Volume 15 Number 3 1987 Nucleic Acids Research

Structure of the gene for the stringent starvation protein of *Escherichia coli*

Hiroaki Serizawa and Ryuji Fukuda*

Department of Molecular Genetics, National Institute of Genetics, Mishima, Shizuoka 411, Japan

Received November 19, 1986; Accepted December 30, 1986

ABSTRACT

The nucleotide sequence of the gene for the stringent starvation protein (SSP) of E. coli was determined. The deduced amino acid sequence shows that the SSP is composed of 212 amino acid residues, rich in both positively and negatively charged amino acids and has a molecular weight of 24,305. Primer extension experiments and nuclease S1 mapping analysis showed a site on the chromosome DNA corresponding to the 5' end of the transcript of the SSP gene. However, the consensus promoter sequences were not found at the upstream region. In the 3' flanking region a long coding frame was found immediately following the SSP gene, suggesting that the SSP gene is a member of a multicistronic operon.

INTRODUCTION

It is considered that a set of transcription factors participate in the control of transcription by interacting with RNA polymerase (1). Up to ten polypeptides have been identified, which associate with RNA polymerase and may modulate the functions of RNA polymerase during transcription (2). Stringent starvation protein (SSP) is one such polypeptide. It forms an equimolar complex with the RNA polymerase holoenzyme, but not with the core enzyme (3). The protein is unique in that it is synthesized predominantly, occupying more than 50% of the total protein synthesis, when cells are exposed to amino acid starvation (4). To study the physiological functions of SSP, we cloned the DNA fragments containing the gene coding for this protein (5). For the cloning, we first determined the partial amino acid sequences of the protein, and then chemically synthesized the oligonucleotide mixtures that represented possible codon combinations for parts of these amino acid sequences. Using the labeled oligonucleotide mixtures as the probe, clones containing the gene for SSP were screened employing a colony hybridization technique. Using one of these clones, pSS-1, which is pBR322 inserting 5 kbp HindIII fragment of E. coli chromosome and carrying ampicilline-resistant gene, the gene for the SSP was mapped to the region

between gltB and glnF at min 69.5 on the E. coli chromosome (Fukuda, R., Nishimura, A. and Serizawa, H., manuscript in preparation). It was thus indicated that the SSP gene is a hitherto unknown gene.

In this study, we have determined the DNA sequence of 1,616 bp, which contains the complete coding region of SSP and the flanking regions. We also determined a transcription start site employing both nuclease S1 mapping and primer extension analyses.

MATERIALS AND METHODS

Preparation of DNA fragments

Plasmid DNAs were prepared by the alkaline lysis procedure (6), purified by Sepharose 4B column chromatography, and digested with appropriate restriction enzymes. Restriction fragments were separated by electrophoresis on polyacrylamide gels, and recovered from gels by the diffusion method (7).

Cloning of the restriction fragments in M13 DNA

Fig. 1A shows the restriction map of the 1.75 kbp SalI fragment for enzymes which were used in subcloning in M13 DNAs. The SalI fragment contained the complete structural gene for SSP (5). The restriction fragments shown by filled bars in Fig. 1A were ligated to restriction enzyme-cleaved M13 mp10 and mp11 replicative form I DNAs, after adding proper linkers to the repaired ends of the fragment if necessary (8).

DNA Sequence Analysis

DNA sequence analysis was performed by the dideoxy chain termination method (9) using single strand DNA prepared from clones of M13 mp10 and mp11 phages, a synthetic 15 base universal primer, and [α-^{32}P]dCTP as a radiolabel. The reaction products were analysed by electrophoresis through 0.35 mm polyacrylamide gels (in 0.1 M Tris borate (pH 8.3), 2 mM EDTA and 8 M urea). The gels were fixed with 10% acetic acid and 10% methanol, dried and exposed to Fuji RX X-ray films at -80°C.

Amino Acid Analysis

Amino acid composition of SSP was determined as described previously (10).

Nuclease S1 Mapping of in vivo SSP Transcripts

Cellular RNA was prepared from K802 cells which were grown in L-broth and harvested at late-log phase of growth. The cells were rapidly chilled to 0°C immediately after addition of chloramphenicol (50 µg/ml), and RNA was isolated by repeated extraction with phenol at 64°C as described (11). DNA restriction fragments used as the probe are shown in Fig. 4A, b. One 5'

terminus of the probes was uniquely labeled with [γ-^{32}P]ATP using T4 polynucleotide kinase. The ^{32}P-probe fragments thus prepared (1.7×10^6 cpm/pmol) and RNA (approximately, 150 µg) were dissolved in 30 µl of hybridization buffer (0.4 M NaCl, 0.02 M PIPES (pH 6.5), 80% formamide)(12). The solution was incubated at 75°C for 10 min to denature duplex DNA and RNA and then cooled down gradually to 30°C during about 2 h. It was incubated at 30°C for further 2 h, and mixed with 268 µl of pre-warmed S1 nuclease buffer (0.03 M NaOAc (pH 4.6), 0.05 M NaCl, 1 mM $ZnSO_4$ and 5% glycerol). After addition of various units of nuclease S1, the mixture was incubated at 30°C for 15 min, followed by phenol extraction and ethanol precipitation. RNA in the mixture was digested with DNase-free RNaseA (50 µg/ml) at 37°C for 15 min. The protected DNA fragments were analysed by electrophoresis through 8% polyacrylamide sequencing gels in parallel with Maxam-Gilbert sequencing ladders (7) of the original probe DNA.

Primer Extension Reaction

Cellular RNA (150 µg) was hybridized with a primer DNA fragment by the same procedure as described above for nuclease S1 mapping. The primer DNA fragments shown in Fig. 4, c and d was uniquely labeled at the 5' ends (specific activity, 1.7×10^6 and 9×10^6 cpm/pmol, respectively). RNA-primer hybrids were precipitated with ethanol, washed with 70% ethanol, dried and dissolved in 20 µl of reverse transcriptase buffer [0.05 M Tris-HCl (pH 8.0), 0.01 M $MgCl_2$, 0.06 M NaCl, 1.5 mM each of 4dNTP and 5 mM DTT]. After addition of 10 units AMV reverse transcriptase (provided by Dr. A. Ishihama), the reaction was carried out at 37°C for 90 min. The reverse transcripts were analysed by electrophoresis through polyacrylamide sequencing gels as described above for nuclease S1 mapping analysis.

Enzymes and Biochemicals

Restriction endonucleases were purchased from Takara Shuzo Co., Japan and New England Biolabs, USA. Nuclease S1 was the product of PL Biochemicals Inc., USA. DNA sequencing kits for the dideoxy method, phosphorylated linkers and other enzymes used in DNA manipulations were obtained from Takara Shuzo Co., Japan. [γ-^{32}P]ATP (>5,000 Ci/mmol) and [α-^{32}P]dCTP (>400 Ci/mmol) were purchased from the Radiochemical Centre, Amersham, England.

RESULTS AND DISCUSSION

Nucleotide sequence and the predicted amino acid sequence of SSP

In the previous report(5), we indicated that the structural gene for SSP is contained in the 1.75 kbp SalI fragment of the plasmid pSS-1, and that its

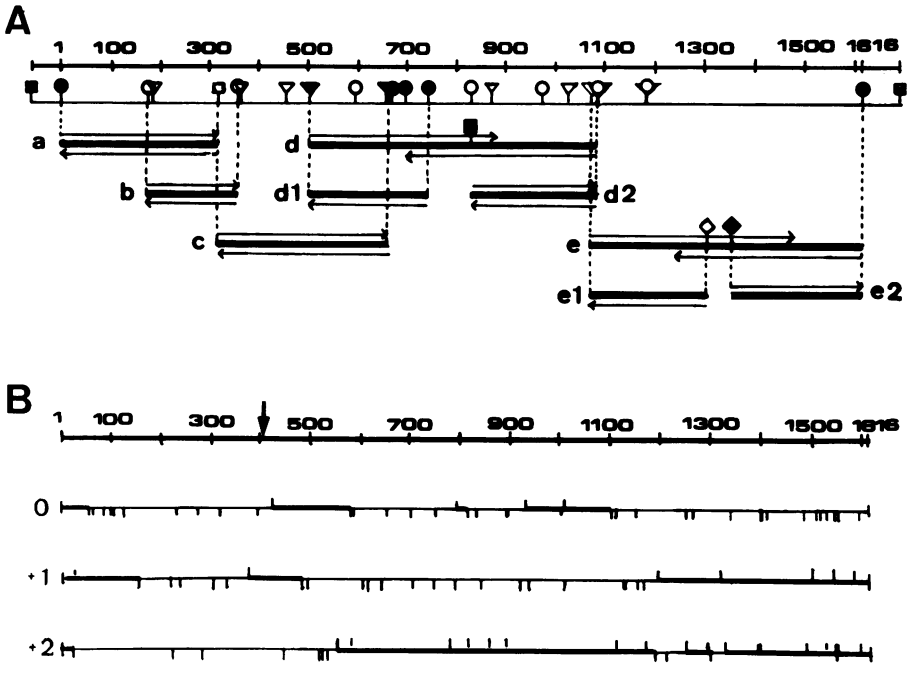

Fig. 1. Sequence strategy for the SSP gene (A), and ATG codons and termination codons in all three reading frames (B). (A) Nucleotides of the 1.75 kb SalI fragment are numbered from the first base of the left-most Sau3Al site to the last base of the right-most Sau3Al site (upper thick line). The restriction sites for endonuclease used in the subcloning are shown in the middle; SalI (⊠), Sau3Al (●), HapII (○), HincII (□), HaeIII (▼), TaqI (▽), NciI (■), FokI (◇) and RsaI (◆). The restriction fragments (filled bars) were isolated and cloned in M13 phages. Nucleotide sequences were determined (as shown by arrow lines). (B) ATG codons (upward vertical lines) and termination codons (downward vertical lines) are shown for all three reading frames. Open reading frames starting with ATG and ending with a termination codon are shown by filled bars. The vertical arrow indicates the 5' end of transcript(s) of the SSP gene (see text below).

structural gene starts at about 500 bases from the left SalI site and continues rightward (see Fig. 1A)(5). Fig. 2 shows the nucleotide sequence of 1,616 bp from the left-most to the right-most Sau3Al sites on the SalI fragment. AUG codons and termination codons in all three reading frames are shown by upward and downward vertical lines, respectively, in Fig. 1B. In (+2) frame we can find the N-terminal amino acid sequence (from Ala^2 to Arg^8) of SSP which was determined directly on the purified protein (5) (underlined in Fig. 2). The other two amino acid sequences of the cyanogen bromide-

```
        10         20         30         40         50         60         70         80         90
GATCCGTCACGGTATCACCCGCGCTCTGATGGAATACGACGAGTCCCTGCGTTCTGAACTGCGTAAAGCTGGCTTCGTTACTCGTGACGC

       100        110        120        130        140        150        160        170        180
TCGTCAGGTTGAACGTAAGAAAGTCGGTCTGCGTAAAGCACGTCGTCGTCCGCAGTTCTCCAAACGTTAATTGGCTTCTGCTCCGGCAGA

       190        200        210        220        230        240        250        260        270
AAACAATTTTCGAAAAAACCCGCTTCGGCGGGTTTTTTTATAGCTAAAATCTGAATCAGCGTAAAAACTGGAAAGTTGCTTTTTGCTGCC

       280        290        300        310        320        330        340        350        360
ACCTGACAGACAGGTAAAACAAAAACCACATCGCCAATAAGGGACTAAGTCAACTATTTCAGACTAAAGCGCATCTCTTTTTCCCCATTT

       370        380        390        400        410        420        430        440        450
CCGGCATCGACTCACCACAATGGTCGCAAAATCTGGTAAACTATCATCCAATTTTCTGCCCAAATGTCGGGTATTGCTCATTTTTTGTTT

       460        470        480        490        500        510        520        530        540
GATTTTCGAACAAAGAGAGTCGTTCCTTATTGGGTAACACAACTTCTGACTGGCCACCTGGTGGCTGGTAGCAGTAAAAATTCTGACTAT

       550        560        570        580        590        600        610        620        630
ACCTGGAGGTTTTCATGGCTGTCGCTGCCAACAAACGTTCGGTAATGACGCTGTTTTCCGGTCCTACTGACATCTATAGCCATCAGGTCC
             MetAlaValAlaAlaAsnLysArgSerValMetThrLeuPheSerGlyProThrAspIleTyrSerHisGlnValArg

       640        650        660        670        680        690        700        710        720
GCATTGTGCTGGCTGAGAAAGGTGTAAGTTTCGAGATCGAACACGTGGAAAAGGACAATCCGCCTCAGGATCTGATTGACCTCAACCCGA
  IleValLeuAlaGluLysGlyValSerPheGluIleGluHisValGluLysAspAsnProProGlnAspLeuIleAspLeuAsnProAsn

       730        740        750        760        770        780        790        800        810
ATCAGAGCGTTCCGACCCTGGTGGATCGTGAGCTGACCCTGTGGGAATCTCGCATCATTATGGAATATCTGGATGAGCGTTTCCCGCATC
  GlnSerValProThrLeuValAspArgGluLeuThrLeuTrpGluSerArgIleIleMetGluTyrLeuAspGluArgPheProHisPro

       820        830        840        850        860        870        880        890        900
CGCCACTGATGCCTGTTTACCCGGTAGCTCGCGGTGAAAGCCGTCTGTACATGCATCGCATCGAAAAAGACTGGTACACGCTGATGAACA
  ProLeuMetProValTyrProValAlaArgGlyGluSerArgLeuTyrMetHisArgIleGluLysAspTrpTyrThrLeuMetAsnThr

       910        920        930        940        950        960        970        980        990
CCATCATCAACGGTTCAGCTTCTGAAGCAGATGCCGCACGTAAGCAACTGCGCGAAGAACTGCTGGCGATTGCGCCGGTCTTCGGTCAGA
  IleIleAsnGlySerAlaSerGluAlaAspAlaAlaArgLysGlnLeuArgGluGluLeuLeuAlaIleAlaProValPheGlyGlnLys

      1000       1010       1020       1030       1040       1050       1060       1070       1080
AGCCGTACTTCCTGAGCGATGAGTTCAGCCTGGTCGATTGCTATCTTGCTCCGCTGCTGTGGCGTCTGCCGCAACTGGGCATCGAGTTCA
  ProTyrPheLeuSerAspGluPheSerLeuValAspCysTyrLeuAlaProLeuLeuTrpArgLeuProGlnLeuGlyIleGluPheSer

      1090       1100       1110       1120       1130       1140       1150       1160       1170
GCGGCCCGGGTGCGAAAGAGCTGAAAGGCTATATGACCCGCGTCTTTGAGCGTGACTCTTTCCTTGCTTCTTTAACTGAAGCAGAACGTG
  GlyProGlyAlaLysGluLeuLysGlyTyrMetThrArgValPheGluArgAspSerPheLeuAlaSerLeuThrGluAlaGluArgGlu

      1180       1190       1200       1210       1220       1230       1240       1250       1260
AAATGCGTCTGGGCCGGAGTTAATCTGTATGGATTTGTCACAGCTAACACCACGTCGTCCCTATCTGCTGCGTGCATTCTATGAGTGGTT
  MetArgLeuGlyArgSer

      1270       1280       1290       1300       1310       1320       1330       1340       1350
GCTGGATAACCAGCTCACGCCGCACCTGGTGGTGGATGTGACGCTCCCTGGCGTGCAGGTTCCTATGGAATATGCGCGTGACGGGCAAAT

      1360       1370       1380       1390       1400       1410       1420       1430       1440
CGTACTCAACATTGCGCCGCGTGCTGTCGGCAATCTGGAACTGGCGAATGATGAGGTGCGCTTTAACGCGCGCTTTGGTGGCATTCCGCG

      1450       1460       1470       1480       1490       1500       1510       1520       1530
TCAGGTTTCTGTGCCGCTGGCTGCCGTGCTGGCTATCTACGCCCGTGAAAATGGCGCAGGCACGATGTTTGAGCCTGAAGCTGCCTACGA

      1540       1550       1560       1570       1580       1590       1600       1610       1620
TGAAGATACCAGCATCATGAATGATGAAGAGGCATCGGCAGACAACGAAACCGTTATGTCGGTTATTGATGGCGACAAGCCAGATC
```

Fig. 2. Nucleotide sequence of the 1616 bp fragment and the deduced amino acid sequence of the SSP gene. The nucleotides are numbered as shown in the legend of Fig. 1. Three partial amino acid sequences underlined coincide exactly with those determined directly on the N-termini of intact SSP and of two cyanogen bromide-cleaved peptides of SSP. The boxed sequences indicate the initiation and termination codons, and the SD sequences for the SSP gene and also for the next coding region. Thin arrow lines show main inverted sequences, and the thick arrow indicates the initiation site for the SSP transcription.

Table 1. Amino Acid Composition of the SSP

	From purified protein			From the DNA sequence
	Acid hydrolysis for 24h	Acid hydrolysis for 72h	Avarage	
Asp	18.3	18.1	18.2(18)	11
Asn				6
Thr	8.6	8.0	8.9(9)[a]	8
Ser	14.6	12.3	15.8(16)[a]	15
Glu	26.1	25.5	25.8(26)	20
Gln				6
Pro	14.8	16.2	15.5(16)	15
Gly	12.1	12.0	12.1(12)	10
Ala	16.3	15.7	16.0(16)	15
Val	12.4	12.8	12.8(13)[c]	13
Met	5.2	4.8	5.2(5-6)[b]	7[d]
Ile	9.5	10.2	10.2(10)[c]	11
Leu	24.8	25.4	25.1(25)[b]	25
Tyr	3.9	0.8	3.9(4)[b]	8
Phe	8.8	8.6	8.7(9)	9
Lys	8.2	8.0	8.1(8)	8
His	4.7	4.3	4.5(4-5)	4
Arg	12.9	8.5	10.7(11)	16
Cys	-	-	-	1
Trp	-	-	-	3
Total			202-204	211

a) Values extrapolated to those at 0h.
b) Values at 24h.
c) Values at 48h.
d) Excluding the N-terminal one.

cleaved peptides, Met-Glu-Tyr-Leu-Asp and Met-Pro-Val-Tyr-Pro-Val are also found downstream in this frame (also underlined in Fig. 2). An AUG codon precedes the N-terminal alanine, and a SD sequence, GGAGG, is found shortly upstream. It was thus indicated that the AUG at nucleotide 555 is the initiation codon for the translation of SSP. This reading frame ends at the termination codon UAA at nucleotide 1191. The resulting coding region from A (555) to T (1190) encodes 212 amino acid residues. The deduced amino acid

CODON USAGE OF THE SSP GENE

*oXoXOOo ȯ*XOX oXo ȮX OXoXoOȯXOoo OXOO˙oOX XoXO ȮX XOoXO oȯ
oOo oXOoO*O oOX*OXO˙XoOo ooO*ẊoOooooo0 ȯOO* oOOO *ȮXO*OoOOOȯ
o Oo XooẊXOoOOOOoXȯoXOoOXoOOȮ XO OX ẊXooO0*oOoẊOoOXO oooȯ
OXOOoX*ooẊXXo OXo ẊoOoOoO*oOȯX

O = 64, o = 64, X = 42, f = 0.75

Fig. 3. Occurrences of optimal and non-optimal codons in the SSP gene. Optimal codons (O,o), non-optimal codons (x), and Met or Trp (*). Blank space indicates amino acid for which no optimal codon is defined.

sequences of SSP is shown below the nucleotide sequence in Fig. 2. In Table 1, the amino acid composition of this coding region is compared to that determined directly for purified SSP. Except for arginine and tyrosine residues, which degraded rapidly on acid hydrolysis for unknown reasons, two amino acid composition data coincide very well. The molecular weight of the deduced SSP is 24,305 in good agreement with the value obtained by SDS polyacrylamide gel electrophoresis (3). The protein is rich in both positively and negatively charged amino acids; it contains 7.5% arginine, 3.8% lysine, 9.4% glutamic acid and 5.2% aspartic acid.

Codon usage of the SSP gene was examined. The frequency of optimal codon use (Fop) is 0.75, suggesting that about 10^3 molecules of the SSP are present per genome in cells (13). The value is in agreement with the number of the SSP molecules which was estimated in cells growing exponentially in ordinary conditions (Fukuda, R., unpublished observation). Fig. 3 shows occurrences of optimal and nonoptimal codons in the SSP gene.

A computer-assisted comparison of the predicted sequence of the SSP with more than 3,300 published protein sequences (NBRF Protein Data Base) did not reveal any significant homologies.

Analysis of the 5' end of SSP gene transcript(s)

To determine the 5' end of transcript(s) of the SSP gene, S1 nuclease mapping analysis was performed using total cellular RNA which was prepared from cells grown under ordinary growth conditions as described in Materials and Methods. For a preliminary experiment, the 279 bp AvaII (the nucleotide position 600)-HincII (321) fragment was used as the probe. The AvaII site is located in the coding region of the SSP gene and its 5'-end was specifically

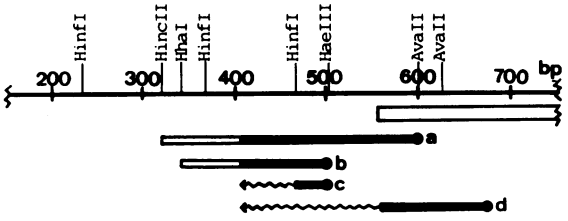

Fig. 4. The hybridization probes for nuclease S1 mapping, and the primers for primer extension experiments. Upper are shown the restriction sites for used endonucleases. Wide open bar indicates the SSP coding region. The filled circles indicates 5'-ends labeled with ^{32}P. For probes a and b, filled bars indicate the regions protected against nuclease S1 digestion, while open regions were digested. The wavy lines indicate the extended regions of the reverse transcripts.

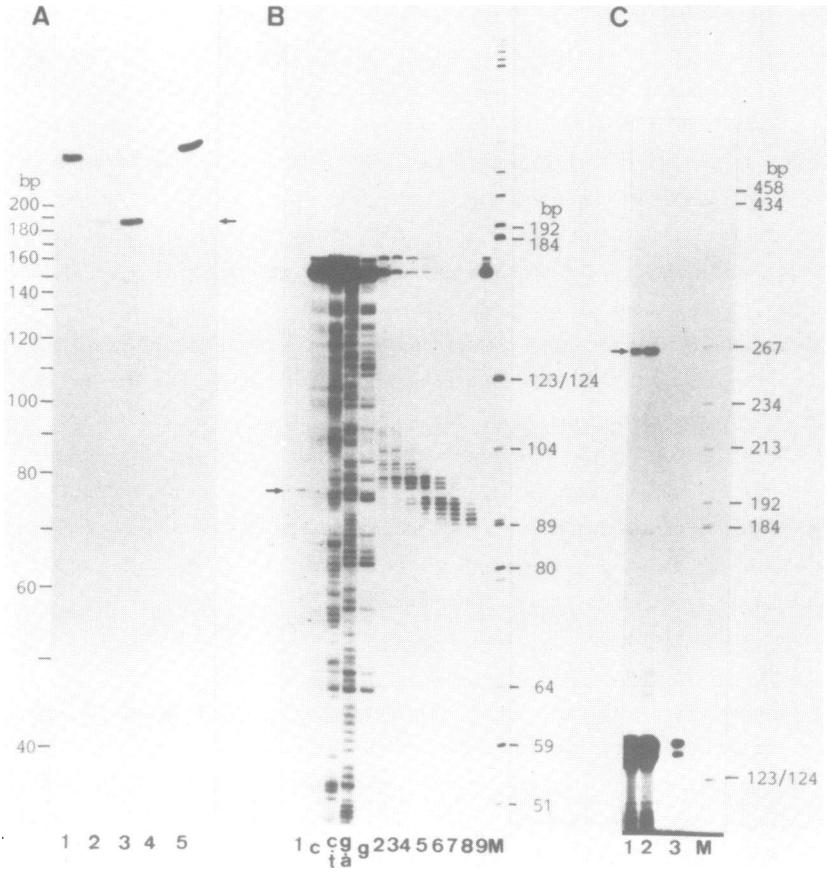

Fig. 5 Nuclease S1 mapping and primer extension analysis of 5' end(s) of the SSP transcripts. [A] Nuclease S1 mapping using the probe a. The probe a (3.4×10^4 cpm) was heat-denatured, and hybridized with total cellular RNA (111 µg). The mixtures were gradually cooled down to 37°C and digested with 100 u of nuclease S1 at 37°C for 15 min. One-fifth and whole reaction products were analysed by electrophoresis on a sequencing gel in lanes, 2 and 3, respectively. In lane 1, nuclease S1 was not added and one-fifth of the reaction products was applied to the gel, while in lane 4, the probe and RNA mixture was quenched at 0°C after heating at 75°C. In lane 5, one-tenth of the probe was directly applied. On the right side, nucleotide lengths obtained from the sequencing ladder of the probe a are indicated. [B] lanes 2 to 9; Nuclease S1 mapping using the probe b. The probe b (4.6×10^4 cpm) and cellular RNA (150 µg) were hybridized and nuclease S1 digestion was performed at 30°C for 15 min in the presence of .4, .8, 1.7, 3.3, 6.6, 13.3 and 26.5 units of the enzyme for lanes 2 to 8, respectively. In lane 9, one-tenth of the probe was directly applied to the gel. Lane 1; products of primer extension reactrion using the primer c (3×10^4 cpm) and cellular RNA (150 µg). Lane m; ^{32}P-DNA markers. [C] Primer extension analysis using primer d (7×10^4 cpm) and cellular RNA (150 µg). Reverse transcript(s), probe only and ^{32}P DNA markers, in lanes 1, 2 and 3, respectively.

labeled with ^{32}P (see Fig. 4, a). As shown in Fig. 5A, lanes 2 and 3, a major fragment of around 190 b remained protected against nuclease S1 digestion. In addition, six to seven very faint fragments were observed which were shorter than the major fragment. But no fragments were seen which were longer than the major fragment. All these fragments could be neither detected when RNA was omitted from the hybridization mixture (data, not shown), nor when the hybridization mixture was quenched at -80°C after heating to 75°C.

To determine the 5' end of the transcript(s) more exactly, another probe was used to get shorter protected fragments on nuclease S1 digestion. The probe was the 162 bp HaeIII (503)-HhaI (341) fragment which was labeled with ^{32}P at the 5' HaeIII end (Fig. 4, b). The nucleotide length of the protected ^{32}P-DNA fragments was around 100 b, but varied remarkably depending on the amount of nuclease S1 added (Fig. 5B, lanes 2 to 8). The 3' end of the protected probe was roughly coincided with that of the experiment showin in A.

Another analysis, the primer extension experiment was thus undertaken. The primer used in Fig. 5B, lane 1 was the 36 bp HaeIII(503)-HinfI(467) fragment, which was labeled with ^{32}P at the 5' end of the HaeIII site (Fig. 4, c). The denatured primer was hybridized with total cellular RNA, and extended with AMV reverse transcriptase. The products were electrophoresed in parallel with Maxam and Gilbert sequencing ladders prepared for the ^{32}P-HaeIII-HhaI fragment (Fig. 4, b). Two bands of the ^{32}P reverse transcripts were observed, a major band and a minor one which was one base shorter (Fig. 5B, lane 1). The upper band of the 5'-^{32}P transcript located between C (408) and T (407), indicating that the base of its 3' end was complementary to A (406), and the transcript was 98b long. On prolonged exposure, several faint bands were seen between the major band and the primer, but no bands were observed above the major band.

To confirm the result, another primer was used, which was complementary to nucleotides 566 to 669 in the coding region of the SSP gene (see Fig. 2), and was uniquely labeled with ^{32}P at the 5' end of the strand which was complementary to the mRNA (Fig. 4, d). As shown in lanes, 1 and 2 in Fig. 5C, only one major band was observed in the nucleotide length of 270 bases in agreement with the nucleotide position in Fig. 4C, lane 1. Both primer extension experiments thus indicated a single major 5' end for the SSP transcript at the position, A (406).

On re-examine the bands of the nuclease S1 mapping analysis (Fig. 5B,

lanes 2 to 8), it was observed that the protected fragments became shorter as the dose of nuclease S1 was increased, finally reaching to a length 5 to 6 nucleotides shorter than the reverse transcript. Probably this is because five or six bases at the 5' end of the transcript (5' AUCCAA$(U)_4$ ----, see Fig. 2) can not form a stable hybrid with the DNA, that the corresponding bases of the probe DNA is very labile to the attack of S1 nuclease. It is noteworthy that the band corresponding to A (406) was always faint, indicating the phosphodiester bond between G and T of the probe strand is resistant to nuclease S1.

If this 5' end is the transcription starting site for the SSP gene, we should be able to find a promoter structure upstream. However, we can not identify the consensus promoter sequences for σ^{70} nor for σ^{32}. Another possibility is that the 5' end we have identified is that of processed RNA, and the transcription starts further upstream. This is, however, unlikely because both nuclease S1 mapping and primer extension analyses did not show any transcript that was initiated upstream (see above). Furthermore a plasmid containing a fragment deleting upstream 321bp from the HincII site on the SalI fragment can express nearly normal level of the SSP gene (Serizawa, H. and Fukuda, R., unpublished observations), indicating the promoter is located between the nucleotide 322 and 406.

The nucleotide sequences flanking the SSP gene.

At the 5' flanking region was found an inverted repeat sequence, which can form a stem and loop structure that is commonly found for the ρ-independent transcription terminator (indicated by inverted arrows in Fig. 2).

At the 3' flanking region, we can find a long coding region in (+1) frame. It starts from ATG (1199∿1201) and continues up to the end of the sequence determined until now. A SD sequence of GGAG is found upstream of the ATG (1186∿1189), one base upstream of the termination codon of the SSP gene. It is thus suggested that another protein is encoded in this 3' flanking region. Since we can neither detect the 3' end of the SSP mRNA, nor 5' end of initiating RNA in this flanking region by nuclease S1 mapping analysis, nor can find any stem and loop structure resembling the ρ-independent terminator in this region, the unidentified gene might be transcribed continuously from the SSP gene.

Acknowledgements

We would like to thank Dr. T. Hase for determination of the amino acid composition of SSP, and Dr. N. Fujita for computer analysis of the DNA

sequence data. We are also very grateful to Dr. A. Ishihama for critical reading of the manuscript. This work was supported by grants from the Ministry of Education, Science and Culture of Japan.

*To whom correspondence should be addressed

REFERRENCES

1. Ishihama, A., Fukuda, R., Kawakami, K., Kajitani, M. and Enami, M. (1980) in Genetics and evolution of RNA polymerase, tRNA and ribosomes, Osawa, S., Ozeki, H., Uchida, H. and Yura, T., Eds., pp105-115, University of Tokyo Press, Tokyo and New York Elsevier/North-Holland Inc.
2. Ishihama, A., Kajitani, M., Enami, M., Nagasawa, H. and Fukuda, R. (1983) in Microbiology-1983, Schlessinger, D. Ed., pp4-6, American Society for Microbiology, Washington DC.
3. Ishihama, A. and Saito, T. (1979) J. Mol. Biol. 129, 517-530
4. Reeh, S., Pederson, S. and Friesen, J.D. (1976) Mol. Gen. Genet. 149, 279-289
5. Fukuda, R., Yano, R., Fukui, T., Hase, T., Ishihama, A. and Matsubara, H. (1985) Mol. Gen. Genet. 201, 151-157
6. Marko, M.A., Chipperfield, R. and Birnboim, H.C. (1982) Anal. Biochem. 121, 382-387
7. Maxam, A.M. and Gilbert, W. (1980) in Methods Enzymol. Grossman, L. and Moldave, K., Eds., Vol. 65, pp 499-580, Academic Press, New York, London.
8. Messing, J. (1983) in Methods Enzymol., Wu, R., Grossman, L. and Moldave, K., Eds., Vol. 101, pp 20-78, Academic Press, New York, London.
9. Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA, 74, 5463-5467
10. Ishii, Y., Hase, T., Fukumori, Y., Matsubara, H. and Tobari, J. (1983) J. Biochem. 93, 107-119
11. Fukuda, R. and Nagasawa-Fujimori, H. (1983) J. Biol Chem. 258, 2720-2728
12. Berk, A.J. and Sharp, P.A. (1978) Proc. Natl. Acad. Aci. USA, 75, 1274-1278
13. Ikemura, T. (1985) Mol. Biol. Evol. 2, 13-34