Localization of F plasmid SopB protein to positions near the poles of Escherichia coli cells

( chromosome partitioning / gene silencing / green fluorescent protein)

SOOK-KYUNG KIM AND JAMES C. WANG*

Department of Molecular and Cellular Biology, Harvard University, 7 Divinity Avenue, Cambridge, MA 02138

Contributed by James C. Wang, December 16, 1997

ABSTRACT The subcellular localization of the SopB protein, which is encoded by the Escherichia coli F plasmid and is involved in the partition of the single-copy plasmid, was directly visualized through the expression of the protein fused to the jellyfish green fluorescent protein (GFP). The fusion protein, but not GFP itself, was found to localize to positions close but not at the poles of exponentially growing cells. Neither the presence of other F-encoded proteins nor the binding of SopB to its recognition sites within the sopC locus of F is required for this localization. Examination of derivatives of the fusion protein lacking various regions of SopB suggests that the signal for the cellular localization of SopB resides in a region close to its N terminus. It is plausible that the near polar localization of SopB may serve the function of keeping a segregated pair of F plasmids apart while the cell septum is being formed. The plausible relation between the specific location of SopB and its suppression of sopC-linked genes when overexpressed is also discussed.

The Escherichia coli fertility or F plasmid encodes two proteins, SopA and SopB, that are known to be required for the stable inheritance of the plasmid (1–3). The 323-residue SopB protein is a DNA-binding protein that specifically binds to the F plasmid sopC locus, which is composed of 12 tandem imperfect 43-bp repeats. Binding of purified SopB to sopC is not dependent on SopA, a 388-residue protein with a DNA-dependent ATPase activity. SopA is involved in the autoregulation of the sopAB operon and appears to bind, by itself or perhaps together with SopB, to four repetitive sequence elements in the regulatory region of the sopAB operon.

Proteins that are structurally and functionally similar to SopA and SopB are known to be encoded by other low-copy-number plasmids such as the P1 prophage (termed the ParA and ParB proteins, respectively, see refs. 4–6). Bacterial chromosomal genes homologous to sopA/parA and sopB/parB have also been identified in Bacillus subtilis, Pseudomonas putida, and Caulobacter crescentus (7–11). In B. subtilis, the subcellular localization of the SopB/ParB homologue SpoOJ was recently reported (10). Immunofluorescence microscopy and visualization of the protein fused to the jellyfish green fluorescent protein (GFP) showed that SpoOJ exhibited a bipolar localization in growing cells dividing symmetrically and in sporulating cells dividing asymmetrically (10). Studies by immunomicroscopy also showed that after the completion of DNA replication in C. crescentus, the ParA and ParB proteins are localized to both poles of the predivisional cell (11). In E. coli, fluorescence in situ hybridization experiments revealed that mini-F plasmid molecules carrying the sopABC region have a midcell location in new born cells but subsequently migrate to positions at one-fourth and three-fourths of cell length, which are termed the quarter-cell positions. In contrast, a mini-F plasmid lacking the sopABC segment appeared to be randomly distributed in cellular spaces not occupied by nucleoids (12). In combination, these findings suggest that the SopB family of proteins and perhaps the SopA-family of proteins as well are involved in anchoring chromosomes to unique cellular locations and that this anchoring may have a significant role in chromosome partition.

The anchoring of B. subtilis chromosomes to unique cellular locations was also examined by direct visualization of the location of the replication origin in living cells (13). In these experiments, tandem copies of the operator for the E. coli lactose operon were inserted near the replication origin, and visualization of this tagged site was made possible by the binding of lac repressor molecules fused to GFP. In cells expressing the fusion protein, two intense fluorescent spots were found to be preferentially located toward the cell poles in sporulating bacteria and in growing cells early in the cell cycle (13). Similar experiments in E. coli cells also showed that the chromosomal replication origin oriC is specifically localized at or near the cell poles, and upon duplication, one copy would move to the site of new pole formation near the site of cell division at midcell (14). P1 and F plasmids, on the other hand, were found to initially localize to the midcell position of newly divided cells but then rapidly migrate upon duplication to the quarter-cell positions (14).

The F plasmid SopB protein has also been implicated in a phenomenon termed the IncG type incompatibility. Inheritance of a single-copy F or E. coli oriC replicon is compromised by the expression of a high level of SopB if the replicon carries the sopABC locus (15). Similarly, a higher than normal cellular level of the ParB protein of P1 prophage can destabilize plasmids carrying parS, which contains a cluster of ParB binding sites (16). Two findings that are likely to be related to the IncG type incompatibility were made more recently. It was found that in E. coli cells expressing a high level of SopB, plasmids bearing either the complete sopC element or as few as one single 43-bp sopC motif have a much higher linking number than the same plasmids without the sopC motif (17–19). Furthermore, in such cells either chromosomally located or plasmid-borne genes that are linked to sopC were found to be silenced; the distance over which sopC can exert this silencing effect is at least 10-kb (20). Examination of the sites of methylation by the dam methylase and the sites of cleavage by DNA gyrase within the silenced domain indicated that the domain is inaccessible to these enzymes (20). Two interpretations of these observations were suggested (20). The first postulates that a nucleoprotein structure emanating from sopC is formed in the presence of a high level of SopB, and the second postulates that SopB mediates the sequestration of a large segment of sopC-linked DNA. In either case, the repres-

Abbreviations: GFP, green fluorescent protein; IPTG, isopropyl β-D-galactopyranoside.

*To whom reprint requests should be addressed.
sion of sopC-linked genes in the presence of a high level of SopB could destabilize the inheritance of single-copy plasmids bearing sopC, leading to the IncG type incompatibility (15). Although the nucleoprotein formation hypothesis appeared attractive, binding of purified SopB to sopC-bearing DNA does not extend beyond the sopC region even at a very high concentration of the protein (21).

To examine further the mechanism by which SopB suppresses the expression of genes linked to sopC, we constructed a plasmid expressing a SopB–GFP fusion protein. Fluorescence and phase-contrast microscopy revealed that the protein is localized to specific positions near but not at the cell poles. Binding of the protein to the sopC locus is unnecessary for this localization, as preferential localization of the protein is neither dependent on the presence of sopC in intracellular DNA nor on the presence of the carboxyl region of the protein known to be required for DNA binding (21). In combination with recent results on SopB and SopA type proteins (10–14), these findings suggest that the binding of a SopB-like protein to a DNA has a major role in targeting the DNA to specific cellular locations. Examination of GFP-tagged truncation derivatives of SopB showed that the specific cellular localization of SopB requires the N-terminal portion of the intact protein, which is known to be necessary for its gene-silencing activity (21).

MATERIALS AND METHODS

Materials. F− E. coli K-12 HB101 was obtained from a commercial source. Strain ASL1270, a derivative of HB101, was constructed by the insertion of a DNA segment containing the sopC sequence and a chloramphenicol-resistance marker into the chromosomal attB locus (20). Plasmids bearing DNA segments encoding the jellyfish GFP were obtained from C. D. Webb and R. Losick (Harvard University), A. Belmont (University of Illinois, Urbana), and a commercial source (Life Technology). Oligodeoxynucleotide primers and other reagents were purchased from various suppliers.

Construction of Plasmids. For the inducible expression of the SopB–GFP fusion protein, a 980-bp NdeI–SacII fragment containing the sopC coding region was excised from pET11b-hmk (21) and purified by agarose gel electrophoresis. A pair of oligodeoxynucleotide primers 5′-CTCCCGCGGTT(TCA-GAGCAAAGGCGGAGGAATC)3′ and 5′-AAAACTGTACAGCTCTGCTCAT-3′, in which sequences derived from the GFP coding sequences are in parentheses, were used to amplify the 720-bp GFP coding region by the PCR. The PCR product was digested withSacII and PstI restriction endonucleases, which cut within the primer sequences (underlined hexamers above), and the DNA fragment was purified by agarose gel electrophoresis. The plasmid pASLR2 (17), constructed for the expression of SopB, was digested with NdeI and PstI, and the backbone of the expression plasmid was gel-isolated. The three purified fragments were then joined to give pSopB-GFP, in which the fused SopB–GFP codons are placed under the control of Ptac, a hybrid trp–lac fusion promoter.

A set of four plasmids for expressing C-terminal truncations of SopB fused to the jellyfish protein were constructing by first cutting pSopB-GFP with SpeI, BspHI, Earl, or AlwNI. Synthetic oligonucleotide linkers were then ligated to the ends of the linearized DNA samples for the in-frame fusion of the truncated SopB coding sequence to the SacII end of the GFP coding region in pSopB-GFP. In each case, after linker ligation the DNA was cut with AgeI and SacII, and the fragment containing the truncated SopB was used to replace the AgeI to SacII fragment in pSopB-GFP. The synthetic linkers used in placing a SacII site downstream of SpeI and BspHI were, respectively, 5′-CTAGAGCGCCGCGCCTTACGCCCGCT-3′ and 5′-CATGCCGGCGCTTACGCCCGG-3′; these linkers were designed to form hairpin structures with the desired 5′-overhangs for joining to the SpeI and BspHI ends. In other constructions, the Earl and AlwNI ends were first treated with the Klenow fragment of E. coli DNA polymerase I, in the presence of all four deoxynucleoside triphosphates, and 5′-GCCGGCGCGC-3′ SacII linkers were then ligated to the repaired ends to add a SacII site. The mutant in which the first 71 codons of SopB was replaced by those coding for a 14-aa peptide Met-Ala-Ser-Met-Thr-Gly-Gly-Gln-Met-Gly-Arg-Gly-Ser was constructed by replacing the small AccI–SpeI fragment of pSopB-GFP by a corresponding fragment from a plasmid previously constructed for the expression of the same 14 residues fused to residues 72–323 of SopB (21).

Visualization of the SopB–GFP Fusion Protein. E. coli ASL1270 and HB101 cells harboring pSopB-GFP were grown to about 107 cells per ml in Luria broth containing ampicillin (100 μg/ml). Each culture was split into two equal portions and isopropyl β-D-thiogalactopyranoside (IPTG) was added to one to a final concentration of 0.1 mM to induce the expression of SopB–GFP. After 2 hr, both IPTG-induced and uninduced cultures were examined in an Olympus BX60 fluorescence microscope equipped with a digital charge-coupled device camera (Princeton Instruments, Trenton, NJ), as described (13, 14). The fluorescence-phase method of Hiraga et al. (22) was used to view the cells. As a control, E. coli ASL1270 cells harboring pGreenLanternTM-I (Life Technology), which expresses jellyfish GFP, were also examined.

RESULTS

Expression of a Functional SopB Protein with a GFP Fused to Its C Terminus. To visualize the subcellular localization of the F plasmid SopB protein, we used gene fusion to link the widely used jellyfish GFP to the C terminus of SopB. Fig. 1 depicts the 6.4-kb pSopB-GFP constructed for the expression of the SopB–GFP fusion protein in E. coli cells.

Functionality of the fusion protein was tested by using the previously observed ability of SopB in silencing genes that are linked to sopC, a cluster of sites to which SopB binds specifically (20). As observed previously, we found that for HB101 cells bearing a pair of plasmids ptacSopB, which expresses SopB from the tac promoter, and pASLS4, which carries the tetracycline-resistance and chloramphenicol-resistance markers and the sopC sequence, induction of SopB expression by the addition of IPTG led to repression of both sopC-linked genes.
drug-resistance markers, and colony formation on chloramphenicol or tetracycline plates was abolished. Under the same conditions, the doubly transformed cells could form colonies on ampicillin plates, because of the absence of the sopC locus on the plasmid carrying the ampicillin-resistance determinant. Identical results were obtained when ptacSopB was replaced by pSopB-GFP in the above experiments, indicating that the SopB-GFP fusion protein is functional in vivo. The functionality of the fusion protein was further confirmed by the use of a strain ASL1270, which was derived from HB101 by the insertion of a DNA segment containing sopC and a chloramphenicol-resistance marker into the attB locus of the E. coli chromosome (20). In the presence of IPTG, colony formation of ASL1270 cells transformed with either ptacSopB or pSopB-GFP was observed on plates containing ampicillin but not on plates containing chloramphenicol.

**Visualization of the Subcellular Localization of the SopB-GFP Fusion Protein by Fluorescence and Phase-Contrast Microscopy.** The subcellular localization of SopB-GFP was viewed by the method of fluorescence-phase combination microscopy (22). In cells overexpressing SopB-GFP, the fusion protein was specifically located to regions close to but not at the cell poles (Fig. 2). Cells bearing pSopB-GFP grown in the absence of IPTG also showed a similar pattern of localization of the fluorescent protein, although the signal was not as strong (result not shown; see also ref. 14). Cells expressing GFP itself showed a uniform distribution of the fluorescent protein (result not shown), indicating that the observed localization of SopB-GFP reflects a property of the SopB protein and not that of the GFP fused to it.

For a population of cells growing exponentially, a single fluorescent dot was usually observed in the smaller cells and a pair of widely separated fluorescent dots was observed in the longer ones (Fig. 2A). The positions of these dots from the cell poles, expressed as fractional cell length, were measured for cells of different lengths and the results are summarized in Fig. 2B. It can be seen from these measurements that in smaller cells showing one fluorescent spot, the position of the spot is close to the middle of the cell; in longer cells with a single spot, the spot is usually located more closely to one of the poles. In cells showing a pair of fluorescent spots, the spots are typically present at about 0.2 cell length from each of the poles.

**Involvement of the N-Terminal Region of SopB in Cellular Localization.** Previous studies indicate that the binding of SopB to specific sites within the sopC locus involves mainly the C-terminal half of the polypeptide (21). The silencing activity of SopB is critically dependent on, however, both its ability to bind sopC and the presence of amino acid residues near its N terminus (21). To deduce which regions of SopB are involved in its specific subcellular localization, plasmids derived from the 6.4-kb pSopB-GFP were constructed for the expression of various N- and C-terminal truncations of SopB fused to the jellyfish GFP (Fig. 3). In each case, immunoblotting with antibodies against the jellyfish protein showed that a polypeptide of the expected size was expressed upon induction of the cells harboring the plasmid by IPTG (results not shown).

Visualization of the cells expressing the various fusion proteins by fluorescence-phase microscopy revealed that deletion of the C-terminal 49 or 79 residues, which is known to abolish binding of SopB to its sites within the sopC locus (21), had little effect on the cellular localization of the protein (Fig. 4 compare B and C to A, for an E. coli cell expressing full-length SopB fused to GFP). In contrast, localization of SopB-GFP to positions close to the poles was no longer observed upon replacing the N-terminal 71 residues of SopB with an insertion of 14 residues unrelated to SopB.

**Fig. 2.** (A) Fluorescence-phase micrographs illustrating the specific cellular localization of the fluorescent SopB-GFP fusion protein. (Bar = 1 μm.) (B) The positions of the green fluorescent spots measured in cells of different lengths. Measurements from cells with a single spot; ○, measurements from cells with a pair of widely separated spots. For each pair of spots, the cell pole with a shorter separation from the proximal fluorescent spot was arbitrarily selected as the origin in the measurement of the distances. In these measurements, an elongated cell with a pair of fluorescent spots was considered as a single cell if no midcell indentation was visible (see the example shown in the upper left corner of A) but as two cells with a single spot in each if a midcell indentation was evident (see image to the right of the cell depicted in the upper left corner).

**Fig. 3.** Illustrations of full-length SopB protein and its truncation derivatives fused to the jellyfish GFP. The number above the SopB and GFP junction in each sketch indicates the particular amino acid residue of SopB at the junction. In the SopB truncation shown at the bottom, the N-terminal 71 amino acid residues of SopB is replaced by a stretch of 14 residues unrelated to SopB.
Recent studies of *E. coli* F plasmid, by either fluorescence in situ hybridization (12) or the introduction of a cluster of lac repressor binding sites into a DNA for the detection of the DNA in the presence of GFP-linked lac repressor molecules (14), indicate that the plasmid is localized to the midcell position in newborn cells but later migrates to positions approximately one-fourth and three-fourths of cell length. It has been suggested that the location of the plasmid to these quarter-cell positions may be related to the presence of the periseleral annular structures at these positions (12, 23). The preferential localization of *E. coli* nucleoids to the same positions was also suggested by experiments examining 4’6-diamidino-2-phenylindole-stained cells (23). In *E. coli* cells treated with a protein synthesis inhibitor, the replicated chromosomes remain located at midcell but move rapidly to the quarter positions upon removal of the inhibitor (23). More recent experiments using the polymerase-taggering method indicate, however, that the *E. coli* chromosomal replication origin *oriC* is preferentially located to positions very near the cell poles and not at the quarter-cell positions (14). The pole-proximal subcellular localization of SopB/ParB and SopA/ParA proteins in *B. subtilis* and *C. crescentus* was also reported (10, 11), as mentioned above.

In the experiments described in this communication, the use of *E. coli* F plasmid SopB protein with a fluorescent protein fragment fused to it revealed that the SopB protein itself is capable of localizing to cellular sites close to but not at the cell poles. The locations of the subcellular SopB protein binding sites are consistent with the quarter-cell locations of F plasmid reported (12, 14), but the low resolution of optical microscopy does not allow a more rigorous conclusion; the overexpression of SopB protein or its derivatives was also found to increase the average cell length in our experiments (see Fig. 2), and this length increment may complicate the interpretation of the subcellular positions measured as fractional cell length. Nevertheless, in combination with previous results, data reported in this work suggest that the specific localization of SopB is responsible for the subcellular localization of DNA containing sopC and that the SopA protein is unnecessary for the positioning of sopC-containing DNA.

The signal for the SopB-protein-mediated subcellular localization appears to require the N-terminal portion of the polypeptide, because the absence of the N-terminal 71 amino acids abolishes this specific localization. The core of SopB sufficient for its specific localization remains to be determined, however. Homology alignment of SopB-like proteins indicates that the region of SopB spanning amino acid residues 85–215 exhibits the strongest homology with other members of this family (21). Interestingly, the beginning portion of SopB, in the absence of the bulk of the protein, appears to localize to positions much closer to the cell poles than the positions assumed by intact SopB or SopB fragments with longer N-terminal regions. The significance of the localization of the short N-terminal peptides of SopB to the polar caps is unknown. A mini-F plasmid lacking the sopC locus, for example, was also found to be preferentially localized to the cell poles, and this localization was interpreted in terms of a preference for cytosol spaces that are not occupied by the nucleoids (12). As GFP itself appears to be distributed more or less uniformly along the entire cell length, it is unclear why the addition of the N-terminal 82 or 107 residues of SopB to N terminus of GFP leads to localization of the protein to the polar caps.

The particular cellular entity (or entities) that interacts with the N-terminal region of SopB remains unknown, and screens using yeast two-hybrid systems (24, 25) yielded no meaningful positive clones (S.-K.K. and J.C.W., unpublished results). The specific subcellular location of SopB suggests that its receptor is membrane-associated, and this interpretation is consistent with a report on the association of SopB with an *E. coli* membrane fraction in the presence of magnesium ions (26).

Because a high level of the SopB–GFP fusion protein was expressed in our experiments, the presence of a pair of bright fluorescence spots at specific locations in the majority of cells suggests that a large number of SopB receptors are present at these locations. This in turn suggests that SopB-mediated localization of F plasmid is probably not a part of the active process that drives the equipartition of a pair of newly replicated F molecules: SopB-mediated distribution of a pair of newly replicated mini-F plasmids to a large number of SopB receptors at these locations would predict that only one-half of all cells would have two plasmids poised for proper segregation upon the formation of a midcell septum. It seems more likely that the SopB-mediated plasmid localization serves a passive role of keeping the segregated plasmids in the proper holding areas until septum formation has been completed.

The SopB-mediated localization of DNA bearing sopC is likely to be related to the silencing of sopC-linked genes in the presence of a high cellular level of SopB (15, 20). Both processes are abolished by replacing the N-terminal 71 residues of the protein (ref. 21 and this work). Examination of a point mutant of SopB incapable of effecting gene silencing also showed that the mutant protein, when fused to GFP, no longer localized to the pole-proximal positions (R. Hanai, personal communication).

We thank R. Losick, C. Webb, and J. Chant for their help in fluorescence microscopy and R. Hanai for helpful discussions. This work was supported by a grant from the National Institutes of Health (GM24544) and a postdoctoral fellowship (to S.-K.K.) from the Korea Research Foundation.