Identification and Characterization of the Single-Stranded DNA-Binding Protein of Bacteriophage P1

HANSJÖRG LEHNHERR,* JANNICK D. BENDTSEN, FABIAN PREUSS, AND TATIANA V. ILYINA

Institute of Molecular Biology, University of Southern Denmark, Main Campus Odense University, DK-5230 Odense M, Denmark

Received 19 April 1999/Accepted 6 July 1999

The genome of bacteriophage P1 harbors a gene coding for a 162-amino-acid protein which shows 66% amino acid sequence identity to the Escherichia coli single-stranded DNA-binding protein (SSB). The expression of the P1 gene is tightly regulated by P1 immunity proteins. It is completely repressed during lysogenic growth and only weakly expressed during lytic growth, as assayed by an ssb-P1/lacZ fusion construct. When cloned on an intermediate-copy-number plasmid, the P1 gene is able to suppress the temperature-sensitive defect of an E. coli ssb mutant, indicating that the two proteins are functionally interchangeable. Many bacteriophages and conjugative plasmids do not rely on the SSB protein provided by their host organism but code for their own SSB proteins. However, the close relationship between SSB-P1 and the SSB protein of the P1 host, E. coli, raises questions about the functional significance of the phage protein.

Bacteriophage P1 infects several enterobacterial species, including Escherichia coli (60). The ability to mediate generalized transduction of chromosomal markers between different strains (35) has gained P1 tremendous practical importance in the construction of new laboratory strains and in the fine mapping of the E. coli chromosome (2). Despite its widespread use in many laboratories around the world, surprisingly little is known about other aspects of the virulent life cycle of bacteriophage P1. Only approximately 60% of the complete nucleotide sequence of the P1 genome is currently accessible in databases. As a consequence, many P1 genes which have been mapped genetically (54, 55, 59) have not yet been identified and characterized physically. One of these genes was described as early as 1982, when Johnson (28) reported that some mutants of bacteriophage P1 were able to suppress a temperature-sensitive defect in the E. coli single-stranded DNA-binding (SSB) protein. E. coli SSB plays an essential role in three fundamental cellular processes, namely, DNA replication, recombination, and repair (for reviews of E. coli SSB, see Chase [5], Lohmann and Ferrari [36], and Meyer and Laine [37]). Also in the 1980s, many bacteriophages and conjugative plasmids were shown to code for their own SSB proteins, and the nucleotide sequences of most of the respective genes have been determined (references 15 and references therein). For reviews of E. coli sensitive defect in the P1 genome is currently accessible in many laboratories around the world, surprisingly little is known about other aspects of the virulent life cycle of bacteriophage P1, it was found that mutations in the auxiliary gene described by Iida and Arber (26).

In this study we report the nucleotide sequence of the P1 ssb gene, show that the expression of ssb-P1 is regulated by the P1 proteins C1 (12, 19) and Lxc (53), and demonstrate that SSB-P1 is sufficient to complement a temperature-sensitive ssb mutant of E. coli. A multiple sequence alignment, including SSB proteins encoded by bacteria, plasmids, and bacteriophages, was constructed. It showed that SSB-P1 has a high degree of sequence similarity to its bacterial counterparts. A possible role of SSB-P1 in the lytic growth cycle of the bacteriophage is discussed.

MATERIALS AND METHODS

Standard procedures and DNA sequencing. Standard DNA techniques, liquid media, and agar plates were used as described by Sambrook et al. (44). Antibiotics were added as appropriate at concentrations of 100 μg/ml for ampicillin, 25 μg/ml for kanamycin, and 25 μg/ml for chloramphenicol. DNA-sequencing reactions were performed as described by Sanger et al. (45), using a Thermo Sequenase-based sequencing kit (Amersham).

Bacterial strains. The E. coli K-12 strains used were UT580 [F' tet' trpΔ36 lacP3ΔlacZM15 proA' B' supD thi Δlac-proAB] (24), KLC438 (F' met thy rha), and KLC436 (F' ssb-1 met thy rha) (51). The ssb-1 allele specifies a temperature-sensitive protein carrying a His55Tyr substitution (37).

Bacteriophages. The bacteriophages used in this study were P1-15:Ts6260 (40), P1Cm (25), P1Cmclcr.100 (25, 43), and P1Cm-lc-2 (42). The lacI gene of P1Cm lxc-2 contains an uncharacterized mutation affecting the function of the auxiliary repressor protein Lxc. The c(ts) genes of P1-15:Ts6260 and P1Cmclcr.100 contain uncharacterized mutations rendering the C1 protein temperature-sensitive. Lysogenic derivatives of different E. coli strains were constructed according to the procedure of Rosner (43). Phage DNA was isolated as described by Iida and Arber (26).

Vectors and plasmids. The vectors pUC19 (58), pBR322 (3), and pACYC184 (4) and the lacZ fusion vector pNM481 (39) were used to clone different P1 restriction fragments. The plasmid pAM1 carries a ColD replication origin and a kanamycin resistance marker (22). The plasmids pAM2b and pAM8 are derivatives of pAM1, carrying in addition the P1 lxc-2 (40), P1Cm (25), P1Cm lxc-2 (42). The lacI gene of P1Cmclcr.100 contains an uncharacterized mutation affecting the function of the auxiliary repressor protein Lxc. The c(ts) genes of P1-15:Ts6260 and P1Cmclcr.100 contain uncharacterized mutations rendering the C1 protein temperature-sensitive. Lysogenic derivatives of different E. coli strains were constructed according to the procedure of Rosner (43). Phage DNA was isolated as described by Iida and Arber (26).

Plasmids constructed in this work. Total P1 DNA was cleaved with the restriction enzymes EcoRI or BamHI, and the P1 restriction fragments EcoRI-lacZ (pHAL245 and pBR322), EcoRI-lacZ (pHAL246 and pUC19), and BamHI-lacZ (pHAL247 and pUC19) were cloned into the indicated vectors cleaved with the corresponding restriction enzymes. These plasmids and several subclones served as templates in sequencing reactions. In order to clone ssb-P1 separately from any other P1 function, we used the plasmid pHAL245 as a template in a PCR, including the two oligonucleotide primers (DNA Technology A/S, Aarhus, Denmark) SSB1 (5'-GGG AGT TCG ATC CCT TTA GAA GAC ACA GGA TA-3') and SSB2 (5'-GGG GAT CCG CGC GTG CCA TTG CCA ACT TTG GGC TT-3'). The 699-bp product of the PCR was used for 5' RACE with the restriction enzyme EcoRI and cloned into the EcoRI/EcoRV site of the cloning vector pACYC184, resulting in plasmid pHAL251. A 327-bp fragment was cloned out of pHAL251 with the restriction enzymes XmnI and EcoRI and was then cloned into the EcoRI/SmaI site of the lacZ fusion vector pNM481. In the resulting indicator plasmid construct, pHAL252, an SSB-P1–LacZ fusion protein was expressed under the control of the ssb-P1 promoter.

Detection of ssb-P1 promoter activity. Cultures of the strain UT580 carrying the indicator plasmid pHAL252 and of derivatives of this strain (carrying in addition one of the following plasmids or P1 prophages: pAM1, pAM2b, pAM8, P1-15:Ts6260, P1Cmclcr.100, P1Cm, or P1Cm-lc-2) were grown into exponential growth phase up to an optical density at 600 nm of 0.6. The cultures were then

* Corresponding author. Mailing address: Institute of Molecular Biology, University of Southern Denmark, Main Campus Odense University, Campusvej 55, DK-5230 Odense M, Denmark. Phone: 65 50 23 74. Fax: 65 93 27 81. E-mail: lehnerr@biobase.dk.
assayed for LacZ activity according to the method of Miller (38). A qualitative indication of promoter activity was obtained by spreading the above-mentioned strains on agar plates containing the lactose analog 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal). Duplicates of the plates were incubated overnight at 30 and 42°C. Blue colonies indicated the expression of the SSB-P1–LacZ fusion protein from pHAL252.

**RESULTS**

**Determination of the nucleotide sequence of ssb-P1.** Intrigued by the fact that the location of ssb-P1 had not been determined previously, we decided to investigate an uncharacterized segment of the P1 chromosome located between map positions 15 and 24 (for a circular map of bacteriophage P1, see Yarmolinsky and Lobocka [59]). We cloned the restriction fragments EcoRI-4 and EcoRI-10 of the P1 isolate P1-15::Tn2680, determined the nucleotide sequence of these two fragments, and aligned them with respect to each other. The resulting 7,885-bp sequence includes the gene lysA, coding for the P1 lysozyme (46), and lies adjacent to the recently published sequence of the P1 dar operon (27). Figure 1 shows a physical map of the sequence, indicating the presence of five open reading frames. Two of them, darB and lysA, are reading in a counterclockwise orientation, while three are oriented clockwise. Two of the latter, orf17 and orf23 (numbered according to their respective map positions on the P1 chromosome [59]), show no significant homology to other known sequences in the databases. The third open reading frame was identified as a new genetic element, which we named the IS1 element.

**Computer analysis.** For nucleotide sequence comparison and handling, the Wisconsin package, version 9.1, of the Genetics Computer Group (10) was used. Database searches were done with the programs Advanced BLAST and PSI-BLAST at the NCBI web server (1, 40a).

**Nucleotide sequence accession number.** The new nucleotide sequence reported in this paper has been submitted to the GenBank Nucleotide Sequence Data Library and has the accession no. AF125376.
found to code for a small, 162-amino-acid protein which showed 66% amino acid sequence identity to the *E. coli* SSB protein, and it was therefore called SSB-P1. Figure 2 shows the nucleotide sequence of ssb-P1, which was further corroborated by determining the corresponding sequence of an independent P1 isolate, P1CM. The P1 ssb gene starts with a GTG codon and is preceded by a weak *E. coli* consensus promoter (17). Immediately downstream of the −10 region of the *E. coli* promoter, a 17-bp asymmetric consensus binding site for the major repressor protein C1 (13, 52) of bacteriophage P1 was found. This C1 binding site, Op21, was identified previously by Citron et al. (6) on a short DNA fragment excluding ssb-P1. A P1-specific late promoter sequence (32, 33), LPr21, was located immediately upstream of the ssb promoter, reading in the opposite direction, expressing the *lysA* gene (46).

### Regulation of ssb-P1 expression

The arrangement of promoter elements shown in Fig. 2 indicated that ssb-P1 is expressed from a weak *E. coli* consensus promoter, regulated by the repressor proteins C1 (12, 19) and Lxc (53). The major P1 repressor protein C1 binds to 17-bp asymmetrical sequences with the consensus ATT GCT CTA ATA AAT TT and reduces the activity of promoter sequences located in the vicinity (21). The auxiliary repressor protein Lxc does not bind DNA on its own but interacts with DNA-bound C1 and in such a ternary complex usually increases the level of repression exerted by C1 (53). However, as Lxc also lowers the concentration of C1 protein in the cell, its effect on different c1-regulated promoter sequences can vary significantly (21, 50). The 17-bp sequence found in the promoter of ssb-P1, ATT GCT CTA ATT AAT TT, shows only a single mismatch (shown in boldface) with the consensus C1 binding site. To experimentally confirm the idea that ssb-P1 is regulated from the promoter shown in Fig. 2, we constructed a fusion of ssb-P1 to *lacZ* (see Materials and Methods). The resulting indicator plasmid, pHAL252, was then assayed in the presence or absence of different P1 functions. In Table 1, the 685 Miller activity units expressed from pHAL252 in the absence of any P1 functions was set to 100%. If the major repressor protein C1 was expressed from plasmid pAM2b, expression of the SSB-P1–*lacZ* fusion protein from pHAL252 was reduced to less than 50%, demonstrating that the ssb-P1 promoter is indeed regulated by C1. If the corepressor protein Lxc was expressed, in addition to C1, from plasmid pAMS, expression from pHAL252 was further reduced. In the presence of a P1 lysogen, either P1-15::Tn2680, P1Cmcr100, or P1CM, expression from pHAL252 was reduced to background levels, showing that SSB-P1 is not expressed during lyogenic growth. These results agree with the finding of Johnson (28) that wild-type P1 does not suppress an *E. coli* ssb mutation. These in vivo results also confirm the in vitro data of Citron et al. (6) showing that Op21 is a functional binding site for the C1 repressor protein.

![FIG. 3. Rescue of an *E. coli* ssb(Ts) mutant. In order to determine if ssb-P1 was essential and sufficient to rescue a temperature-sensitive *E. coli* ssb mutant at 42°C, we cloned the P1 gene under the control of its own promoter into the cloning vector pACYC184. The resulting plasmid, pHAL251, was then transformed into strain KLC436 (51) carrying the ssb-1 allele, con-](http://jb.asm.org/)

<table>
<thead>
<tr>
<th>Strain</th>
<th>Relevant P1 functions</th>
<th>Miller LacZ units</th>
<th>% Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>UT580(pNM481)</td>
<td></td>
<td>51 ± 11</td>
<td>7</td>
</tr>
<tr>
<td>UT580(pHAL252)</td>
<td></td>
<td>685 ± 69</td>
<td>100</td>
</tr>
<tr>
<td>UT580(pHAL252, pAM1)</td>
<td></td>
<td>671 ± 74</td>
<td>98</td>
</tr>
<tr>
<td>UT580(pHAL252, pAM2b)</td>
<td>c1'</td>
<td>309 ± 39</td>
<td>45</td>
</tr>
<tr>
<td>UT580(pHAL252, pAMS)</td>
<td>c1'·bc'</td>
<td>223 ± 36</td>
<td>33</td>
</tr>
<tr>
<td>UT580(pHAL252, P1-15::Tn2680)</td>
<td>c1(Ts); bc</td>
<td>67 ± 12</td>
<td>10</td>
</tr>
<tr>
<td>UT580(pHAL252, P1Cmcr100)</td>
<td>c1(Ts); bc</td>
<td>74 ± 12</td>
<td>11</td>
</tr>
<tr>
<td>UT580(pHAL252, P1Cm)</td>
<td>c1; lxc</td>
<td>86 ± 15</td>
<td>13</td>
</tr>
<tr>
<td>UT580(pHAL252, P1Cm lxc-2)</td>
<td>c1; lxc·</td>
<td>92 ± 18</td>
<td>13</td>
</tr>
</tbody>
</table>

* The P1 functions marked with a prime are plasmid encoded, while all others are located on different P1 prophages. For more information, see Materials and Methods.

* The LacZ values (± standard deviations) are the averages of at least six independent measurements.

* The value for strain UT580(pHAL252) was set to 100%.

<table>
<thead>
<tr>
<th>Phenotype of colonies on X-Gal plates at:</th>
<th>30°C</th>
<th>42°C</th>
</tr>
</thead>
<tbody>
<tr>
<td>White</td>
<td>Blue</td>
<td>Blue</td>
</tr>
<tr>
<td>White</td>
<td>Blue</td>
<td>Blue</td>
</tr>
<tr>
<td>White</td>
<td>Blue</td>
<td>White</td>
</tr>
<tr>
<td>White</td>
<td>Blue</td>
<td>Blue</td>
</tr>
</tbody>
</table>
FIG. 4. Multiple sequence alignment of single-stranded DNA-binding proteins. The corresponding database accession number is given at the right side of each sequence (sp, Swiss-Prot; gb, GenBank; gbu, GenBankUpdate; pir, Protein Identification Resource). The secondary-structure elements shown at the top of the figure are in accordance with the three-dimensional structure of the E. coli SSB protein (Protein Database Brookhaven PDB 1KAW) (41). Residues which are conserved in more than 70% of the sequences are highlighted by colors. Hydrophobic residues are shown in green, glycine and proline are in yellow, serine and threonine are in brown, aromatic residues important for DNA-binding are in red (with numbering according to the E. coli sequence), negatively charged residues and their amines are in violet, and positively charged residues are in blue. The brackets labeled A to C group sequences of bacterial, plasmid, and bacteriophage origin, respectively.
ferring temperature-sensitive lethality due to rapid cessation of DNA replication (14). Overnight cultures of this strain and appropriate controls were grown at 30°C and then spread in duplicate in a sector of an agar plate. The duplicates were incubated overnight at 30 and 42°C. Figure 3 shows that all strains grew at the permissive temperature of 30°C. The KLC438 wild-type parent of KLC436 was not temperature sensitive, while KLC436, as expected, did not grow at 42°C. The lysogenic strain KLC436 (P1Cm), carrying a wild-type P1, was also temperature sensitive, while rescue at high temperature was observed in the presence of P1Cm lxc-2, confirming the results of Johnson (28) and Rosner (42). Growth at 42°C was observed in the presence of pHAL251, but not in the presence of the parent plasmid, pACYC184, indicating that the cloned P1 ssb gene was both essential and sufficient to rescue an E. coli mutant carrying the ssb-1 allele.

Alignment of SSB proteins. Searching the GenBank and Swiss-Prot databases we found many homologues of the P1 ssb protein. These homologues were encoded by bacteria, mitochondria, a number of broad- and narrow-host-range plasmids, and other bacteriophages with host spectra different from that of P1 (data not shown). A multiple sequence alignment including only a relevant subset of P1 SSb homologues is shown in Fig. 4.

There are two highly conserved portions of the SSB proteins. The first includes the amino-terminal amino acids 1 to 115 (the numbers are in reference to the E. coli sequence [Fig. 4]) and corresponds to the DNA-binding domain (30, 56). This domain is also sufficient for the essential tetramerization of the SSb protein, as was demonstrated for an E. coli SSb proteolytic fragment, SSbc, and several deletion mutants (30, 56). The second conserved portion, located at the carboxy-terminal ends of the proteins, is a relatively short stretch of 15 to 23 amino acid residues, containing a remarkably conserved acidic patch located in the last five amino acid residues. According to recent data, this domain is involved in direct interactions between SSb and the χ subunit of the E. coli DNA polymerase III (29).

Intensive studies of a number of SSb homologues demonstrated that they were analogous in many ways (8, 36). The crystal structures of two members of the SSb family have been determined (41, 57), and their comparative analysis supported the idea that the members of the SSb family use common structural principles in order to bind to single-stranded DNA (41).

DISCUSSION

Several studies reporting mutational analyses of bacteriophage P1 failed to identify ssb-P1 (48, 49, 54, 55). Our result showing that the SSb proteins of P1 and E. coli are functionally interchangeable might account for this failure, as any mutations in ssb-P1 might well go unnoticed when assayed in E. coli. Also, several attempts to clone ssb-P1 failed (47), perhaps due to the close proximity of ssb-P1 and lxcA, the gene coding for the P1 lysozyme (46). It was reported that even weak expression of lxcA is very deleterious to host cells (46). While lxcA in its natural context is expressed from a P1-specific late promoter sequence (33) and thus is not expressed in the absence of the phage-specific activator protein gp10 (32), indirect low-level expression from a promoter in the cloning vector might be sufficient to kill cells containing a plasmid carrying the lxcA gene. In the presence of P1 repressor proteins, the inadvertent expression of the lysozyme might be prevented by the C1-Lxc repressor complex binding to Op2 (Fig. 2). However, under such conditions the ssb-P1 gene will not be expressed and thus might again go unnoticed. Only after we managed to separate ssb-P1 from lxcA was it possible to analyze the function of the former gene.

The P1 ssb gene is located in close proximity to the resident IS7 element, and thus it can be speculated that P1 obtained the gene during a transposition event. However, the very strict and phage-specific regulation of ssb-P1 argues against a recent acquisition of the gene by the phage. The expression of SSb-P1 exclusively during lytic growth indicates a function of the protein related to vegetative DNA replication. Some bacteriophages, like T4, T7, and φ29, specify a complete set of replication proteins and are therefore independent of the host replication machinery (31). Unlike these phages, bacteriophage P1 does not specify a complete set of replication proteins, as its vegetative replication depends on DNA polymerase III (DnaE) and primase (DnaG) activities of the host (18). Nevertheless, P1 does specify several replication-associated proteins, like the lytic replication initiator protein RepL (7), a DnA-like helicase (9), a Dam methyltransferase (13), and a homologue of the theta subunit of DNA polymerase III (34), in addition to SSb-P1. These proteins, with the exception of RepL, are homologous to the respective E. coli proteins and thus appear redundant. Indeed, it has been shown that the ssb genes of several conjugal plasmids are dispensable (11, 16, 23). However, the strong conservation of key residues important for SSb function indicates that the maintenance of ssb has to have some selective advantage for the phage or the plasmids. It is conceivable that subtle differences between the proteins of the episome and the host might allow the former to exert specific control over key regulatory steps during vegetative replication or conjugation. Alternatively, it cannot be excluded that P1 or the analyzed conjugal plasmids might encounter host bacteria which differ considerably from E. coli, and in such a host the SSb proteins expressed by the episomal genetic elements might well turn out to be essential.

That highly homologous ssb genes are encoded by both gram-negative and gram-positive bacteria, as well as by some of their plasmids and bacteriophages, raises some evolutionary questions about the possible origin of the gene and the mechanisms by which it is disseminated. A careful phylogenetic analysis might provide some answers to such questions.

ACKNOWLEDGMENTS

We thank J. Lee Rosner for personal communications and for the phage P1Cm lxc-2, Solvej Oestergaard and Finn K. Vogensen for making the complete nucleotide sequence of the ssb gene of TP901-1 available to us prior to publication, and Mathias Velleman and the late Heinz Schuster for personal communications.

This work was supported by a grant from the Statens Naturvidenskabelige Forskningsråd to H.L.

REFERENCES


Rosner, J. L. Personal communication.


Schuster, H., and M. Velleman. Personal communication.


