyAspGlnMetIleMetGluValThrPhe

          1681  GAGAAAACGCGCCGTGGCCTGACCCGCTTTAAAGGGGTTGCGCTGGTCGAC
             GluLysThrArgArgGlyLeuThrArgPheLysGlyValAlaLeuValAsp
```

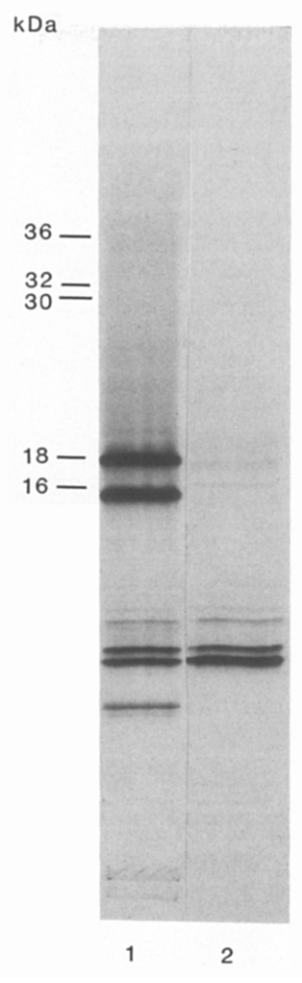

Figure 3. Autoradiography of proteins synthesized in *E. coli* minicells harboring the control plasmid pUC19 (lane 1) or pUCHS16 (lane 2). The control plasmid elaborates ß-lactamase (30 kDa) and its precursor (32 kDa). pUCHS16 encodes three additional proteins, the OmpH protein (16 kDa), its precursor (18 kDa), and the 36-kDa protein.

bilayer in 10 to 12 residues, the ORF-36 product could be a membrane protein and span the membrane repeated times.

The now found ORF-36 is probably analogous to the *E. coli* gene *firA*. According to recent data (see 9,18) and in contrast to previous reports (19, see also 20), *firA* is a gene distinct from *hlpA* (also named as *skp*) and does not code for the 17 kDa HIpI (Skp) protein which is

Figure 2. Nucleotide sequence of the 3′ terminal and downstream region of *ompH* of *S. typhimurium*. Nucleotides 1-195 constitute the 3′ terminus of *ompH* and we have sequenced them previously (6). In addition, we have preliminary sequenced nucleotides 196-389 (6), the sequence now shown includes two corrections to that earlier sequence (see positions 204 and 296). Downstream of *ompH*, two open reading frames (ORF-36, ORF-II) are shown, their initiation codons are underlined. Underlined are also the putative Shine-Dalgarno sequence (positions 180-188) and the putative promoter -35 and -10 regions (positions 11-16 and 34-39, respectively) of the ORF-36 gene. The deduced amino acid sequences are given below the nucleotide sequence.

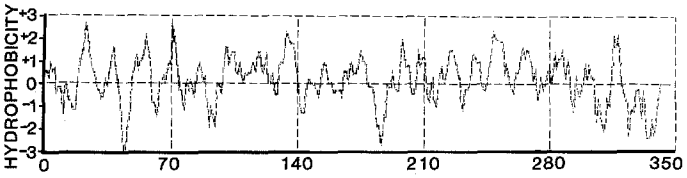

Figure 4. Hydrophobicity-hydrophilicity plot of the ORF-36 protein. Hydrophobicity is indicated by a positive value of the ordinate, and hydrophilicity is indicated by a negative value.

the *E. coli* analogue of the OmpH protein of *S. typhimurium* (21). *firA* is known to be located between *hlpA* and ORF17.4. Because the homology of the *hlpA* gene of *E. coli* and the corresponding *ompH* gene of *S. typhimurium* is 87% (21) and because the ORF-17.4 regions are 90% homologous in these species (this paper), it can be expected that the *firA* and ORF-36 genes might be considerably homologous, too.

We have not yet elucidated the function of the ORF-36 product. The close location of ORF-36 to *lpxA* and *lpxB* might suggest that it is involved in LPS biosynthesis. A direct approach to understand the function is to construct mutants which produce defective ORF-36 product. Fortunately, such a mutant is already available. As reviewed in Introduction, the antibiotic supersensitive phenotype of our *S. typhimurium* mutant SH7622 which carries the SS-C mutation is completely reversed by the plasmid pUCHS16. As shown in the present paper, this plasmid contains only two *Salmonella* genes, *ompH* and the ORF-36 gene. Data to be presented elsewhere indicates that SH7622 has a point mutation in ORF-36 and that also other mutations in ORF-36 affect the antibiotic sensitive phenotype of *S. typhimurium* (P, Koski, L. Hirvas, and M. Vaara, manuscript in preparation). Further work is underway in our laboratory. We suggest that the ORF-36 gene product will be called as the Ssc protein until its function is better known.

ACKNOWLEDGMENTS

We thank Birgit Kuusela for her excellent technical assistance. This investigation was supported by Grant 01/690 from the Academy of Finland (to M.V.) and by Sigrid Juselius Foundation (to M.V. and P.K.)

REFERENCES

1. Nikaido, H., and Vaara, M. (1985) Microbiol. Rev. 49,1-32.
2. Sukupolvi, S., and Vaara, M. (1989) Biochim. Biophys. Acta 988,377-387.
3. Sukupolvi, S., Vaara, M., Helander, I., Viljanen, P., and Mäkelä, P.H. (1984) J. Bacteriol. 159,704-712.
4. Sukupolvi, S., Helander, I., Hukari, R., Vaara, M., and Mäkelä, P.H. (1985) FEMS Microbiol. Lett. 30,341-345.
5. Vaara, M., Viljanen, P., Sukupolvi, S., and Vaara, T. (1985) FEMS Microbiol. Lett. 26,289-294.
6. Koski, P., Rhen, M., Kantele, J., and Vaara, M. (1989). J. Biol. Chem. 264,18973-18980.
7. Koski, P., Hirvas, L., and Vaara, M. (1990) Gene 88,117-120.

8. Vaara, M., Hirvas, L., and Koski, P. (1990) In Endotoxin Research (J. Spitzer,Ed.), Vol. I. Elsevier, Amsterdam, in press.
9. Bachmann, B.J. (1990) Microbiol. Rev. 54,130-197.
10. Hattori, M., and Sakaki, Y. (1986) Anal. Biochem. 152,232-238.
11. Sanger, F., Nicklen, S., and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74,5463-5467.
12. Mizusawa, S., Nishimura, S., and Freela, F. (1986) Nucl. Acids Res. 14,1319-1324.
13. Dougan, G., and Kehoe, M. (1984) Meth. Microbiol. 7,233-258.
14. Laemmli, U.K. (1970) Nature (L.) 227,680-685.
15. Kyte, J., and Doolittle, R.F. (1982) J. Mol. Biol. 157,105-132.
16. O'Neill, M.C. (1989) J. Biol. Chem. 264,5522-5530.
17. Coleman, J., and Raetz, C.R.H. (1988) J. Bacteriol. 170,1268-1274.
18. Thome, B.M., Hoffschulte, H.K., Schiltz, E., and Müller, M. (1990) FEBS Lett. 269,113-116.
19. Aasland, R., Coleman, J., Holck, A.L., Smith, C.L., Raetz, C.R.H., and Kleppe, K. (1988) J. Bacteriol. 170,1268-1274.
20. Kröger, M., Wahl, R., and Rice, P. (1990) Nucl. Acids Res. 18, Suppl. 2549-2587.
21. Hirvas, L., Coleman, J., Koski, P., and Vaara, M. (1990) FEBS Lett. 262,123-126.