Proc. Natl. Acad. Sci. USA
Vol. 80, pp. 5480–5484, September 1983
Biochemistry

F sex factor encodes a single-stranded DNA binding protein (SSB) with extensive sequence homology to *Escherichia coli* SSB

(DNA and protein sequences/functional domains/secondary structure/evolution)

JOHN W. CHASE*, BARBARA M. MERRILL†, AND KENNETH R. WILLIAMS†

*Department of Molecular Biology, Albert Einstein College of Medicine, Bronx, NY 10461; and †Department of Molecular Biophysics and Biochemistry, Yale University School of Medicine, New Haven, CT 06510

Communicated by Evelyn M. Witkin, May 23, 1983

ABSTRACT **We have determined the sequence of the gene encoding a single-stranded DNA (ss DNA) binding protein (SSB) from the *Escherichia coli* F sex factor and the amino acid sequence of the protein it encodes. The protein has extensive homology with *E. coli* SSB, particularly within its NH_2-terminal region, where 87 of the first 115 amino acid residues are identical to those of the *E. coli* protein. We have previously shown that this portion of *E. coli* SSB contains the DNA binding region. The sequences diverge extensively in their COOH-terminal regions, although small areas of homology exist in several places. Six of the last seven amino acid residues of the two proteins are identical, which may have implications in terms of the direct interactions of these proteins with other proteins required for DNA replication, recombination, and repair. The coding region of the F plasmid *ssf* gene is 537 base pairs. The protein encoded by the gene contains 178 amino acids (one more than *E. coli* SSB) and has a calculated molecular weight of 19,505. Other than the presumptive Shine–Dalgarno sequence, the promoter and terminator regions of both genes are not similar. The most significant feature in this regard may be the lack of a region of dyad symmetry within the presumptive promoter of the F plasmid *ssf* gene as is found in the region of the presumptive *E. coli ssb* promoter. In this report the predicted secondary structures of both the F plasmid and *E. coli* SSB proteins are compared and the evolutionary significance of their sequence and structural similarities to the functional domains of the proteins are discussed.**

It is now well established that the *Escherichia coli* single-stranded DNA (ss DNA) binding protein (SSB) plays a necessary role in DNA replication (1), recombination (2, 3), and repair (4, 5). Corresponding proteins from several other systems, including eukaryotic organisms, have also been described and similarly implicated in DNA metabolic processes (6, 7). Recently it has been shown that the *E. coli* F sex factor (F plasmid) encodes a protein that binds strongly to ss DNA and partially complements defects caused by a chromosomal *ssb-1* mutation (8). We have determined the amino acid sequence of this protein and the nucleotide sequence of the gene that encodes it and have found it to contain regions of extensive homology to *E. coli* SSB. Although evidence has been presented for corresponding functional domains of the *E. coli* and phage T4 DNA binding proteins (9), obvious sequence homology to a previously characterized SSB has not been described for a protein of this type.

Our results are discussed in terms of the functional domains of *E. coli* and F plasmid SSB and the possible interaction of these proteins with other replication, recombination, and repair proteins.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

MATERIALS AND METHODS

Bacterial Strains and Plasmids. The *E. coli* K-12 strains used were HMS91 (W3110 *rha*$^-$, *thy*$^-$) and KLC850 (*ssb-1* derivative of HMS91). pKAC51 is an ampicillin-resistant tetracycline-sensitive derivative of pBR322 that was constructed by isolating the *Pvu* II/*Cla* I DNA fragment from pBR322 encoding β-lactamase and the origin of replication before nuclease S1 treatment and subsequent ligation and transformation. pKAC50 was constructed by ligating the *Pvu* II/*Cla* I DNA fragment from pBR322 described above for pKAC51 to the *Cla* I/*Sma* I DNA fragment containing the *Bam*HI cleavage site from *Eco*RI fragment 3 from F plasmid. This DNA fragment was isolated from pKL1 (see diagrams of pKL1 and pKAC50 in ref. 8). pEG310 was derived from pKAC50 by cleavage with *Bam*HI, filling in the resulting termini with deoxynucleoside triphosphates in the presence of the large subunit of *E. coli* DNA polymerase I, and ligating (8).

Media and Reagents. All restriction enzymes were from New England BioLabs. Polynucleotide kinase was from P-L Biochemicals. All radioactively labeled nucleoside triphosphates were from Amersham. F plasmid SSB for tryptic peptide analysis was purified from KLC850/pKAC50 by a modification of the previously described procedure (10).

Sequence Determination of the F Plasmid *ssf* Gene. All DNA fragments were purified from 5% or 8% (wt/vol) polyacrylamide gels. Termini were labeled and the sequence was determined as described (11).

Protein Chemistry of F Plasmid SSB. F plasmid SSB was purified by a procedure to be described later. The protein was denatured by adding trichloroacetic acid (10%, vol/vol). The acid-precipitated SSB was redissolved in 8 M urea and then diluted to a final concentration of 2.2 M urea, 0.07 M NH_4HCO_3. After trypsin digestion (protein-to-enzyme weight ratio of 25, 24 hr at 37°C) the resulting peptide mixture was fractionated by HPLC and then characterized as described (12–14). The sequence of amino acid residues 2–14 in F plasmid SSB was determined with 1.5 nmol of SSB that had been coupled to diisothiocyanate glass (15). Some SSB tryptic peptides were also sequenced after they were coupled to aminopolystyrene by using a water-soluble carbodiimide (15). The secondary structure of F plasmid SSB was predicted from a computer program (16) based on the method of Chou and Fasman (17).

Abbreviations: ss DNA, single-stranded DNA; SSB, the *Escherichia coli* or F plasmid ss DNA binding protein [the *E. coli* SSB has also been referred to as a DNA helix-destabilizing protein (*Eco*HD-protein I)]; *ssf*, structural gene for F plasmid-encoded SSB as previously suggested by Kolodkin *et al.* (8); EOP, efficiency of plating.

RESULTS

F Plasmid-Encoded SSB Complements the ss DNA Binding Deficiency and Temperature-Sensitive Conditional Lethality of an *E. coli ssb-1* Mutant Strain. Studies by Kolodkin *et al.* (8) showed that a high copy number plasmid containing *Eco*RI DNA fragment 3 from F plasmid complemented the conditional lethality of an *ssb-1* strain and that mutation of the unique *Bam*HI cleavage site in this DNA fragment prevented this effect. We inserted the *Sma* I to *Cla* I DNA fragment that includes the *Bam*HI cleavage site of *Eco*RI DNA fragment 3 into the *Pvu* II to *Cla* I portion of pBR322 encoding β-lactamase and the origin of replication. This plasmid (pKAC50) encoded two proteins as visualized by maxicells, one of which bound tightly to ss DNA and appeared to have a subunit molecular weight of 22,000. This latter protein was not made by a derivative of pKAC50 that was altered at the *Bam*HI cleavage site (pEG310) (8).

Measurements of the binding activity of extracts of *ssb-1* mutant strains containing these plasmids are shown in Table 1. We have previously shown that extracts of *ssb-1* strains do not contain measurable amounts of ss DNA binding activity as determined under the conditions of our filter binding assay (18). Nearly 80% of wild-type binding activity is, however, restored in an *ssb-1* strain containing pKAC50, but none is detected in the same strain carrying a derivative of pKAC50 with a mutation at the *Bam*HI cleavage site (pEG310). This binding activity also directly correlates with the EOP of these strains measured at 43°C. We conclude, in agreement with the previous studies of Kolodkin *et al.* (8), that these complementation effects of the *ssb-1* mutation are due to a ss DNA binding protein encoded by F plasmid.

Nucleotide Sequence of the F Plasmid-Encoded *ssf* Gene. From the analyses of pKAC50 performed in our laboratory and in Low's laboratory (8), it was known that the *Bam*HI cleavage site was within either the coding sequence or a control region of the F plasmid *ssf* gene. We subsequently determined that a cleavage site for *Bal* I was very close to the *Bam*HI site, and

Table 1. Complementation of ss DNA binding activity by F plasmid-encoded SSB in extracts of *ssb-1* mutant cells

Strain*	Genotype: Chromosome	Genotype: Plasmid	μg SSB† per mg protein	Binding activity,‡ nmol DNA bound per mg protein	EOP (43°C)§
HMS91	*ssb*⁺	—	0.31	0.78	0.93
KLC850	*ssb-1*	—	0.36	<0.1	0.00004
KLC850/ pKAC51	*ssb-1*	—	0.38	<0.1	0.00001
KLC850/ pKAC50	*ssb-1*	F *ssb*⁺	0.46	0.61	0.85
KLC850/ pEG310	*ssb-1*	F *ssb*⁻	0.48	<0.1	0.000001

* pKAC51 is the vector portion of pKAC50 and pEG310. pEG310 was derived from pKAC50 by mutation of the *Bam*HI cleavage site.

† Determined by solid-phase radioimmunoassay of crude extracts using antibody to wild-type SSB. Although precise measurements of the quantity of F plasmid SSB produced by KLC850/pKAC50 have not been made, qualitative estimates have been made on the basis of sodium dodecyl sulfate/polyacrylamide gels of fractions analyzed during purification; it is interesting that more SSB was not detected. Despite extensive homology between F plasmid and *E. coli* SSB it is possible that antibody to *E. coli* SSB does not react well with the F plasmid protein; however, this point has not been directly examined.

‡ ss DNA binding activity determined in crude extracts by method of Whittier and Chase (18).

§ Efficiency of plating (EOP) at 43°C is the ratio of viable cells at 43°C to those at 30°C determined on LB agar plates without NaCl.

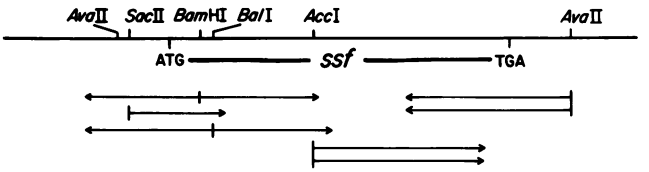

FIG. 1. Diagram of the portion of pKAC50 sequenced in this study. The ATG start and TGA termination codons of the F plasmid *ssf* gene are indicated. Also indicated is the sequencing strategy (see *Results*). The arrows below the map indicate the direction of sequencing and length of sequences determined in individual experiments.

we also determined the location of *Ava* II cleavage sites in pKAC50. Cleavage sites for other restriction enzymes utilized in the sequencing of the gene were found as sequences were determined, proceeding from the 5′ termini of the *Bal* I site and the 3′ termini of the *Bam*HI site. The region of pKAC50 whose sequence was determined and the sequencing strategy are diagrammed in Fig. 1. It was apparent when the first sequences were determined that extensive homology existed between the *E. coli ssb* (13) and F plasmid *ssf* genes, at least within the regions of the genes encoding the NH_2 termini of the proteins. Further consideration of the implications of the similarities in sequence and structure of these ss DNA binding proteins are discussed below. The DNA and protein sequences of the F plasmid *ssf* and *E. coli ssb* genes and proteins are both presented in Fig. 2 for comparison.‡ The coding region of the F plasmid *ssf* gene is 537 base pairs. However, because the methionine of the initiation codon is not present in the protein, F plasmid SSB contains 178 amino acid residues—one more than *E. coli* SSB. Both SSBs utilize a UGA termination codon. The presumptive promoter sequences of the *E. coli ssb* and F plasmid *ssf* genes are not similar beyond their presumptive Shine–Dalgarno sequences (Fig. 3*A*). The coding region of the F plasmid *ssf* gene is followed by a dyad symmetry of 19 base pairs compared to 27 base pairs for the *E. coli ssb* gene (13) (Fig. 3*B*). The sequences of these regions are not similar.

Protein Chemistry of F Plasmid SSB. Solid-phase sequence analysis of F plasmid SSB resulted in the identification of residues 2–6 and 8–14, all of which were in complete agreement with those predicted on the basis of the DNA sequence given in Fig. 2. Dansylation of 0.3 nmol of the protein revealed alanine as the NH_2-terminal amino acid, also in agreement with the DNA sequence. Carboxypeptidase A digestion of 1 nmol of F plasmid SSB released 0.92 nmol of phenylalanine and no other amino acids, again in accord with the DNA sequence.

The remainder of the amino acid sequence was confirmed by determining amino acid compositions for both the intact F plasmid SSB and its tryptic peptides. In addition, the sequences of three of the tryptic peptides were determined. The total amino acid composition of F plasmid SSB agrees well with the predicted composition (data not shown). Hydrolysis and analysis of performic acid-oxidized SSB and of tryptic peptide T-8 (Table 2) confirmed the finding that (in contrast to the *E. coli* SSB) the F plasmid protein does contain a single cysteine. A better confirmation of the DNA sequence for F plasmid SSB was obtained by isolating 14 of its tryptic peptides, which together account for 161 out of the 178 amino acids in this protein (Table 2). The amino acid compositions of these purified peptides as well as the sequences determined for T-2, T-7, and T-8 are all in ex-

‡ Two errors in the *E. coli ssb* sequence (13) have now been corrected. The codon for serine-39 was reported to be TCA and has now been shown to be TCC. It should be noted that this change introduces a *Bst*NI recognition site. Amino acid residue 133 was reported to be serine (codon AGC). This residue has now been correctly shown to be glycine (codon GGC). These corrections have been confirmed by both DNA and protein sequencing (unpublished data).

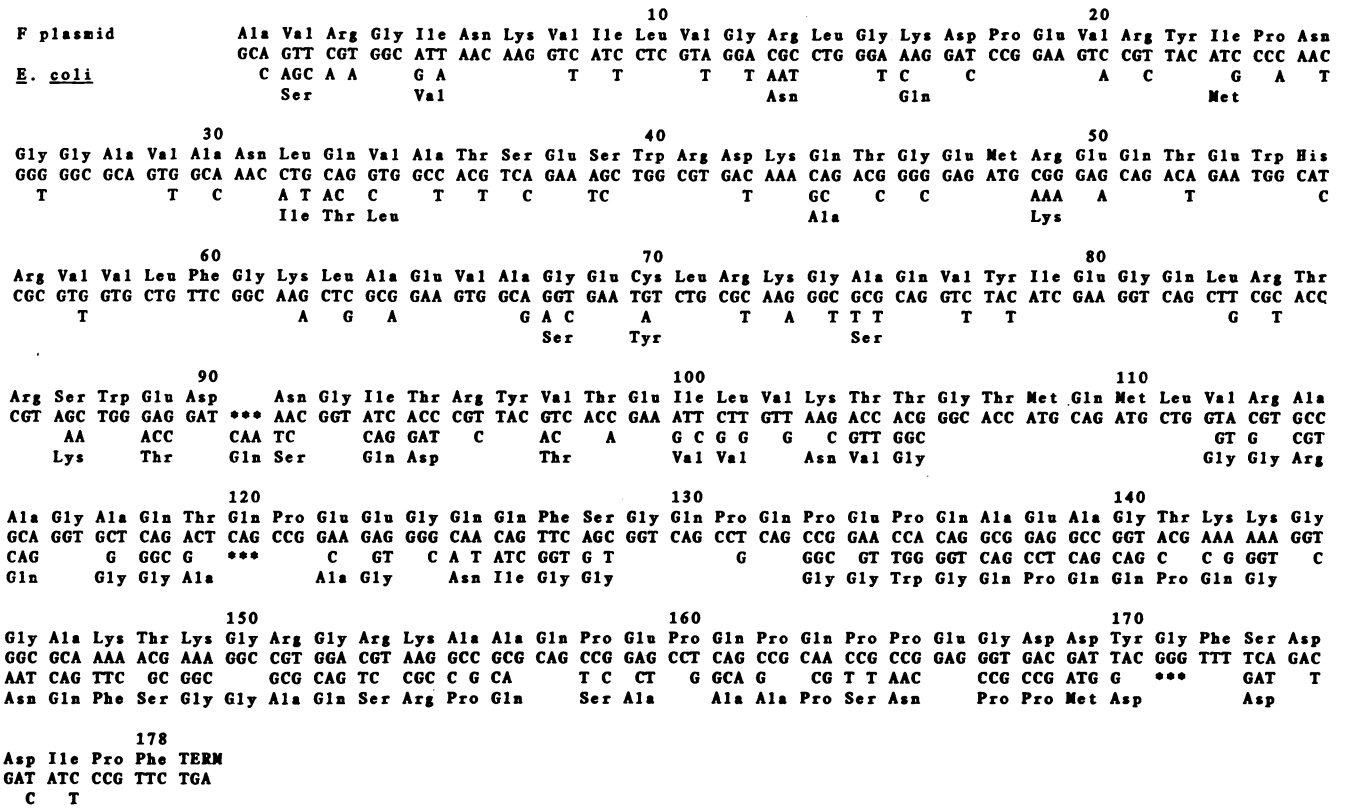

FIG. 2. Nucleotide and amino acid sequences of the F plasmid *ssf* and *E. coli ssb* genes and proteins. The complete sequences of F plasmid *ssf* are given and differences that exist in the *E. coli* sequences are listed below. At one place in the F plasmid *ssf* sequence and at two places in the *E. coli ssb* sequence a space equivalent to one codon has been introduced in order to maximize homology between the amino acid sequences. These regions are indicated by asterisks. F plasmid SSB contains 178 amino acid residues and *E. coli* SSB contains 177. TERM, termination codon.

cellent agreement with the compositions expected from the DNA sequence. Carboxypeptidase A digestion of T-14 released 0.8 nmol of phenylalanine per mol of peptide and no other amino acids, thus confirming that this peptide is at the COOH terminus of the F plasmid SSB.

A secondary structure for F plasmid SSB was predicted by an empirical method based on its amino acid sequence (Fig. 4). The first two-thirds of the protein (approximately residues 1–115) is highly ordered, with 70% of the amino acids predicted to occur in either α-helical or β-pleated structures. In contrast, the COOH-terminal third of SSB has no predicted β-pleated sheets and only three short stretches of α-helix, which together account for less than 30% of this region. While charged amino acids are distributed uniformly throughout the first two-thirds of the F plasmid SSB, the COOH-terminal third of the molecule contains both an unusually basic region (residues 142–154 have a net charge of +7) and an unusually acidic region (residues 166–175 have a net charge of −5). Overall, the F plasmid SSB should have a total net charge at pH 7 of only +1 compared to −2 for *E. coli* SSB.

DISCUSSION

We determined the sequence of a gene from F plasmid that encodes a ss DNA binding protein and confirmed 90% of the predicted amino acid sequence by direct sequence determination or by amino acid analysis of tryptic peptides. The protein encoded by this gene has extensive homology to *E. coli* SSB. This is most evident in the NH_2-terminal two-thirds of the protein (residues 1–115), where nearly 80% of the sequence exactly corresponds to the *E. coli* protein. Analysis of the secondary structure by the method of Chou and Fasman (17) predicts this region to be very similar to *E. coli* SSB (Fig. 4). Our previous studies utilizing proteolytic products of *E. coli* SSB have shown that this portion of the protein contains the DNA binding region (9). We have also shown that the *ssb-1* mutation of

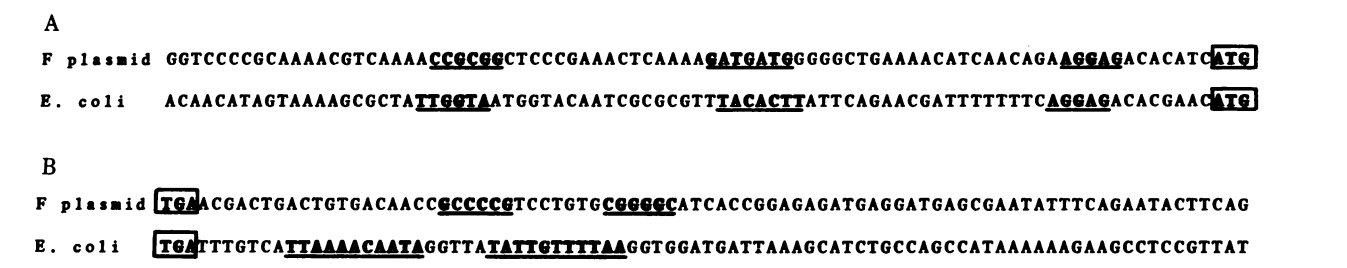

FIG. 3. Comparison of the presumptive F plasmid *ssf* and *E. coli ssb* promoter (*A*) and terminator (*B*) sequences. The presumptive RNA polymerase recognition sequences, Pribnow sequences, and Shine–Dalgarno sequences are underlined. Those indicated for the *E. coli ssb* gene are as suggested previously (13). The dyad symmetry following the coding region of each gene is underlined. The ATG initiation codon and the termination codon of each gene are enclosed in boxes.

Table 2. Amino acid compositions* of F plasmid SSB tryptic peptides

Amino acid	T-1	T-2†	T-3	T-4	T-5	T-6	T-7‡	T-8‡	T-9	T-10	T-11	T-12	T-13	T-14§
Cys	–	–	–	–	–	–	–	0.8 (1)	–	–	–	–	–	–
Asx	–	–	1.0 (1)	2.0 (2)	0.8 (1)	–	–	–	0.3	1.7 (2)	–	–	0.6	3.8 (4)
Thr	–	–	–	0.9 (1)	1.1 (1)	0.9 (1)	–	–	–	0.9 (1)	1.1 (1)	2.8 (3)	2.1 (2)	0.3
Ser	–	0.4	–	2.0 (2)	–	–	–	–	–	0.8 (1)	–	–	1.2 (1)	1.2 (1)
Glx	–	0.6	1.1 (1)	2.2 (2)	2.2 (2)	3.0 (3)	–	1.9 (2)	2.9 (3)	1.4 (1)	1.2 (1)	1.2 (1)	10.5 (11)	4.8 (5)
Pro	–	–	1.2 (1)	1.1 (1)	–	–	–	–	–	–	–	–	4.6 (4)	6.3 (6)
Gly	–	1.4 (1)	1.0 (1)	2.7 (2)	1.1 (1)	0.4	1.0 (1)	1.0 (1)	2.1 (2)	1.2 (1)	–	1.1 (1)	4.1 (4)	2.4 (2)
Ala	1.0 (1)	–	–	2.6 (3)	–	–	–	1.8 (2)	1.0 (1)	–	–	–	4.5 (5)	2.1 (2)
Val	1.1 (1)	1.7 (2)	0.9 (1)	1.8 (2)	–	–	1.5 (2)	0.7 (1)	1.0 (1)	–	1.9 (2)	1.2 (1)	–	0.3
Met	–	–	–	–	0.8 (1)	–	–	–	–	–	–	1.4 (2)	–	–
Ile	–	0.6 (1)	–	0.8 (1)	–	–	–	–	0.9 (1)	0.8 (1)	0.9 (1)	–	–	1.0 (1)
Leu	–	1.0 (1)	0.9 (1)	1.1 (1)	–	–	0.9 (1)	1.8 (2)	1.1 (1)	–	1.1 (1)	1.1 (1)	–	0.5
Tyr	–	–	–	0.8 (1)	–	–	–	–	0.8 (1)	–	0.9 (1)	–	–	0.8 (1)
Phe	–	–	–	–	–	–	1.0 (1)	–	–	–	–	–	0.9 (1)	1.6 (2)
His	–	–	–	–	–	0.9 (1)	–	–	–	–	–	–	–	–
Lys	–	–	0.9 (1)	–	0.7 (1)	–	1.0 (1)	–	1.0 (1)	–	1.0 (1)	–	1.0 (1)	1.0 (1)
Arg	1.0 (1)	1.3 (1)	1.1 (1)	1.1 (1)	1.2 (1)	1.2 (1)	–	1.4 (1)	1.2 (1)	1.3 (1)	–	1.1 (1)	–	0.4
Trp	–	–	–	+ (1)	–	+ (1)	–	–	–	+ (1)	–	–	–	–
Yield, %	57	52	71	25	32	54	53	35	45	54	48	58	50	34
Residue numbers	1–3	8–13	14–21	22–41	42–49	50–56	57–62	63–72	73–84	87–95	96–103	104–113	114–142	154–178

*Presented as residues per peptide molecule. Numbers in parentheses are calculated from the DNA sequence.
†Identity confirmed by solid-phase sequence analysis of the peptide.
‡These two peptides were isolated as a 60:40 ratio of T-7 to T-8. The sequence of the mixture was then determined, to confirm the identification.
§An additional 3.7 nmol (30%) of this peptide was recovered as T-14', residues 155–178.

E. coli SSB results in the substitution of tyrosine for histidine at position 55, resulting in the production of a protein defective in DNA binding (refs. 12 and 18 and unpublished results). This alteration should interrupt the extensive α-helical region predicted to occur between residues 45 and 67 of both proteins. These correlations are striking and further suggest the importance of this predicted α-helical region in DNA binding.

Our previous characterization of proteolytic products of *E. coli* SSB produced by digestion with trypsin and chymotrypsin has led to the conclusion that regions of the protein in the area

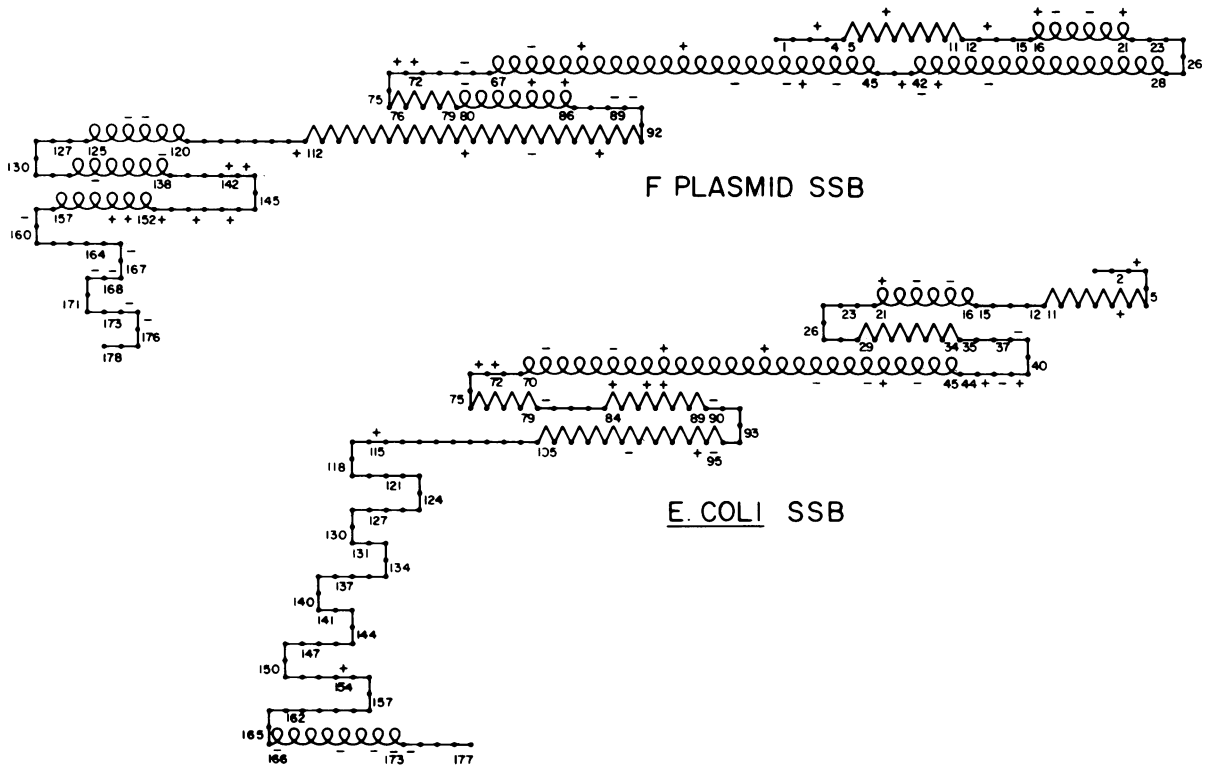

FIG. 4. Secondary structure of F plasmid SSB. ⌇, ⩘, and ⟿ denote α-helical, β-pleated-sheet, and random-coil residues, respectively; β-structure turns are indicated by chain reversals. The + and − signs refer to charged residues (pH 7), and the numbers refer to residues at conformational boundaries.

of residues 116–135 are important in maintaining the integrity of its tetrameric structure and may also participate in interactions between tetramers that result in cooperative binding to ss DNA (9). Although only 7 of the 19 amino acids in this region (residues 116–135) are identical in both the *E. coli* and F plasmid SSB, there is one stretch of 4 amino acids (residues 129–132) that are homologous and that are predicted to form a β-bend in both proteins.

There is very little homology between the two proteins within their COOH-terminal regions. Between residues 136 and 170 only 4 residues of the F plasmid protein directly correspond to *E. coli* SSB. In addition, the F plasmid SSB contains 12 charged amino acids in this region compared to only 3 for the *E. coli* SSB. Even so, the difference in *net* charge is only 3 because the region containing residues 136–170 should have a net charge of +2 for the F plasmid SSB compared to −1 for the *E. coli* SSB. Although this entire region is devoid of extensive secondary structure, α-helical regions are predicted to occur in the F plasmid SSB between residues 133–138 and 152–157. This differs from *E. coli* SSB, which is predicted to be completely random coil in this region except for several β-turns. It also differs in that the α-helical region predicted to occur in the *E. coli* protein between residues 166 and 173 is not predicted to occur in the F plasmid-encoded protein. It is striking, however, that 6 of the last 7 amino acids (residues 172–178) of the F plasmid protein exactly correspond to the COOH-terminal sequence of *E. coli* SSB. It is important to note that the *ssb-113* mutation of *E. coli* SSB results in the substitution of serine for proline-176 (the next-to-the-last amino acid residue in the *E. coli* protein). This alteration does not greatly affect DNA binding as does the *ssb-1* mutation (unpublished results) but does affect DNA replication, repair, and recombination in much the same manner as *ssb-1*, although with a number of subtle differences (4). The known effects of this alteration and the fact that the COOH-terminal amino acid sequences of both proteins are nearly identical strongly suggests that the integrity of this region is essential for the functioning of both proteins. On the basis of genetic and biochemical arguments, we have previously suggested that the pleiotropic effects exhibited by a deficiency in SSB could best be explained if the protein exerts its effects not only through binding ss DNA but also by directly interacting with proteins intimately involved in DNA replication, recombination, and repair (4). We have suggested that the COOH terminus of the protein is the most likely region for such interactions (19), and we believe that the observations we have made on F plasmid SSB strengthen these arguments.

The sequences immediately preceding the initiation codon of each gene are very similar; however, the sequences in the presumptive promoter and terminator regions of these genes are generally not homologous (Fig. 3) and they show minimal homology to the consensus sequence (20). The dyad symmetry that exists in the *E. coli* SSB promoter (13) is not present in the F plasmid gene. The significance of this dyad symmetry, if any, is still unknown. We speculated that it may be a recognition site for a regulator that affects transcription of *ssb* (13). *In vivo* studies in our laboratory (unpublished results) and recent studies of Salles *et al.* (21) have failed to produce any evidence for either positive or negative regulation of *ssb*. The question of *ssb* regulation in *E. coli* is particularly important due to the known involvement of SSB in induction of the SOS repair process and the proximity of the *ssb* and *uvrA* promoters, the latter gene being a known damage inducible (*din*) locus (22). *In vitro* transcription or gene fusion studies of both genes are needed in order to evaluate their regulatory mechanisms. Because the F plasmid and *E. coli* ss DNA binding proteins almost certainly did not evolve independently, it is interesting that their control regions are quite different, at least in sequence. It may be of basic importance that we understand their regulation in order to arrive at a complete understanding of their participation in biochemical reactions.

It is important to note the absence of the region of DNA from F plasmid encoding its ss DNA binding protein from mini-F plasmids capable of autonomous replication (23, 24). This observation suggests an additional role for F plasmid SSB apart from a role in replication that may be compensated for by the corresponding host protein. It has been noted by Kolodkin *et al.* (8) that certain other transmissible plasmids also produce an activity that suppresses the phenotype of *ssb-1*. Analysis of the ss DNA binding proteins from these systems may lead to a better understanding of the interactions of various DNA metabolic proteins that is crucial to a complete understanding of DNA replication, recombination, and repair, as well as DNA transfer at the molecular level.

It is a pleasure to acknowledge the free exchange of ideas between our laboratory and that of Dr. Brooks Low throughout this study. We are also grateful to Dr. William Konigsberg for his advice and encouragement during this work and to Janet Murphy and Elke Lorenson for their expert technical assistance. We especially thank Laura Dalheim for operating the Sequemat Mini-15 sequencer. This investigation was supported by Public Health Service Grants GM11301, GM23451, and CA13330 (to J.W.C.), Public Health Service Grant GM12607 (to William H. Konigsberg), and National Science Foundation Grant PCM-81-04118 (to K.R.W.). J.W.C. is an established investigator of the American Heart Association, Grant 78-129. B.M.M. is a Predoctoral Fellow of the National Science Foundation.

1. Meyer, R. R., Glassberg, J. & Kornberg, A. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 1702–1705.
2. Glassberg, J., Meyer, R. & Kornberg, A. (1979) *J. Bacteriol.* **140**, 14–19.
3. Golub, E. I. & Low, K. B. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 1401–1405.
4. Whittier, R. F. & Chase, J. W. (1983) *Mol. Gen. Genet.* **190**, 101–111.
5. Lieberman, H. B. & Witkin, E. M. (1983) *Mol. Gen. Genet.* **190**, 92–100.
6. Coleman, J. & Oakley, J. (1979) *Crit. Rev. Biochem.* **7**, 247–289.
7. Falaschi, A., Cobianchi, F. & Riva, S. (1980) *Trends Biochem. Sci.* **5**, 154–157.
8. Kolodkin, A. L., Capage, M. A., Golub, E. I. & Low, K. B. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 4422–4426.
9. Williams, K. R., Spicer, E. K., LoPresti, M. B., Guggenheimer, R. A. & Chase, J. W. (1983) *J. Biol. Chem.* **258**, 3346–3355.
10. Chase, J., Whittier, R., Auerbach, J., Sancar, A. & Rupp, W. D. (1980) *Nucleic Acids Res.* **8**, 3215–3227.
11. Maxam, A. & Gilbert, W. (1980) *Methods Enzymol.* **65**, 499–560.
12. Williams, K. R., L'Italien, J. J., Guggenheimer, R. A., Sillerud, L., Spicer, E., Chase, J. W. & Konigsberg, W. (1982) in *Methods in Protein Sequence Analysis*, ed. Elzinga, M. (Humana, Clifton, NJ), pp. 499–507.
13. Sancar, A., Williams, K. R., Chase, J. W. & Rupp, W. D. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 4274–4278.
14. Williams, K. R., LoPresti, M. B., Setoguchi, M. & Konigsberg, W. H. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 4614–4617.
15. L'Italien, J. R. & Laursen, R. A. (1982) in *Methods in Protein Sequence Analysis*, ed. Elzinga, M. (Humana, Clifton, NJ), pp. 383–408.
16. Cohen, F. (1979) Dissertation (Yale Univ., New Haven, CT).
17. Chou, P. Y. & Fasman, G. D. (1978) *Adv. Enzymol.* **47**, 45–148.
18. Whittier, R. F. & Chase, J. W. (1980) *Anal. Biochem.* **106**, 99–108.
19. Chase, J. W., Murphy, J. B., Whittier, R. F., Lorensen, E. & Sninsky, J. J. (1983) *J. Mol. Biol.* **164**, 193–211.
20. Rosenberg, M. & Court, D. (1980) *Annu. Rev. Genet.* **13**, 319–353.
21. Salles, B., Paoletti, C. & Villani, G. (1983) *Mol. Gen. Genet.* **189**, 175–177.
22. Kenyon, C. J. & Walker, G. C. (1981) *Nature (London)* **289**, 808–810.
23. Timmis, K., Cabello, F. & Cohen, S. N. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 2242–2246.
24. Lovett, M. A. & Helinski, D. R. (1976) *J. Bacteriol.* **127**, 982–987.