

The Scc Spirochetal Coiled-Coil Protein Forms Helix-Like Filaments and Binds to Nucleic Acids Generating Nucleoprotein Structures

Khalil Mazouni,¹ Gérard Pehau-Arnaudet,² Patrick England,³ Pascale Bourhy,¹
Isabelle Saint Girons,¹ and Mathieu Picardeau^{1*}

Laboratoire des Spirochètes,¹ Plate-Forme de Cryomicroscopie Moléculaire,² and Plate-Forme de Biophysique des Macromolécules et de leurs Interactions,³ Institut Pasteur, Paris, France

Received 18 August 2005/Accepted 12 October 2005

The analysis of the genome of *Leptospira* spp., a group of bacteria of the phylum of spirochetes with several unique evolutionary and morphological features, has allowed the identification of a gene encoding a coiled-coil protein, called Scc, which is completely unrelated to any other eukaryotic or prokaryotic protein. Since coiled-coil proteins are often key elements of the cytoskeleton, we analyzed the protein Scc, which is a 24-kDa protein composed of a N-terminal coiled-coil domain, a proline-rich intermediate domain, and an acidic tail. The gene *scc* is located in an operon which also contains the genes encoding the initiation factor IF3 and the two ribosomal proteins L20 and L35. In this study, we showed that the presence of the coiled-coil domain was responsible for the polymerization of Scc in helix-like structures, in an ATP-independent manner, in both *Escherichia coli* living cells and in vitro. Analysis of the Scc polymers by electron microscopy showed filaments with a width of 6 to 10 nm, similar to that of eukaryotic intermediate filaments. Scc was also found to bind both RNA and double-stranded DNA without detectable sequence specificity. By electron microscopy, we showed that Scc polymer assembly was affected by the presence of nucleic acids, giving rise to rod-shaped structures with a width ranging from 45 to 155 nm. Finally, *Leptospira biflexa* cells depleted in Scc form small colonies, but the morphology of their helicoidal cell body was not affected. These results provide the first insight into a unique DNA binding filament-forming coiled-coil protein that could play an important role in the subcellular architecture of the spirochetal microorganism.

The eukaryotic cytoskeleton is composed of three major fibrous structures (actin filaments, tubulin-based microtubules, and intermediate filaments) that constitute a dynamic filament network in the cytoplasm. This network is involved in the maintenance of cellular shape, cell division, cell motility, and other cellular processes. In bacteria, despite the observation of internal network-like systems (4), it was the general belief that the cytoplasm of prokaryotic cell was an unstructured entity. By the development of microscopic techniques, our understanding of the organization of the bacterial cytoplasm has been greatly modified over the past few years. Structural and functional homologues of tubulin (FtsZ) and actin (MreB, Mbl, and ParM) in prokaryotes have been described (17). Recently, the third major cytoskeletal element, i.e., intermediate filaments, in bacteria has been identified as well (3). The crescentin protein, CreS, has been shown to confer the cell shape of *Caulobacter crescentus* by polymerization into a helical structure beneath the cytoplasmic membrane. The protein CreS is a α -helical coiled-coil protein. Coiled-coil structures were described >50 years ago, at the time of the discovery of the double-helix model of DNA, as the main structural elements of eukaryotic fibrous proteins such as keratin or myosin (7, 22). The amino acids in a coiled-coil motif reside on seven different structural positions on the coil, forming a heptad periodical repeat (a-b-c-d-e-f-g). The first and fourth positions (a and d) of the heptad repeat are generally apolar or hydro-

phobic amino acids and form an interface that promotes the coiled association between two parallel α -helices (18). Coiled-coil structures are formed by two or more α -helices wound around each other in a superhelix structure. The α -helical coiled-coil is the most widespread subunit oligomerization motif found in proteins. In eukaryotic cells, coiled-coil proteins form large structures, and many of them play essential roles in cytoskeleton structures. Besides crescentin, only a restricted number of cytoplasmic coiled-coil based polymerizing proteins in prokaryotes have been described. For example, AglZ is a filament-forming protein involved in gliding motility in *Myxococcus xanthus* (27). Filamentous polymers are formed by the coiled-coil protein p1 from *Bacillus subtilis* bacteriophage phi29, which is a membrane-associated protein involved in the initiation of viral DNA replication (6). The cytoplasmic filament protein CfpA from the spirochetes of the *Treponema* genus also possesses a coiled-coil region (28).

Leptospira spp. belong to the bacterial phylum of *Spirochetes*, an evolutionary and structurally unique group of bacteria. These bacteria are composed of both saprophytic and pathogenic members, such as *Leptospira biflexa* and *Leptospira interrogans*, respectively (9). Genome-wide analysis of *L. interrogans* (24) by BLAST search (1) using coiled-coil proteins as a query and the COILS program (18) allowed the identification of several putative coiled-coil proteins. After discarding proteins with annotated functions such as structural maintenance of chromosome (SMC)-like and flagellin-like proteins and those with putative transmembrane segments, we identified a small gene whose product, 212 amino acids in length, did not exhibit significant homologies with other proteins in databases. Here,

* Corresponding author. Mailing address: Laboratoire des Spirochètes, Institut Pasteur, 28 rue du docteur Roux, 75724 Paris Cedex 15, France. Phone: 33 (1) 45 68 83 68. Fax: 33 (1) 40 61 30 01. E-mail: mpicard@pasteur.fr.

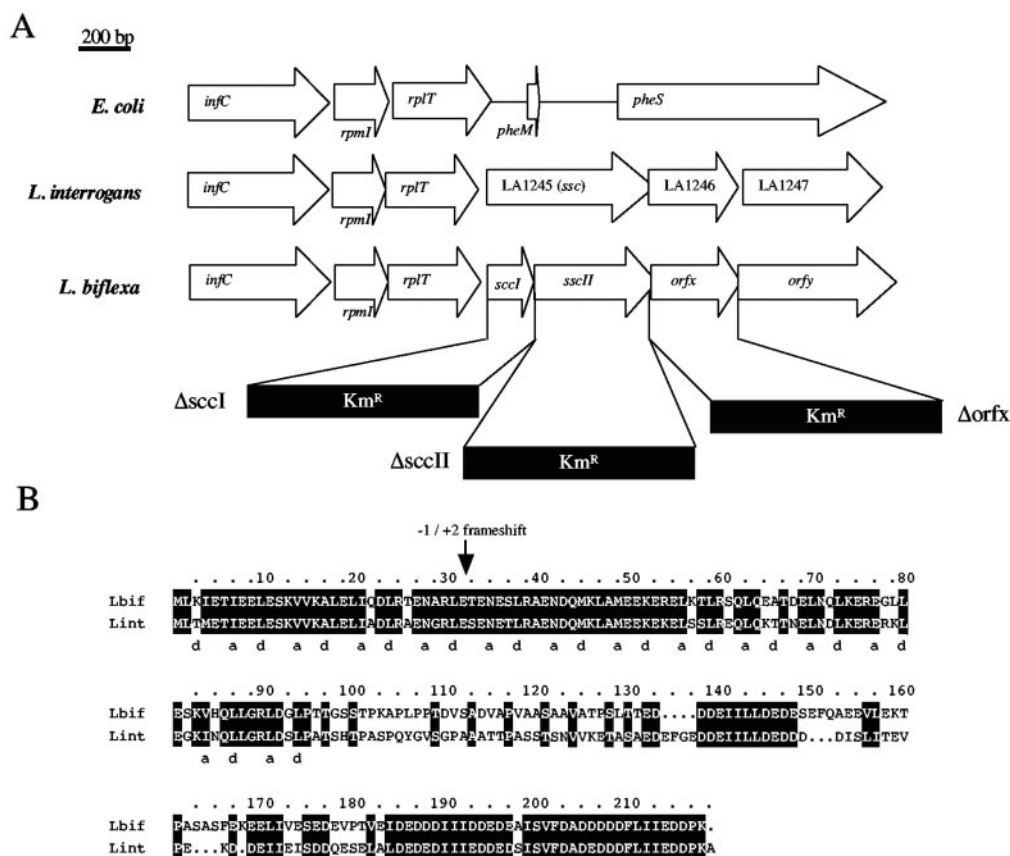


FIG. 1. Scc, a coiled-coil protein from *Leptospira* spp. (A) Genetic organization of the *rplT* locus in *E. coli* and *L. interrogans*. Constructs used for the generation of *L. biflexa* mutants are also indicated. Km^r, kanamycin resistance cassette. (B) Sequence alignment of Scc proteins of *L. interrogans* (Li nt) and *L. biflexa* (Lbif). The arrow maps the frameshift in *L. biflexa*. Residues conserved in the two proteins are shaded. The a and d positions of the heptad repeat of the putative coiled-coil domain are shown below the alignment.

we report the identification and biochemical analysis of this novel structural protein, termed Scc (for spirochetal coiled-coil protein) in *Leptospira* spp. The protein Scc exhibits a tripartite structure which includes an N-terminal coiled-coil domain, followed by an intermediate domain rich in Pro, Ala, Ser, and Thr residues, and an acidic tail (pI of 2.8). Using fusions to the green fluorescent protein (GFP), we demonstrated that Scc proteins can self assemble into helix-like structures in vitro and in living *Escherichia coli* cells. This protein also binds to both DNA and RNA molecules, generating rod-shaped nucleoprotein structures.

MATERIALS AND METHODS

Bacterial strains and growth conditions. *L. biflexa* serovar Patoc strain Patoc I and *L. interrogans* serovar Lai strain Lai (National Reference Center for *Leptospira*, Paris, France) were grown at 30°C in EMJH (8, 13) medium. *Escherichia coli* XL2 (Stratagene, Cedar Creek, TX) and MG1655 strains were grown in Luria-Bertani (LB) rich medium supplemented with 50-μg/ml kanamycin or 100-μg/ml ampicillin when appropriate.

DNA and RNA manipulations. Genomic DNA of *Leptospira* was isolated as previously described (23). Plasmid DNA was purified using the QIAGEN Plasmid Miniprep kit (QIAGEN GmbH, Hilden, Germany). Reverse transcription-PCR (RT-PCR) of RNA was carried out as described by the manufacturer's instructions (SuperScript One Step RT-PCR with Platinum Taq; GibcoBRL, Rockville, MD) with primer pairs (primer nucleotide sequences are available on request) corresponding to selected open reading frames (ORFs) from the *scc* locus as previously described (5).

Identification of the *L. biflexa* *scc* locus and gene inactivation. Recombinant colonies of an *L. biflexa* genomic library were transferred onto a nylon filter (N⁺ Hybond; Amersham Biosciences, Little Chalfont, England) and probed with the radiolabeled *L. interrogans* *scc* gene under low-stringency conditions as previously described (25). DNA of candidate clones was sequenced, thus allowing the identification of the *L. biflexa* *scc* locus (2,511 bp), which was amplified by PCR with primer pairs CcpA (5'-GCAAAGGGTGATAAGGTAAAAGTGACAC TTCG-3') and CcpR (5'-CCTTTCAATGGTAAACGTAAGAAATTTGTTCAA GG-3') and inserted into the pGEM-T vector (Promega), resulting in plasmid pGEMTccp. Blunt-ended restriction sites, SspI and/or SnaBI, were introduced at the 5' and 3' ends of *sccI*, *sccI*, and *orfx* coding sequences (Fig. 1) with the Quick Change mutagenesis kit (Stratagene). A PvuII fragment containing the kanamycin resistance cassette was then inserted into the blunt-ended sites of pGEMTccp to generate plasmids pGK2(Δ*sccI*), pGK3(Δ*sccII*), and pGK5(Δ*orfx*), respectively. Recombinant plasmids pGK2, pGK3, and pGK5, which are not replicative in *Leptospira* spp., were then subjected to UV irradiation and used to deliver the inactivated allele in *L. biflexa* as previously described (23). Kanamycin-resistant colonies were picked and tested for the insertion of the kanamycin resistance cassette in the target gene by PCR and Southern blot analysis as previously described (23). For complementation studies, a transcriptional fusion between the *L. interrogans* *hsp10* promoter and the *L. interrogans* or *L. biflexa* *scc* coding sequences was ligated into the SmaI site of the spectinomycin-resistant *L. biflexa*-*E. coli* shuttle vector pGSLe24 (5). Since the *L. biflexa* *scc* gene contains a frameshift, complementation of mutants was performed with either the *L. biflexa* wild-type *scc* gene or the *L. biflexa* *scc* gene corrected by site-directed mutagenesis as indicated below. Transformants were selected on EMJH plates containing 40-μg/ml spectinomycin. Production of Scc by transformants was confirmed by immunoblot analysis.

Recombinant protein expression and purification. To clone the studied ORF for overexpression in *E. coli*, *scc* coding sequences (regions corresponding to

residues 1 to 212, 1 to 95, 1 to 130, and 95 to 212) from *L. interrogans* were amplified by PCR (nucleotide sequences of primer pairs available on request) and cloned into the SspI site of pCGFPHis (20) for in-frame cloning of the studied ORF upstream of the EGFP-6His gene. The *gfp* gene was obtained from pEGFP1 (Clontech). The nucleotide sequence 5'-CGCCTA GAG AGA-3' of *sccI* (between nucleotide 88 and 98) was changed to 5'-CGC CTA GAG A-3' (the deleted nucleotides are in boldface type) by site-directed mutagenesis (Quick Change mutagenesis kit; Stratagene) of plasmid pGEMTccp, generating plasmid pGEMTccpM, to produce a full-length protein (214 amino acids) in *E. coli*. We also made GFP fusions with the *L. biflexa* Scc bearing the segments at positions 1 to 95, 1 to 130, and 95 to 214 by amplification from plasmid pGEMTccpM and ligation into the SspI site of pCGFPHis as described above. Plasmid constructs for Scc-6His expression were also made. Recombinant proteins were purified with Ni-nitrilotriacetic acid agarose beads as previously described (20) and eluted in purification buffer (PB) (50 mM Tris [pH 7.5] and 150 mM NaCl) containing 500 mM imidazole. Protein concentration was determined with the BCA Protein Assay Reagent kit (Pierce, Rockford, IL) using bovine serum albumin as a standard. In every case, the purity of the studied proteins was >90%, as judged by Coomassie-stained sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) gels.

Nucleic acid binding experiments. For agarose gel electrophoretic mobility-shift assays, increasing amounts (0.5 μ g to 40 μ g) of Scc-6His or Scc-GFP-6His proteins and deletion derivatives were incubated for 10 min in PB at room temperature with either *L. biflexa* total RNA (1 μ g) or circular plasmid pGEM7Zf⁺ DNA (1 μ g). All reactions were carried out in a total volume of 40 μ l. The reaction products were visualized by ethidium bromide staining after electrophoresis for 2 h through 0.8% (wt/vol) agarose gels run in TBE buffer (90 mM Tris-borate [pH 8], 2 mM EDTA). GFP, used as a negative control, was not observed to bind either substrate under these conditions. For surface plasmon resonance (SPR) assays, experiments were performed on a Biacore 2000 instrument equilibrated at 25°C with HBS (10 mM HEPES, 150 mM NaCl, pH 7.2) containing 0.005% Tween 20. The Penta-His monoclonal antibody (QIAGEN) was covalently immobilized on the carboxymethylated surface of the four flow cells of a CM5 sensor chip to a level of 10,000 to 12,000 resonance units, using the Amine Coupling kit (Biacore). Recombinant GFP-6His, *L. biflexa* Scc-GFP-6His, and *L. biflexa* Scc (95-212)-GFP-6His were captured to a level of 400 to 500 resonance units in three independent flow cells of the anti-6His surface. Eight different concentrations of plasmid pGEM7Zf⁺ DNA (7.8 to 1,000 μ g/ml) and of total RNA (1.2 to 300 μ g/ml) were then injected across the four surfaces for 10 min at a flow rate of 20 μ l/min, and the dissociation of the complexes was followed for 5 min. The raw SPR profiles were double subtracted from those measured on the anti-6His reference flow cell and from those obtained by injection of the sample buffer onto the GFP-6His, Scc-GFP-6His, and Scc (95-212)-GFP-6His surfaces.

Antiserum preparation and immunoblot analysis. The *L. biflexa* Scc-GFP-6His protein (100 μ g in 1 ml of phosphate-buffered saline [PBS]) was emulsified 1:1 in Freund adjuvant (Calbiochem-Behring, La Jolla, CA). A total of 100 μ l of the preparation was injected subcutaneously into each of the BALB/c mice. The injection was repeated three to four times at 3-week intervals. For Western blot analysis, proteins separated on 12.5% (wt/vol) SDS-PAGE gels were transferred to nitrocellulose membranes. Blots were incubated with blocking buffer (PBS containing 0.05% Tween 20 and 5% nonfat dried milk) at room temperature for 1 h. After removal of the blocking buffer, antisera diluted 1:1,000 with blocking buffer was added, and the blots were incubated overnight at 4°C. After the blots were washed in PBS containing 0.05% Tween 20, 1:10,000-diluted peroxidase-labeled anti-mouse immunoglobulin G (Amersham) was added, and the blots were incubated at room temperature for 1 h. After three more washes, blots were developed with BCI/NBT Color Development substrate (Promega). Collected serum reacted with *L. biflexa* and *L. interrogans* Scc proteins but did not detect *E. coli* crude protein extracts.

Microscopy. To investigate the in vitro polymerization activity of Scc, GFP fusion samples (1 μ g/ μ l) were incubated for 1 h at 30°C in PB, then observed by fluorescence microscopy. For in cellulose studies, *E. coli* strain MG1655 harboring plasmid constructs expressing full-length or truncated Scc-GFP-6His proteins were harvested from culture media by centrifugation at 4,000 \times g for 4 min, washed once in 10 mM Tris-HCl (pH 8), and resuspended in 10 mM Tris-HCl (pH 8). The cell suspension was then spread on slides and viewed by fluorescent microscopy. Images of *E. coli* cells were captured with a Leica DMRXA microscope equipped with a Roper Scientific Micromax cooled charge-coupled device camera and MetaMorph software (Universal Imaging). For electron microscopy, protein samples (1 μ g/ μ l) were placed onto carbon-collodion-coated copper grids, negatively stained with 2% uranyl acetate (pH 4.5), and examined with a CM12 transmission electron microscope (FEI, Eindhoven, The Netherlands)

operated at 100 kV. All images were recorded on Kodak SO-163 films. For cryoelectron microscopy, 4 μ l of the sample was applied on holey carbon grids, blotted with filter paper, and vitrified in liquid ethane. Specimen were transferred to a Gatan 626 DH cryoholder and examined with a 2010F electron microscope (Jeol, Tokyo, Japan) operating at 200 kV. Images were recorded at a nominal magnification of 50,000 with 1.5 μ m of underfocus under low-dose conditions.

Sequence analysis and nucleotide sequence accession number. The DNA sequence data were analyzed with the BLAST program (1) at the National Center for Biotechnology Information and with the SMART <http://smart.embl-heidelberg.de/> (16), the TMpred http://www.ch.embnet.org/software/TMPRED_form.html (11), and the COILS (http://www.ch.embnet.org/software/COILS_form.html) (18) servers. The nucleotide sequence of the locus containing the *L. biflexa scc* gene was deposited at the National Center for Biotechnology Information under accession number DQ094179.

RESULTS

Sequence analysis of Scc. Computer analysis of the *L. interrogans* genome (24), led us to the identification of ORF LA1245, the product (212 amino acids) of which exhibits an α helix at the N-terminal region with coiled-coil motifs. The coiled-coil prediction of the N terminus (residues 3 to 94) of this protein, using COILS server (18), is high (100%). The α -helical coiled-coil domain continuously spans the first 95 residues of the protein, and it comprises 12 heptads. The positions of the predominantly hydrophobic residues of the heptad repeats characteristics of coiled-coil proteins are shown in Fig. 1. This protein was further referred as Scc. The *scc* gene of the saprophyte *L. biflexa*, which can be genetically engineered, was identified by screening a genomic library with a *L. interrogans scc* probe. Surprisingly, sequence alignment of Scc proteins of *L. interrogans* and *L. biflexa* suggests that the *L. biflexa scc* gene contains a frameshift, leading to the emergence of a premature stop codon. The GAG-AGA mRNA sequence, located 42 nucleotides upstream of the stop codon of *L. biflexa sccI*, may predispose translational (or potentially transcriptional) slippage, producing a -1 or $+2$ frameshift that restores the normal reading frame (Fig. 1). Site-directed mutagenesis was used to artificially restore the normal reading frame, allowing the production of recombinant proteins in *E. coli*. To address the possibility of a frameshift within *L. biflexa scc*, immunoblot analysis was performed on the *L. biflexa* wild-type strain, using antiserum from mice immunized with purified Scc-GFP-6His. The Scc antiserum is reactive with a single band with a molecular mass similar to the 24 kDa of *L. interrogans* Scc (Fig. 2), further suggesting the presence of a frameshift in *L. biflexa scc*.

The putative Scc protein (214 amino acids) of *L. biflexa* (Scc-bi) shares >70% similarity to that of *L. interrogans* (Scc-int). The Scc proteins of both *L. interrogans* and *L. biflexa* can be divided into three regions, based on pI and sequence features. Immediately following the coiled-coil domain is a second domain (from residues 94 to 130), rich in residues Pro, Ala, Ser, and Thr (representing 73% of the residues), called the PAST domain. The C-terminal domain (residues 130 to 212) has a pI of 2.8 and 3 in Scc-int and Scc-bi, respectively. BLAST analysis of the Scc-int coiled-coil region (90 amino acids) shows 57% similarity (50/87 residues) with the myosin-like protein from *Oryza sativa* and 53% similarity (43/80 residues) with the chromosome segregation protein SMC1 from *Synchocystis* sp. PCC 6803. In addition, BLAST analysis of the

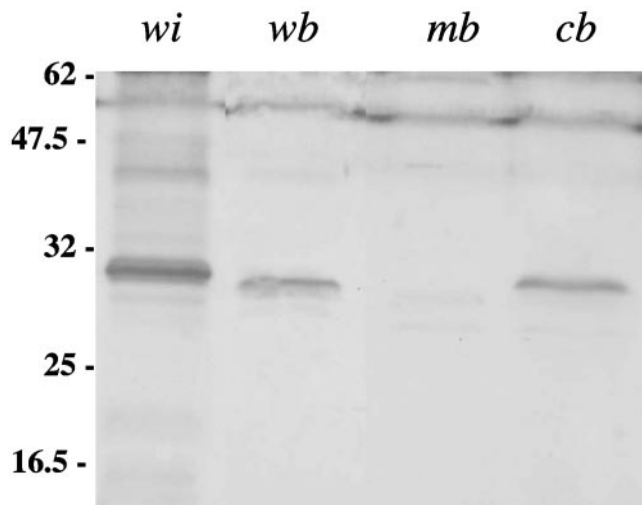


FIG. 2. Western blot analysis of Scc. Total protein from *L. interrogans* wild type (*wi*), *L. biflexa* wild type (*wb*), *L. biflexa* $\Delta sccI$ (*mb*), and *L. biflexa* $\Delta sccI$ -complemented (*cb*) strains was resolved by SDS-PAGE and transferred to nitrocellulose. The membrane was probed as described in Materials and Methods with a polyclonal antibody to *L. biflexa* Scc. Standards are indicated at left (in kilodaltons).

C-terminal domain of Scc-int (residues 130 to 212) exhibits 55% (44/79 residues) and 53% (48/90 residues) similarity with RpoE of *Bacillus licheniformis* and *Staphylococcus aureus*, respectively.

Sequence analysis of the *scc* locus reveals that *scc* is clustered with five other genes in the same direction of transcription, including the *infC-rpmI-rplT* operon, which encodes the translation IF3 protein and the two ribosomal proteins, L35 and L20 (Fig. 1). Downstream of the *scc* gene, two other genes are found in the same direction of transcription. The first gene, *orfx/LA1246*, encodes a protein that exhibits 44% similarity (42/94 residues) with the *Pseudomonas aeruginosa* bacterial cell division protein ZapA. The second gene, *orfy/LA1247*, encodes a putative protein with 40% (74/185 residues) similarity with a 5-formyltetrahydrofolate cycloligase protein from *Yersinia pestis*. The compact organization of *infC-rpmI-rplT* is similar to that of most bacteria such as *E. coli*. In *E. coli*, the gene downstream *rplT* is *pheM*, encoding a leader peptide of 14 residues involved in the regulation of the *pheS* operon (Fig. 1). The gene *pheS* and *pheT* encode, respectively, the small and large subunits of phenylalanyl-tRNA synthetase. Transcriptional analysis by RT-PCR revealed that the *L. interrogans* and *L. biflexa infC-rpmI-rplT-scc-orfx* genes form an operon (data not shown).

Scc forms polymeric structures in vitro and in *E. coli*. Since coiled-coil domains usually mediate oligomerization, we investigated the ability of Scc-bi and Scc-int proteins to form polymeric structure in vitro. Under our experimental conditions, purified polyhistidine-tagged versions of Scc-GFP were purified to homogeneity and allowed to assemble in vitro. After a 1-h incubation, fluorescence microscopy showed that Scc proteins self assembled, in an ATP-independent manner, into helix-like structures, even in water. The average filament length was 2 to 3 μm (Fig. 3A), and polymeric structures were stable for several days. Electron microscopy of the polymeric

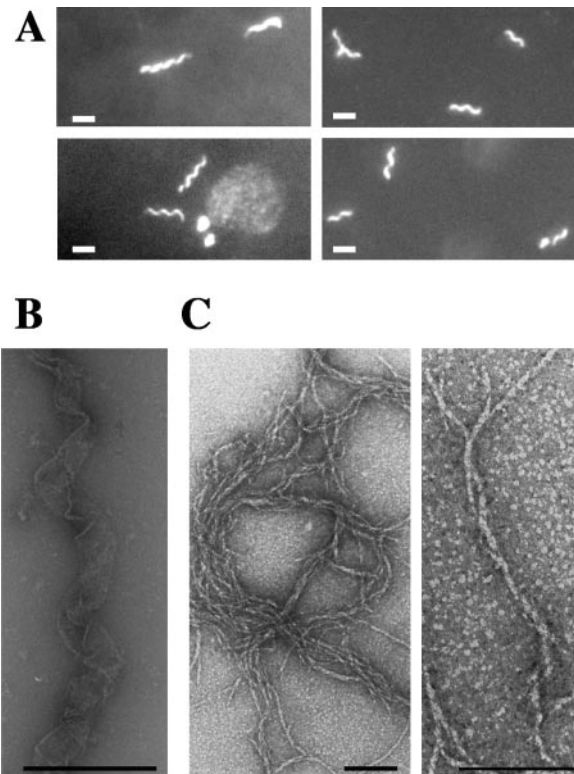


FIG. 3. Scc forms polymeric structures in vitro. (A) Fluorescence microscopy imaging of polymers of *L. biflexa* Scc fused to GFP. Typical filaments are 2 to 3 μm long. Scale bar, 1 μm . (B and C) Electron microscopy of in vitro self-assembly of Scc-6His alone. Helix-like structures (B) and filaments with a cross-sectional diameter of 6 to 10 nm (C) were observed. Similar pictures were observed for Scc-GFP. Purified recombinant Scc-GFP or Scc-6His protein (1 $\mu\text{g}/\mu\text{l}$) was incubated 1 h at 30°C in 50 mM Tris, pH 7.5, and 150 mM NaCl. Scale bar, 500 nm (B) or 100 nm (C).

Scc-6His protein, as well as of the Scc-GFP-6His protein, also revealed helix-like structures (Fig. 3B). In addition, we observed well-separated, individual filaments with a width of 6 to 10 nm that usually adopted a rope-like structure (Fig. 3C). This indicates that the fusion to GFP did not alter Scc polymerization. In the majority of *E. coli* cells (88%), Scc-bi-GFP was colocalized with the membrane, mainly at the inner curvature of the rod-shaped cells (Fig. 4). A helical filament appeared to extend from pole to pole. In dividing cells, two helical filaments appeared to extend from mid-cell to opposite poles, forming pole-to-pole helical structures (Fig. 4). The remaining cells contained a fluorescent focus at one of the poles or no signal.

To define the minimum structural determinant required for polymerization, the size of the Scc-bi and Scc-int proteins was reduced, and the tagged molecules were examined by fluorescence microscopy in vitro and in *E. coli*. Deletion of the C-terminal domain or both the central and C-terminal domains did not abolish filament formation in vitro, but mutants yielded smaller filaments. Similarly, *E. coli* cells expressing Scc (1-95)-GFP and Scc (1-130)-GFP showed fluorescent aggregates at one of the two poles of the rod-shaped bacterium, but no helical filaments (data not shown). This result suggests that oligomers can be formed by interaction of the N-terminal

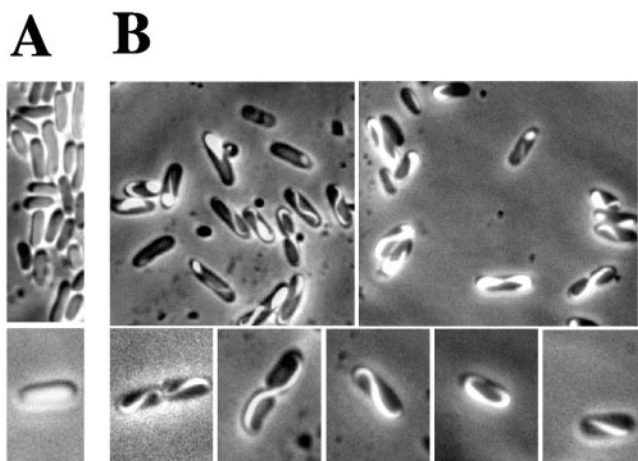


FIG. 4. Localization of Scc-GFP in *E. coli*. (A) *E. coli* cells expressing GFP alone giving rise to a uniform diffuse fluorescence. (B) *E. coli* cells (48-h culture at 30°C) expressing *L. biflexa* Scc-GFP. The top and bottom rows show images at different magnifications.

coiled-coil domain but that other domains may be important for the correct filament assembly. To obtain more evidence for this mode of assembly, we constructed mutants in which the N-terminal domains were deleted. Fluorescence microscopy showed that Scc (95-130)-GFP was not able to form oligomers in vitro. In *E. coli* cells expressing Scc (95-130)-GFP, the fusion protein fluorescence was distributed throughout the cytosol, with a diffuse pattern similar to that observed with control cells expressing the GFP alone.

Scc is a nucleic acid binding protein. We found that purified Scc-6His and Scc-GFP-6His were able to bind heparin agarose beads, unlike GFP alone, indicating that Scc could be a DNA

binding protein. To characterize nucleic acid-protein interactions in solution, a gel shift assay was used. In these experiments, native or truncated proteins were incubated with nucleic acids and then subjected to electrophoresis in an agarose gel. The full-length Scc-bi protein formed retarded complexes that migrated into the agarose gel at protein:RNA ratios as low as 5:1 (Fig. 5). With increasing protein:RNA ratios, increasing proportions of RNA were involved in retarded complexes (data not shown). A similar retardation pattern was found for the Scc (1-130) truncated proteins (Fig. 6), indicating that the observed profile is characteristic of the coiled-coil and PAST domains of the Scc protein. Similarly, incubation of increasing amounts of Scc with plasmid DNA affected the migration of supercoiled DNA, as well as that of relaxed circular or linear plasmid DNA (Fig. 6). Again, the amount of retarded DNA increased with the protein:DNA ratio. Interestingly, the DNA was also retarded by the Scc truncated protein containing the coiled-coil domain alone (Fig. 5). For all the Scc truncated forms, similar results were obtained with the GFP fusion and the His-tagged proteins. The nucleic acid binding capacity of Scc was further demonstrated by surface plasmon resonance assays. DNA and RNA were both able to bind and in a concentration-dependent manner to native Scc, but not to truncated Scc (95-212) or to GFP (Fig. 6).

These experiments prompted us to test whether Scc could generate nucleoprotein structures upon binding to nucleic acids. By electron microscopy, we found that the polymerization of Scc-bi was altered in response to either circular-linear DNA or RNA, generating rod-shaped structures with a width of 45 to 155 nm, and an average diameter of 84 nm (Fig. 7). Similar results were obtained by cryoelectron microscopy, further confirming that these structures are not due to uranyl acetate staining (data not shown). These structures were altered after

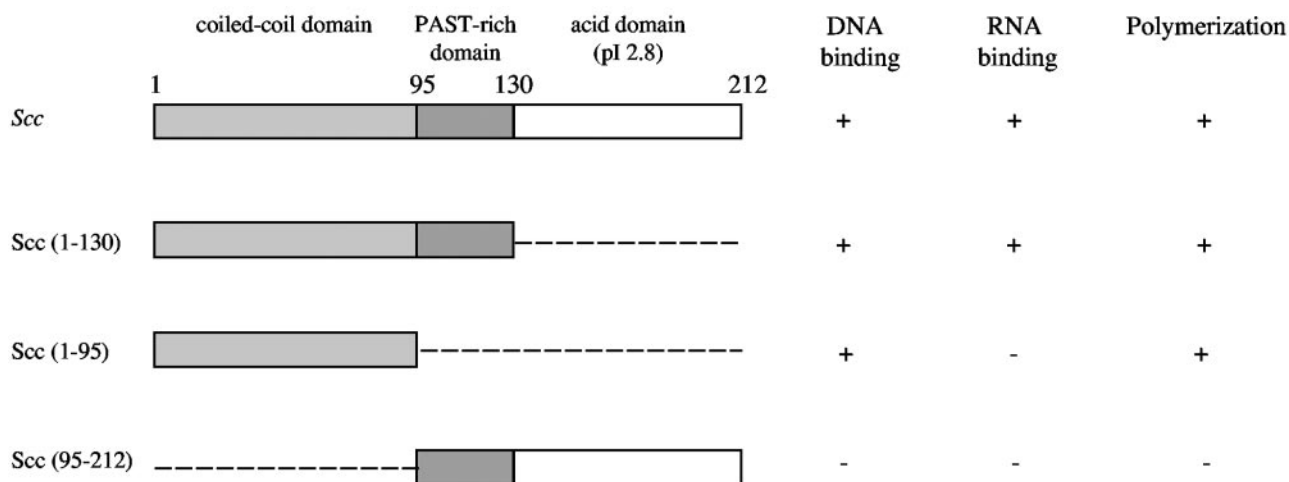


FIG. 5. Schematic representation of Scc deletion mutants and their polymerization and nucleic acid binding abilities. Scc can be divided into three domains, based on amino acid content and pI. The numbers correspond to the amino acid sequences. Binding to nucleic acids was determined by gel shift assays with supercoiled plasmid DNA pGEM7Zf⁺ and total RNA from *L. biflexa* as described in Materials and Methods. Polymerization of GFP fusions was determined by fluorescence microscopy of both *E. coli* cells and in vitro. In addition, electron microscopy was used to determine polymerization of Scc-6His protein derivatives into helix-like filaments. The ability to form rod-shaped structures in the presence of circular plasmid DNA was also examined. -, no nucleic acid interaction or polymerization (when experiments were done with GFP fusions, the results were similar to those obtained with GFP alone); +, nucleic acid interaction or polymerization. Similar results were obtained with Scc from *L. biflexa* and *L. interrogans*.

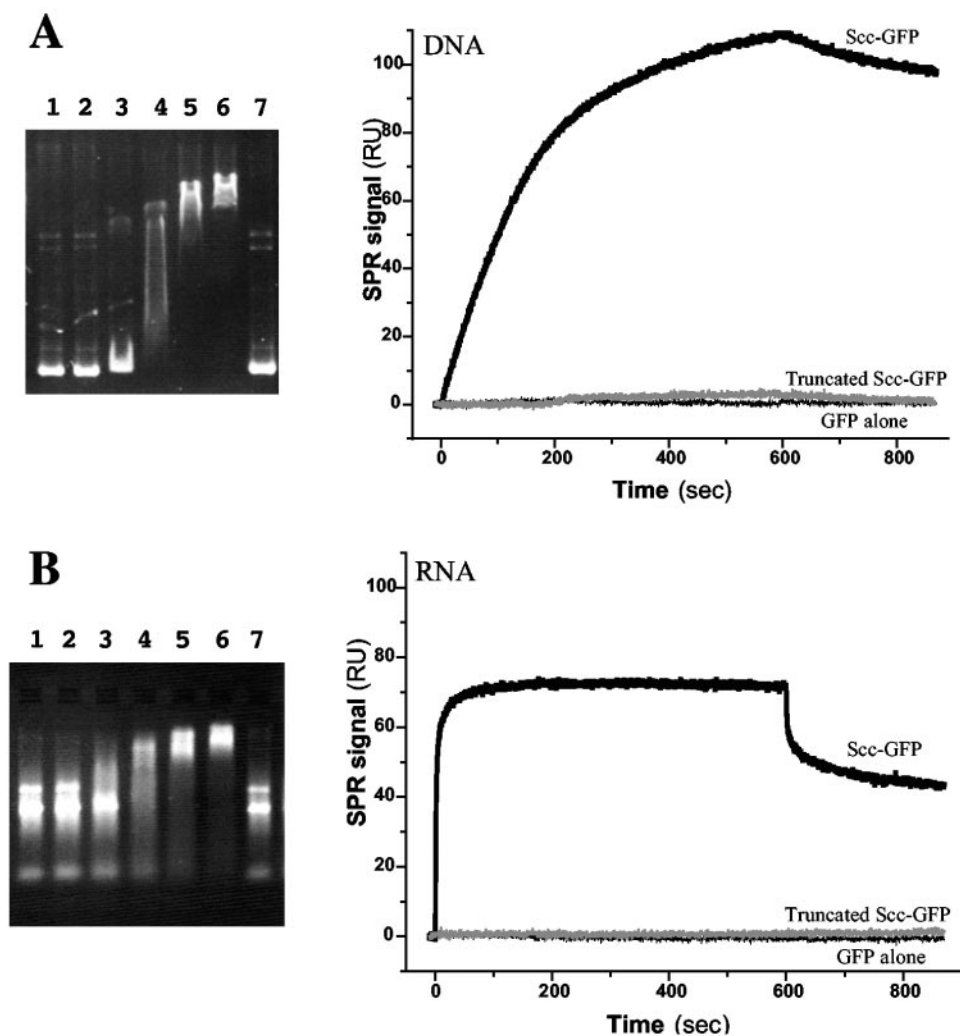


FIG. 6. Nucleic acid binding assays of Scc. (Left) Electrophoretic mobility shift assays of *L. biflexa* Scc. Agarose gel electrophoresis staining with ethidium bromide for visualizing DNA (A) and RNA (B). Lanes 1 and 7, 1 μ g plasmid DNA (A) or 1 μ g total RNA from *L. biflexa* (B). Lanes 2 to 6, 1 μ g of plasmid DNA (A) or total RNA (B) with 2, 5, 10, 20, and 40 μ g of Scc-6His protein. All reactions were carried out in a total volume of 40 μ l. (Right) Real-time surface plasmon resonance (SPR) profiles showing the interaction between immobilized *L. biflexa* Scc-GFP-6His and 250- μ g/ml plasmid DNA (A) or 150- μ g/ml total RNA (B) in solution. No binding was observed on GFP or truncated Scc (95-212)-GFP-6His.

nuclease treatment (Fig. 7B) and not visualized after proteinase K or detergent treatment (data not shown). These data indicate that these structures are effectively composed of protein and nucleic acids. We also examined in-gel-prepared Scc-DNA complexes that had been subjected to electrophoresis. Agarose that was taken for the areas of the gel containing retarded complexes (Fig. 6) was melted and spread for microscopy. Again, rod-shaped structures were seen in these preparations and only in these preparations (data not shown). These results indicate that Scc alone cannot generate such superstructures. The Scc-truncated protein containing the coiled-coil domain alone was also able to generate similar rod-shaped structures in the presence of nucleic acids (Fig. 5).

The *L. biflexa* scc gene is not essential. Since gene inactivation is not feasible in pathogenic *Leptospira*, the *L. biflexa* scc gene was inactivated by allelic exchange. The double-crossover disruptants were confirmed by PCR, Southern hybridization, and immunoblot analysis. Inactivation of *L. biflexa* scc did not

affect cell growth or cell morphology in liquid medium. The plating efficiency (ratio of cell numbers to CFU) of an exponentially growing scc mutant in liquid medium was similar to that seen with the parental strain. However, colony morphology showed a dramatic difference between the scc mutant and wild-type strains, i.e., the scc mutant formed small colonies (Fig. 8). Cells of the scc mutant remained highly motile, as observed by dark-field microscopy, suggesting that Scc is not involved directly in the mechanisms of swimming. The small-colony phenotype was repeatedly obtained upon gene knock-out, suggesting that the phenotype was not due to an unrelated secondary mutation. Although scc is the fourth gene of an operon, RT-PCR showed that the scc mutation did not prevent transcription of genes downstream of the insertion site (data not shown). In addition, inactivation of the scc downstream gene, orfx, resulted in mutants with large colonies on solid medium. Therefore, the phenotype observed with the *L. biflexa* scc mutant cannot be attributed to the transcriptional inacti-

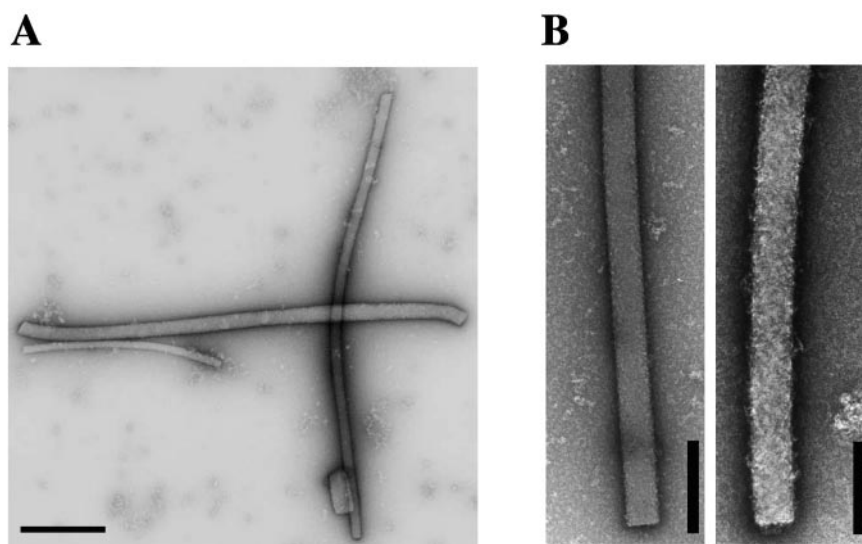


FIG. 7. Electron microscopy of Scc in combination with DNA. (A) Rod-shaped structures observed in samples containing *L. biflexa* Scc with DNA. Scale bar, 500 nm. (B) Electron microscopy of Scc-RNA structures treated with RNase. Rod-shaped structures observed in samples containing recombinant Scc with RNA (left) in the presence of RNase (right). Scale bar, 200 nm.

vation of *orfX*. Introduction of a low-copy-number plasmid carrying the *L. biflexa* or *L. interrogans* wild-type *scc* gene under the control of a spirochetal constitutive promoter into the *L. biflexa* Δscc mutants restored Scc production, as observed by Western blot analysis (Fig. 2), but not a wild-type colony morphology. Although the abolition of the *scc* gene expression was not lethal, the phenotype alterations observed on solid medium underlined the physiological importance of this protein.

DISCUSSION

In eukaryotes, coiled-coils motifs are the main structural elements of fibrous proteins that form the cytoskeleton involved in many cellular processes. Recent studies of bacterial cells suggest that most of the basic constituents of the cytoskeleton are conserved between prokaryotes and eukaryotes. Among the order *Spirochaetales*, cytoskeleton-like systems were observed a long time ago in the cytoplasm of some spi-

rochetal cells by electron microscopy (4). Cytoplasmic filaments that were estimated to be approximately 7 nm in diameter, like Scc filaments, were found in *Treponema* and *Leptospira* spp. (4). Further work is necessary to ascertain if these filaments are related to the protein Scc. Furthermore, a filamentous ribbon-like structure composed of a unique 82-kDa protein, CfpA, is present in the cytoplasm of *Treponema* cells. These cytoplasmic filaments are involved in chromosome structure, segregation, and/or in the cell division process (12).

In this study, we show that the protein Scc from both the saprophyte *L. biflexa* and the pathogen *L. interrogans* can spontaneously self assemble into helical filaments via its coiled-coil domain. The most distinctive property of the Scc protein is its ability to form filaments without a requirement for energy, cofactors, or other exogenous factors. This characteristic was also found in the phage phi29 p1 protein (6) and in the *C. crescentus* crescentin protein, which forms helix-like structures (3) similar to those of Scc. In addition, electron microscopy enables visualization of Scc protofilaments with a width of about 6 to 10 nm. By electron microscopy, eukaryotic intermediate filaments and CreS exhibit diameters ranging between 7 and 11 nm (3, 10). By analogy to eukaryotic intermediate filaments, the first step in Scc oligomerization may be the formation of dimers by coiled-coil interactions. Electron microscopy suggests that the helices of Scc can rope together by multiple interactions. However, a hallmark feature of proteins from eukaryotic intermediate filaments is their insolubility in buffers of physiological ionic strength and pH (10), which is not the case of Scc.

Microscopic observations of Scc-GFP in *E. coli* revealed filaments that were similar to the helical structures of ParM (21) or CreS (3). In rod-shaped bacteria, actin-like proteins such as MreB and ParM form helical filaments just beneath the bacterial cell surface that help to determine cell shape (MreB) and contribute to DNA segregation (ParM). In addition, filamentous structures found underneath the cell membrane in

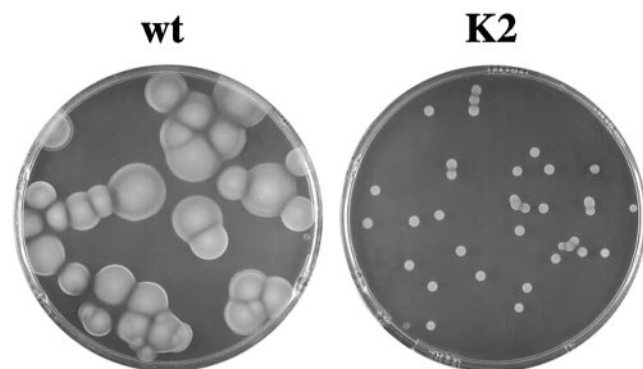


FIG. 8. Phenotype of the *L. biflexa* *scc* mutant. Growth of *L. biflexa* wild-type strain (wt) and *sccI* mutant strain (K2) on solid medium. Bacterial cells were spread into EMJH plates and incubated 15 days at 30°C.

helical mollicutes could be involved in the movement of the cells by length changes of the filaments (14). We showed that there is no role of Scc in the maintenance of the spiral cell-shape of *L. biflexa*, as in the case of the *C. crescentus* CreS protein, or in cell movement, as it seems to be the case of filaments in mollicutes (14). However, failure to swarm on solid medium may indicate that *L. biflexa* scc mutants are nonchemotactic mutants (2).

Rod-shaped structures, resembling those of tobacco mosaic virus, were observed only when DNA or RNA was added to Scc protein samples. The assembly of the rod-shaped tobacco mosaic virus, 300 nm long and 18 nm in diameter, occurs by the threading of the RNA through the central hole of a growing rod consisting of a 17-kDa capsid protein (26). However, no specific diffraction patterns were obtained by optical diffraction of the electron micrographs of Scc nucleoprotein structures (data not shown).

In bacteria and other microorganisms, genes corresponding to related functions are often grouped in large transcription units, known as operons. In *Leptospira* spp., the scc gene was found colocalized with the *infC-rpmI-rplT* operon, encoding the translation IF3 protein and the two ribosomal proteins L35 and L20. In *E. coli*, the ribosome is a ribonucleoprotein complex composed of two subunits. The 50S large subunit is formed by two RNAs (23S and 5S) and 33 proteins (L1 to L36), whereas the small 30S subunit consists of one RNA (16S) and 21 proteins (S1 to S21). Previous studies had shown that the pattern of transcription in the *E. coli* *infC-rpmI-rplT* operon is highly complex (15). A compact organization of these genes with the downstream genes, including scc, may have regulatory advantages. Immunoblotting with polyclonal antibodies raised against Scc shows evidence that *L. biflexa* Scc is produced by a frameshift, which might represent an additional means of controlling gene expression under specific conditions. Since the expression of scc is likely regulated by a complex regulatory mechanism, it may prevent the complementation in *trans* of the *L. biflexa* scc mutant. At the protein level, the Scc copy number and/or the mechanism of assembly of a putative protein complex may also prevent complementation in *trans*. A hypothetical role of Scc in the translation process would be consistent with its ability to bind nucleic acids and the genetic link between scc and the *infC-rpmI-rplT* operon. Taken together, these results suggest that Scc may form a ribonucleoprotein complex that could be involved in translation; this complex might be a component of the translation machinery or could be involved for example in tracking the ribosomes. In previous studies, ribosome-like and/or polysome-like particles were localized attached to intracellular protofilaments, primarily composed of the elongation factor-Tu protein, in the inner face of the cytoplasmic membranes of bacterial cells (19). Scc could also play a role in the protection of nucleic acids from degradation.

Further studies are required to understand the structure and the mechanism of assembly of Scc, as well as its precise physiological role. After the description of crescentin, our study opens up the possibility that intermediate filament-like systems may be present in an important number of prokaryotes. The determination of the functional and structural similarities between these systems will no doubt be of fundamental importance.

ACKNOWLEDGMENTS

We thank T. Blisnick and M. C. Prevost for technical help with antiserum preparation and electron microscopy, respectively.

K.M. is funded by the Cantarini Grant and the Institut Pasteur.

REFERENCES

- Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. Basic local alignment search tool. *J. Mol. Biol.* **215**:403–410.
- Armstrong, J. B., J. Adler, and M. M. Dahl. 1967. Nonchemotactic mutants of *Escherichia coli*. *J. Bacteriol.* **93**:390–398.
- Ausmees, N., J. R. Kuhn, and C. Jacobs-Wagner. 2003. The bacterial cytoskeleton: an intermediate filament-like function in cell shape. *Cell* **115**:705–713.
- Bermudes, D., G. Hinkle, and L. Margulis. 1994. Do prokaryotes contain microtubules? *Microbiol. Rev.* **58**:387–400.
- Bourhy, P., L. Frangeul, E. Couve, P. Glaser, I. Saint Girons, and M. Picardeau. 2005. Complete nucleotide sequence of the LE1 prophage from the spirochete *Leptospira biflexa* and characterization of its replication and partition functions. *J. Bacteriol.* **187**:3931–3940.
- Bravo, A., and M. Salas. 1998. Polymerization of bacteriophage phi 29 replication protein p1 into protofilament sheets. *EMBO J.* **17**:6096–6105.
- Crick, F. H. C. 1953. The packing of α -helices: simple coiled-coils. *Acta Crystallogr.* **6**:689–697.
- Ellinghausen, H. C., and W. G. McCullough. 1965. Nutrition of *Leptospira pomona* and growth of 13 other serotypes: fractionation of oleic albumin complex and a medium of bovine albumin and polysorbate 80. *Am. J. Vet. Res.* **26**:45–51.
- Faine, S., B. Adler, C. Bolin, and P. Perolat. 1999. *Leptospira* and leptospirosis. MedScience, Melbourne, Australia.
- Herrmann, H., and U. Aebi. 2004. Intermediate filaments: molecular structure, assembly mechanism, and integration into functionally distinct intracellular scaffolds. *Annu. Rev. Biochem.* **73**:749–789.
- Hofmann, K., and W. Stoffel. 1993. TMbase-A database of membrane spanning proteins segments. *Biol. Chem. Hoppe-Seyler* **374**:166.
- Izard, J., W. A. Samsonoff, and R. J. Limberger. 2001. Cytoplasmic filament-deficient mutant of *Treponema denticola* has pleiotropic defects. *J. Bacteriol.* **183**:1078–1084.
- Johnson, R. C., and V. G. Harris. 1967. Differentiation of pathogenic and saprophytic leptospires. *J. Bacteriol.* **94**:27–31.
- Kurner, J., A. S. Frangakis, and W. Baumeister. 2005. Cryo-electron tomography reveals the cytoskeleton structure of *Spiroplasma melliferum*. *Science* **307**:436–438.
- Lesage, P., H. N. Truong, M. Graffe, J. Dondon, and M. Springer. 1990. Translated translational operator in *Escherichia coli*. Auto-regulation in the *infC-rpmI-rplT* operon. *J. Mol. Biol.* **213**:465–475.
- Letunic, I., R. R. Copley, S. Schmidt, F. D. Ciccarelli, T. Doerks, J. Schultz, C. P. Ponting, and P. Bork. 2004. SMART 4.0: towards genomic data integration. *Nucleic Acids Res.* **32**:D142–D144.
- Lewis, P. 2004. Bacterial subcellular architecture: recent advances and future prospects. *Mol. Microbiol.* **54**:1135–1150.
- Lupas, A., M. Van Dyke, and J. Stock. 1991. Predicting coiled coils from protein sequences. *Science* **252**:1162–1164.
- Mayer, F. 2003. Cytoskeletons in prokaryotes. *Cell Biol. Int.* **27**:429–438.
- Mazouni, K., F. Domain, C. Cassier-Chauvat, and F. Chauvat. 2004. Molecular analysis of the key cytokinetic components of cyanobacteria: FtsZ, ZipN and MinCDE. *Mol. Microbiol.* **52**:1145–1158.
- Moller-Jensen, J., R. B. Jensen, J. Lowe, and K. Gerdes. 2002. Prokaryotic DNA segregation by an actin-like filament. *EMBO J.* **21**:3119–3127.
- Pauling, L., and R. B. Corey. 1953. Compound helical configurations of polypeptide chains: structure of proteins of the alpha-keratin type. *Nature* **171**:59–61.
- Picardeau, M., A. Brenot, and I. Saint Girons. 2001. First evidence for gene replacement in *Leptospira* spp. Inactivation of *L. biflexa* *flaB* results in non-motile mutants deficient in endoflagella. *Mol. Microbiol.* **40**:189–199.
- Ren, S., G. Fu, X. Jiang, R. Zeng, H. Xiong, G. Lu, H. Q. Jiang, Y. Miao, H. Xu, Y. Zhang, X. Guo, Y. Shen, B. Q. Qiang, X. Q., A. Danchin, I. Saint Girons, R. L. Somerville, Y. M. Weng, M. Shi, Z. Chen, J. G. Xu, and G. P. Zhao. 2003. Unique and physiological and pathogenic features of *Leptospira interrogans* revealed by whole genome sequencing. *Nature* **422**:888–893.
- Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. *Molecular cloning: a laboratory manual*, 2nd ed. Cold Spring Harbor Press, Plainview, N.Y.
- Scholthof, K. B. 2004. Tobacco mosaic virus: a model system for plant biology. *Annu. Rev. Phytopathol.* **42**:13–34.
- Yang, R., S. Bartle, R. Otto, A. Stassinopoulos, M. Rogers, L. Plamann, and P. Hartzell. 2004. AglZ is a filament-forming coiled-coil protein required for adventurous gliding motility of *Myxococcus xanthus*. *J. Bacteriol.* **186**:6168–6178.
- You, Y., S. Elmore, L. L. Colton, C. Mackenzie, J. K. Stoops, G. M. Weinstein, and S. J. Norris. 1996. Characterization of the cytoplasmic filament protein gene (*cfpA*) of *Treponema pallidum* subsp. pallidum. *J. Bacteriol.* **178**:3177–3187.